Preventing E-cadherin aberrant *N*-glycosylation at Asn-554 improves its critical function in gastric cancer

S Carvalho[1,2], TA Catarino[1], AM Dias[1,2], M Kato[3], A Almeida[4,5], B Hessling[6], J Figueiredo[1], F Gärtner[1,2], JM Sanches[7], T Ruppert[6], E Miyoshi[8], M Pierce[9], F Carneiro[1,10,11], D Kolarich[4], R Seruca [1,10], Y Yamaguchi[3], N Taniguchi[3], CA Reis [1,2,10] and SS Pinho [1,2]

[1]Instituto de Investigação e Inovação em Saúde (Institute for Research and Innovation in Health), University of Porto, Portugal/Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal;

[2]Institute of Biomedical Sciences of Abel Salazar (ICBAS), University of Porto, Porto, Portugal;

[3]Systems Glycobiology Research Group, RIKEN-Max Planck Joint Research Center, RIKEN Global Research Cluster, Wako, Saitama, Japan;

[4]Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, Potsdam, Germany;

[5]Freie Universität Berlin, Institute of Chemistry and Biochemistry, Berlin, Germany;

[6]Center for Molecular Biology, University of Heidelberg, Heidelberg, Germany;

[7]Institute for Systems and Robotics (ISR/IST), LARSyS, Instituto Superior Técnico, University of Lisbon, Lisboa, Portugal;

[8]Molecular Biochemistry and Clinical Investigation, Osaka University Graduate School of Medicine, Osaka, Japan;

[9]Complex Carbohydrate Research Center, Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, USA;

[10]Medical Faculty, University of Porto, Porto, Portugal and [11]Department of Pathology, Hospital S. Joao, Porto, Portugal.

Correspondence: Dr SS Pinho, Glycobiology in Cancer group, Institute of Molecular Pathology and Immunology, University of Porto, Rua Doutor Roberto Frias, Porto 4200-465, Portugal.

E-mail: salomep@ipatimup.pt

Received 14 December 2014; revised 15 May 2015; accepted 18 May 2015

Originally published in Oncogene. 2016 Mar 31;35(13):1619-31. doi: 10.1038/onc.2015.225.

ABSTRACT

E-cadherin is a central molecule in the process of gastric carcinogenesis and its posttranslational modifications by N-glycosylation have been described to induce a deleterious effect on cell adhesion associated with tumor cell invasion. However, the role that site- specific glycosylation of E-cadherin has in its defective function in gastric cancer cells needs to be determined. Using transgenic mice models and human clinical samples, we demonstrated that N-acetylglucosaminyltransferase V (GnT-V)-mediated glycosylation causes an abnormal pattern of E-cadherin expression in the gastric mucosa. In vitro models further indicated that, among the four potential N-glycosylation sites of E-cadherin, Asn-554 is the key site that is selectively modified with β1,6 GlcNAc-branched N-glycans catalyzed by GnT-V. This aberrant glycan modification on this specific asparagine site of E-cadherin was demonstrated to affect its critical functions in gastric cancer cells by affecting E-cadherin cellular localization, cis-dimer formation, molecular assembly and stability of the adherens junctions and cell–cell aggregation, which was further observed in human gastric carcinomas. Interestingly, manipulating this site-specific glycosylation, by preventing Asn-554 from receiving the deleterious branched structures, either by a mutation or by silencing GnT-V, resulted in a protective effect on E-cadherin, precluding its functional dysregulation and contributing to tumor suppression.

Oncogene advance online publication, 20 July 2015; doi:10.1038/onc.2015.225

INTRODUCTION

Epithelial cadherin (E-cadherin) is a calcium-dependent cell–cell adhesion molecule with a critical role in epithelial tissue morphogenesis.[1–7] The mature E-cadherin protein is organized into three major structural domains: an ectodomain of about 550 amino acids comprised of five tandemly repeated subdomains (EC1–EC5), a single transmembrane domain and a short cytoplasmic domain of about 150 amino acids. The human E-cadherin ectodomain is comprised of four potential *N*-glycosylation sites: two putative sites located in the EC4 subdomain (Asn-554 and Asn-566) and the remaining two potential sites in the EC5 subdomain (Asn-618 and Asn-633).[6]

Epithelial cell–cell adhesion is achieved through homophilic interactions between cadherin molecules of neighboring cells. The extracellular domains of E-cadherins on the same cell surface establish lateral or *cis*- interactions and then among adjacent cells to form *trans* adhesive bonds, leading to the formation of a zipper-like structure.[8–10] Cell–cell adhesion is further accomplished through the molecular interaction of the cytoplasmic domain of E-cadherin and the catenins β-, γ-, α- and p120-catenin. The stability of this cadherin–catenin complex, and its interaction with the actin cytoskeleton, forms the core adherens junctions, which is essential for cell–cell adhesion and precludes individual epithelial cell motility, thus providing a normal and homeostatic tissue architecture.[11,12]

The functional inactivation or downregulation of E-cadherin is considered to be a hallmark of the epithelial carcinogenic process, being closely associated with tumor cell invasion and metastases.[13] Several mechanisms have been proposed to explain the loss of function of E-cadherin in cancer,[7] and among those mechanisms, the modification of E-cadherin by glycosylation has been demonstrated to be instrumental for the regulation of E-cadherin functions in the context of cancer.[14,15] Moreover, O-mannosylation of E-cadherin was recently demonstrated to be important for E-cadherin-mediated cell–cell adhesion.[16,17]

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

In fact, during malignant transformation, the glycosylation profile of E-cadherin undergoes a significant alteration,[18] with implications for its biological functions. Of those glycosylation alterations, two are generally considered to be fundamental for the regulation of the protein: the bisecting GlcNAc *N*-glycan structures (catalyzed by the *N*-acetylglucosaminyltransferase III (GnT-III) glycosyltransferase), and the β1,6 GlcNAc-branched *N*-glycan structure (catalyzed by the GnT-V glycosyltransferase). These two key glycan structures were recently found to precisely control, in an opposite manner, the functions of E-cadherin in cancer cells.[15,19–24] E-cadherin has different *N*-glycosylation sites that are potentially responsible for controlling its biological functions. Under normal physiological circumstances, E-cadherin was described to be primarily modified with high-mannose/hybrid *N*-glycans;[25,26] however, the key site(s) that are crucial for the regulation of its functions in the context of cancer remain to be determined.

In this study, we report that E-cadherin becomes functionally impaired when GnT-V glycosyltransferase catalyzed modification of site 1 (Asn-554) with β1,6 GlcNAc-branched *N*-glycan structures. Further, the presence of a mutation of this specific *N*-glycosylation site (which abrogates the availability of this site to be modified by GnT-V) has a protective effect on E-cadherin function, preventing its dysregulation.

RESULTS

Impact of GnT-V-mediated glycosylation in E-cadherin expression in the mice stomach and in human gastric carcinoma

In the present study, we used GnT-V transgenic mouse models to show that differential expression of E-cadherin in the gastric mucosa is influenced by the presence of GnT-V. We used two transgenic mouse models: *MGAT5* knockout (with no GnT-V activity) and *MGAT5* transgenic mice that overexpress GnT-V activity. We observed that the gastric mucosa of the wild-type (WT) mice showed a moderate expression of β1,6 GlcNAc-branched structures detected by the *biotinylated Phaseolus vulgaris leucoagglutinin* (L-PHA) lectin. The gastric mucosa of the *MGAT5* knockout mice showed no L-PHA reactivity (Figure 1a). On the other hand, mice overexpressing GnT-V showed high levels of expression of β1,6 GlcNAc-branched structures in the gastric mucosa. Regarding E-cadherin localization, *MGAT5* knockout mice showed E-cadherin normally expressed in the basolateral cell membrane of gastric epithelial cells. In GnT-V-overexpressing mice, E-cadherin was also displayed in the cytoplasm of cells at the neck zone and in deep glands of the gastric mucosa (Figure 1a). No major histopathological lesions were observed in the gastric mucosa of these mouse models.

Similarly, the levels and profile of E-cadherin modifications with the aberrant β1,6-branched *N*-glycans (catalyzed by GnT-V) in human gastric carcinomas (diffuse gastric cancer sub-type) and the correlation with prognostic features were demonstrated by performing *in situ* Proximity Ligation Assay (PLA) technique (Figure 1b). PLA is a technology that extends the capabilities of traditional immunoassays to include direct detection of proteins, protein interactions and posttranslational modifications such as glycosylation with high specificity and sensitivity.[27] Our results showed that gastric carcinoma cells display a marked increase of positive PLA signals demonstrating an increased modification of E-cadherin with β1,6-branching *N*-glycans. The normal gastric mucosa was either negative or showed a weak PLA signal, suggesting that E-cadherin is not modified by GnT-V-mediated branching glycans in normal gastric cells (Figure 1b). Furthermore, and in order to evaluate whether the GnT-V-mediated glycosylation of E-cadherin is correlated with prognosis, a series of 19 patients with diffuse gastric cancer was analyzed and associated with prognostic variables. Patients positive for E-cadherin+β1, 6-branched *N*-glycans (positive PLA) display a poorer

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

survival rate (assessed by Kaplan–Meier survival analysis) when compared with diffuse gastric cancer patients negative for PLA E-cadherin/L-PHA (P = 0.093; Figure 1, Graphic b1). Additionally, we further showed that all patients who do not survive were positive for PLA E-cadherin/ L-PHA, as represented in Table 1 (P = 0.077).

These results demonstrate the importance of this specific E-cadherin aberrant glycoform *in vivo* in the pathogenesis of gastric carcinoma.

Bioinformatics evaluation of E-cadherin N-glycosylation site occupancy

Glycosylation is known to be a cell- and tissue-specific event and the levels and pattern of glycosylation of a specific protein vary accordingly with cell and tissue location[28] (Supplementary Figure S1). *N*-linked glycosylation does not occur at every potential glycosylation site. Depending on the protein and on the physiopathological context, some glycosylation sites are more important than others for the regulation of protein function.[29]

Using an *in silico* bioinformatics analysis that takes three main criteria into consideration, E-cadherin *N*-glycosylation site occupancy could be predicted. The first criterion is based on the species homology of the E-cadherin peptide sequence (Figure 2a). A comparison of the E-cadherin peptide sequence between different species revealed that the *N*-glycosylation sites Asn-554 (site 1) and Asn-633 (site 4) are the most conserved among species.

The second criterion used in the evaluation is the site-specific modification of *N*-linked glycans through the solvent accessible surface area (ASA) calculation of four Asn residues. Very low ASA values at Asn residue indicate a reduced probability of *N*-glycans attachment. As shown in Figure 2b, the potential *N*-glycosylation site Asn-618 (site 3) displays a low ASA value, suggesting a low probability of being *N*-glycosylated.

The third criterion is the evaluation of the β-turn propensity at the Asn-x-Ser/Thr acceptor sequence (Figure 2c). The variable *x* represents the sequence position around the sequon and a high value of β-turn propensity at an Asn residue can positively affect the spatial configuration of Asn and Ser/Thr side chains towards the approach of an oligosaccharyltransferase. For *x*=0 (Asn residue), Asn-554 (site 1) and Asn-633 (site 4) show higher β-turn propensity values demonstrating a high probability for the occurrence of *N*-glycosylation. It has also been reported that the presence of Pro at *x*=3 position inhibits *N*-glycosylation.[30,31] Taken these issues into account, Asn-566 (site 2) and Asn-618 (site 3) are not likely to be glycosylated.

Overall, the combination of these *in silico* parameters (Figure 2d) suggests that *N*-glycosylation is likely to occur at Asn-554 (site 1) and Asn-633 (site 4) and not likely to occur in the other two putative sites (sites 2 and 3).

E-cadherin Asn-554 (site 1) and Asn-633 (site 4) are occupied with complex- and high-mannose/hybrid-type N-glycans, respectively

In order to further characterize *in vitro* the E-cadherin *N*-glycosylation site occupancy, we applied site-directed mutagenesis to generate different E-cadherin *N*-glycan mutants lacking selected potential *N*-glycosylation sites by replacing asparagine for glutamine residues in each *N*-glycosylation consensus sequence (NXS/T) either individually or in different combinations (Figure 3a).

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

The different E-cadherin *N*-glycan mutants were transfected into gastric carcinoma AGS cell line (that lacks endogenous E-cadherin both at mRNA and protein levels[32]), and the mobility shift of the E-cadherin band was evaluated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE; Figure 3b). Concerning the individual E-cadherin N-glycan mutants (M1, M2, M3 and M4), the results showed that the E-cadherin mutant M1 (lacking site 1, Asn-554) revealed the most pronounced mobility shift followed by the E-cadherin mutant M4 that displayed a slight mobility compared with E-cadherin WT. E-cadherin mutants M2 and M3 did not exhibit a mobility shift compared with the WT. These results suggest that sites 1 (Asn-554) and 4 (Asn-633) are N-glycosylated. In addition, and consistent with the bioinformatics prediction, site 2 (Asn-566) and site 3 (Asn-618) are not occupied with *N*-glycans. Validating these observations, the multiple E-cadherin *N*-glycan mutants M1234 (total *N*-glycan mutant) and M14 (with Asn-554 and Asn-633 not present) showed the same mobility shift that was slightly higher than that for E-cadherin M1. E-cadherin M23 did not display a mobility shift compared with E-cadherin WT, further suggesting that sites 2 and 3 are not *N*-glycosylated.

Next we determined the susceptibility of the E-cadherin *N*-glycan mutants towards different endoglycosidases, namely *PNGase F* and *Endo H*. As shown in Figure 3c, E-cadherin WT showed partial sensitivity to *Endo H* but, being predominantly *PNGase F*-sensitive, indicating that E-cadherin WT is modified with both complex and high-mannose/hybrid-type N-glycans. The E-cadherin mutant M234 showed a slight shift upon *Endo H* but exhibited a significant mobility shift after *PNGase F* digestion. On the other hand, treatment of M123 with *Endo H* resulted in a substantial mobility shift that was not further enhanced by *PNGase F*. These observations indicate that Asn-554 is modified mostly by complex-type structures, whereas Asn-633 appears to contain mainly high-mannose/hybrid-type *N*-glycans. This is also consistent with the higher mobility shift of M1 compared with M4 (Figure 3b).

These observations were further supported by a glycomics analysis of the released *N*-glycans of E-cadherin WT, M123 and M234 by porous graphitized carbon (PGC) nano-liquid chromatography–electronspray ionization tandem mass spectrometry (nanoLC-ESI MS/MS). The respective E-cadherin proteins were immunoprecipitated, SDS–PAGE separated and electro-blotted onto polyvinylidene difluoride (PVDF) membranes prior on-blot *PNGase F* digestion as described earlier.[33,34] E-cadherin WT and M234 showed higher levels of sialylated and complex type N-glycans, respectively (Figure 3d). E-cadherin WT showed the highest levels of complex type, triantennary *N*-glycans (Supplementary Figure S2, Supplementary Tables S1 and S2). These structures were found reduced in M123 and could not be detected in M234 (Supplementary Figure S2), further indicating the presence of tetra-antennary *N*-glycans in M234 (though these analyses do not allow an absolute glycan quantitation).

Regarding M134, M124 and M14 (all with a mutation at sites 1 and 4), no detectable changes were observed after *Endo H/PNGase F* treatment, supporting the hypothesis that Asn-566 (site 2) and Asn-618 (site 3) are not *N*-glycosylated. As expected, E-cadherin with all potential *N*-glycosylation sites mutated (M1234) was neither *Endo H* nor *PNGase F* sensitive. Taken together (see also Supplementary Figure S3), these results indicate that, in gastric cancer cells, E-cadherin is modified with mostly complex-type *N*-glycans at site 1 (Asn-554) and with mainly high-mannose/hybrid-type at site 4 (Asn-633), whereas sites 2 and 3 (Asn-566 and Asn-618) were not *N*-glycosylated. A statistical analysis suggests that glycosylation sites carrying highly processed complex-type glycans are significantly more solvent-accessible than those carrying less processed high-mannose-type glycans.[35] Consistent with this, ASA of Asn-633 is lower than that of Asn-554 (Figure 2b).

Site-specific occupancy of Asn-554 with complex type N-glycans has a deleterious effect on E-cadherin localization and dimerization

The glycosylation of E-cadherin with complex-type *N*-glycans, such as bisecting GlcNAc and the β1,6 GlcNAc-branched *N*-glycans structures, has been reported to precisely regulate E-cadherin functions in an opposing manner.[15,20] Having confirmed that site 1 (Asn-554) of E-cadherin is the one that is mainly occupied with complex-type *N*-glycans (Figure 3, Supplementary Figure S2), we proceeded to further evaluate the functional impact of this site-specific *N*-glycan occupancy by comparing the mutants M1 and M234, representing the absence versus the presence of a complex *N*-glycan on site 1, respectively. By immunofluorescence analysis, we evaluated the cellular localization of E-cadherin and cell morphology, and the results showed that AGS cells expressing the E-cadherin M1 mutant (Figure 4a) or M123 (Supplementary Figure S4A) exhibited a membrane localization of E-cadherin that showed a more focused membrane staining than the E-cadherin WT (as revealed by confocal microscopy). By contrast, AGS cells expressing E-cadherin M234 (or E-cadherin M4, Supplementary Figure S4A) displayed an aberrant E-cadherin expression with mis-localization into the cytoplasm. Quantitative analyses of *in situ* immunofluorescence images that consider internuclear profiles further showed a focused membrane localization of E-cadherin after mutation at Asn-554 (E-cadherin M1) and a decreased E-cadherin membrane expression in M234 (Figure 4A1). Regarding the E-cadherin M2, M3 and M23 mutants, no alterations were observed in the cellular localization of E-cadherin compared with E-cadherin WT (Supplementary Figure S4A).

We then attempted to determine whether *cis*-dimer formation of E-cadherin was affected by the occupancy/mutation of site 1 with complex-type *N*-glycans, by performing a *cis*-dimerization assay using BS^3, a membrane-impermeable chemical cross-linker. The formation of E-cadherin *cis*-dimers was observed after AGS cells expressing E-cadherin WT, M1 or M234 were treated with BS^3 (Figure 4b). However, the levels of E-cadherin *cis*-dimerization were significantly higher in E-cadherin M1 than in E-cadherin WT and M234 (Figure 4B1) or than in M4 (Supplementary Figures S4B and B1), indicating that the mutation of Asn-554 (M1) enhanced the formation of E-cadherin ^{cis}-dimers.

In addition, a calcium switch assay was performed (Supplementary Figure S5). These results showed that the calcium-binding property appear to be independent of E-cadherin glycosylation; however, it is observed that E-cadherin M234 showed a more evident impact in the calcium-binding property of E-cadherin.

Consistent with this, the deletion of site 1 (M1) induced a significantly increased rate of cell–cell adhesion together with the formation of larger compact cellular aggregates when compared with E-cadherin WT and M234 (Figures 4c and c1; Supplementary Figure S4C).

Site-specific occupancy of Asn-554 with complex-type N-glycans impairs a competent adhesion complex formation

We further evaluated whether the mutation of Asn-554 and/or the occupancy of this specific *N*-glycan site could affect the molecular organization and assembly of the cadherin–catenin complex. No significant differences were observed in the total expression levels of the catenins, β- and p120-catenin, among the different E-cadherin *N*-glycan mutants (Figure 4d). However, an increased E-cadherin/β-catenin interaction was verified in E-cadherin M1 in comparison with E-cadherin WT (0.7-fold) and a significantly decreased interaction was found between E-cadherin and β-catenin when Asn-554 (site 1) was *N*-glycosylated (M234 or M4, Figures 4e and e1; Supplementary Figures 4D and D1). In addition, we verified a 1.3- and 0.9-fold increased interaction between E-cadherin and p-120-catenin when Asn-554 (site 1) was mutated (M1) when compared with E-cadherin WT

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

and E-cadherin M234, respectively, suggesting that mutation of site 1 resulted in an increased recruitment of catenins by E-cadherin (Figures 4e and e1). These differences in the recruitment of catenins by E-cadherin reflect the positive impact on the assembly and stability of the cadherin–catenin complex induced by the mutagenesis of Asn-554 (site 1).

Preventing Asn-554 site-specific glycosylation improves E-cadherin functions

Having demonstrated the deleterious effects of site 1 *N*-glycan occupancy for the biological functions of E-cadherin in gastric tumor cells (Figure 4), we further identified the specific type of complex *N*-glycans that are attached to Asn-554. Figure 5a demonstrates that E-cadherin M234 showed an increased reactivity with the L-PHA lectin compared with M1, which is indicative of the presence of β1,6 GlcNAc-branched structures at Asn-554 (site 1). Both E-cadherin WT and M234 show similar β1,6 GlcNAc-branched expression patterns, as both retain the Asn-554 complex glycan. In contrast, no positive reactivity with L-PHA on the E-cadherin band in the M1 mutant was observed, showing that site 1 is specifically modified with β1,6 GlcNAc-branched structures. Despite that it was not possible to unambiguously substantiate the β1,6 GlcNAc nature of the branching in M234 with the minimal amounts of E-cadherin material available for glycomic analyses, the triantennary *N*-glycans were only detected in E-cadherin WT and M123 and not in M234. Nevertheless, the combined data presented here clearly indicate that also a subtle increase of a specific structure feature on E-cadherin induces a strong functional response.

To confirm these results, we performed silencing of the *MGAT5* gene using the small interfering RNA (siRNA) technique (Figure 5b). As shown in Figure 5c, the knockdown of GnT-V resulted in a significant decrease in the expression of β1,6 GlcNAc-branched structures (decreased L-PHA reactivity) in total cell lysates compared with non-silenced (NS) cells. The decreased L-PHA reactivity was also observed in total cell lysates of AGS Mock siRNA cells when compared with NS cells (Supplementary Figure S6). As a consequence of the GnT-V knockdown, we observed a competitive increase in the expression of bisecting GlcNAc structures (increased *biotinylated Phaseolus vulgaris erythroagglutinin* (E-PHA) reactivity) in the total cell lysates of the GnT-V knockdown cells compared with NS cells, as previously described. The silencing of *MGAT5* led to a decrease in the reactivity of E-cadherin M234 with the L-PHA lectin compared with the control (NS), which further supports the conclusion that site 1 is specifically modified with β1,6 GlcNAc-branched structures catalyzed by GnT-V enzymatic activity (Figure 5d). Furthermore, our results showed that, upon *MGAT5* knockdown, the gastric tumor cells expressing E-cadherin M234+siGnT-V displayed a notable change of its pattern of expression, showing a decreased cytoplasmic staining and an increase in membrane expression compared with NS cells (Figure 5e). In addition, we also demonstrated that, after the knockdown of GnT-V, the ratio of E-cadherin cis-dimers in gastric tumor cells expressing E-cadherin M234+siGnT-V increased to levels similar to that for E-cadherin M1 (Figures 5f and f1), suggesting that the inhibition of GnT-V- mediated modification with β1,6 GlcNAc-branched structures at Asn-554 (site 1) induces increased E-cadherin cis-dimer formation and, consequently, the strengthening of cell–cell adhesion.

These results reveal that the GnT-V selectively modifies N-glycosylation site 1 (Asn-554) on E-cadherin in the context of gastric tumor cells and that its silencing clearly increases E-cadherin-mediated cell–cell adhesion and expression.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

DISCUSSION

In the present study, using two different *MGAT5* mouse models, we showed that the pattern of E-cadherin expression in gastric mucosa cells depends on the presence of β1,6 GlcNAc-branched N-glycans that are formed as the result of enzymatic catalysis by GnT-V (*MGAT5*). *MGAT5* transgenic mice (which overexpress GnT- V-mediated glycosylation) display an abnormal pattern of E-cadherin cellular expression in gastric epithelial cells that overexpress β1,6 GlcNAc-branched *N*-glycans compared with *MGAT5* knockout mice (with absence GnT-V-mediated enzymatic activity). Additionally, we verified that human diffuse gastric carcinoma is characterized by an aberrant pattern of E-cadherin *N*-glycosylation mediated by GnT-V. Moreover, we observed that diffuse gastric cancer patients displaying an aberrant glycosylation of E-cadherin mediated by GnT-V tend to exhibit a poorer survival rate when compared with gastric cancer patients without GnT-V-mediated glycosylation on E-cadherin.

We further investigated the specific *N*-glycosylation site(s) that are required for controlling E-cadherin functions in gastric cancer, particularly the key E-cadherin *N*-glycosylation site that can be modified with the deleterious β1,6 GlcNAc-branched structures catalyzed by the action of the GnT-V glycosyltransferase. Through combination of bioinformatics prediction tools with mutagenesis and glycans enzymatic digestion approaches as well as glycomic analyses, our results consistently revealed that, among the four potential *N*-glycosylation sites available in the extracellular domain of E-cadherin, Asn-554 (site 1) and Asn-633 (site 4) are N-glycosylated, respectively, with complex-type and high-mannose *N*-glycans structures. Conversely, Asn-566 (site 2) and Asn-618 (site 3) of E-cadherin are not occupied with N-glycans, suggesting that these N-glycosylation sites are not critical for controlling the biological functions of E-cadherin in gastric tumor cells. Consistent with our observations, previous reports have shown that E-cadherin from a CHO cell line is modified with *N*-glycans at Asn-554 and Asn-633[ref. 36] and that Asn-633 is modified with high-mannose *N*-glycans that were demonstrated to be important for the folding of E-cadherin.14 In addition, our glycomics data showed that E-cadherin WT carries triantennary, sialylated N-glycans, whereas the glycosylation was found to be altered on the M123 and M234 mutants. These data indicate that site-specific mutations of the *N*-glycosylation sites on E-cadherin appear to alter the protein conformation and thus also the accessibility for the glycosyltransferases required for the formation of described *N*-glycans.

We clearly demonstrated that site 1 (Asn-554) in the extracellular domain of E-cadherin is the selected site modified with the β1,6 GlcNAc-branched structures and the key site for the functional regulation of E-cadherin in gastric tumor cells. Whenever site 1 (Asn-554) is modified with β1,6 GlcNAc- branched structures, the localization and functions of E-cadherin are impaired (Figure 6) by: compromising its cis-dimerization capability, by decreasing the rate of cell–cell aggregation, and also by interfering in the molecular assembly and stability of the adhesion complex, that altogether promotes its cellular mis-localization and non-functional role. In fact, a mutation of this specific site 1 (E-cadherin M1 or M123) resulted in a protective effect, precluding the functional impairment of E-cadherin through the recovery of E-cadherin-mediated cell–cell adhesion, an increased stability of adherens junctions and the correct membrane localization of the molecule. Similarly, the silencing of the *MGAT5* gene that encodes GnT-V also resulted in a restoration of E-cadherin expression and functions. Although the GnT-V knockdown is likely to affect various cell surface receptors, in this study we focused on E-cadherin, showing that upon GnT-V knockdown, the E-cadherin *N*-glycan mutant M234 (in which site 1 is available) loses the β1,6 GlcNAc-branched structures and consequently has increased cis-dimerization capability, increases recruitment of catenins and, importantly, it reaches the cell membrane, behaving similar to the M1 mutant. These results on M234 (and M4) mutant also suggest that loss of high-mannose glycan in site 4 may also

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

contribute to this process, as it was previously shown the importance of this site for E-cadherin folding.[14] Therefore, blocking Asn-554 (site 1), either by mutation or by GnT-V silencing, resulted in a protective effect on E-cadherin to not become aberrantly glycosylated thus avoiding its functional dysregulation (Figure 6). Taken together, these observations are in agreement with the results using transgenic mice regarding an E-cadherin phenotype where the *MGAT5* knockout resembles the M1 mutant and the M234+siGnT-V, whereas the GnT-V overexpressing mice are parallel with the M234 or M4 mutant. Furthermore and importantly from the clinical standpoint, we showed that gastric cancer patients with poor prognosis displayed a marked increased modification of E-cadherin with β1,6-branching *N*-glycans highlighting the importance of this specific E-cadherin aberrant glycoform in the gastric carcinogenic process.

Overall, this is the first study to clearly demonstrate that, among the four putative *N*-glycosylation sites, site 1 (Asn-554) in the EC4 domain of E-cadherin is the one that is selectively modified by the GnT-V glycosyltransferase in gastric carcinoma cells leading to its functional impairment. The molecular mechanism governing the specific biosynthesis of the β1,6 GlcNAc-branched structures at site 1 of E-cadherin, and not on any of the other three sites, remains unknown. Nevertheless, it is possible that, because this *N*-glycosylation site is the most extracellular (the furthest from the membrane), it is therefore more likely to regulate adhesion with neighboring cells. Additionally, the pathological context of gastric tumor cells may also account for this selective remodeling of E-cadherin *N*-glycosylation site 1 by GnT-V, which probably occurs depending on the physiopathological context and in a tissue- and cell-specific manner.[19,37–39]

Taken together, we discover how site-specific E-cadherin glycosylation modification can directly modulate E-cadherin-mediated cell–cell adhesion, a key pathophysiological event in cancer progression. Moreover, preventing E-cadherin site-specific glycosylation modifications was shown to improve its tumor-suppressive functions in gastric cancer. These results may open new avenues for the discovery of a promising (glyco)biomarker for improving early diagnosis, clinical surveillance and patients' risk stratification, as well as for the development of targeted-specific therapeutic strategies applied to the clinical setting.

MATERIALS

Tissue immunohistochemistry

Formalin-fixed paraffin-embedded tissue from normal, MGAT5 knockout (C57BL/6 background; generously provided by Professor Michael Pierce) and MGAT5 transgenic mouse stomach (C57BL/6 background; generously provided by Professor Eiji Miyoshi) were used for E-cadherin and L-PHA staining as previously described.[15]

In situ PLA

A series of human stomach samples from both normal mucosa (n = 4) and gastric carcinoma diffuse-type (n=19) were obtained from the S. João Centre Hospital, Porto, Portugal and subjected to *in situ* PLA.[27] This study was approved by the committee of S. João Centre Hospital. An informed consent was obtained for all the following subjects.

An antibody against E-cadherin (Mouse anti-E-cadherin, clone 4A2C7) and L-PHA biotinylated lectin against the β1,6 GlcNAc-branched *N*-glycan structure (Vector Laboratories, Burlingame, CA, USA) were used. The PLA probes used were, respectively, anti-mouse immunoglobulins PLUS and

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

anti-biotin-MINUS, conjugated to oligonucleotides. Images were acquired using brightfield and/or fluorescence microscope. The PLA signal was evaluated by three independent observers and scored as follows: negative, when absent/low expression of PLA signal was observed; and positive, when moderate/high expression of PLA signal was observed.

Prediction of N-glycosylation site occupancy

Amino-acid sequences of E-cadherin from various species are extracted from NCBI Reference Sequences[40] and selected by means of BLAST search.[41] Multiple top-hit sequences are aligned using Clustal W.[42] 3D homology modeling of E-cadherin was performed using the MODELLER version 9.8 software (San Francisco, CA, USA).[43] The crystal structures of the N-terminal domain of human E-cadherin (PDB ID 1o6s) and C-cadherin ectodomain (PDB ID 1l3w) were used as templates. The solvent ASA of the model was calculated using the DSSP program (San Franscisco, CA, USA).[44] A sliding window method was used to calculate the averaged ASA (window size=5), and the value was assigned to the central amino-acid residue. The β-turn propensity was predicted by the Chou Fasman method[45] using updated PDB data.[46]

Plasmids construction and site-directed mutagenesis

A plasmid pcDNA3-E-cad containing human full-length E-cadherin cDNA (GenBank accession no. Z13009) was used to produce the E-cadherin *N*-glycan site mutations singly and in different combinations. PCR-based site-directed mutagenesis was carried out using the QuickChange XL Site-directed Mutagenesis Kit (Stratagene, Santa Clara, CA, USA). The mutagenesis primers used to the replacement of asparagine for glutamine in the potential *N*-glycosylation sites (Asn-554, Asn-566, Asn-618 and Asn- 633) were created according to the manufacturer's instructions. All mutations were confirmed by DNA sequencing.

Release of N- and O-glycans and nanoPGC LC-ESI-MS/MS analysis

The *N*-glycans of immunoprecipitated E-cadherin WT and the respective mutants were released from PVDF membranes after SDS–PAGE separation and electro-blotting onto PVDF membranes to immobilize glycoproteins as described previously.[34,35] The released *N*-glycans were dissolved in 10 µl of pure water, and 5 µl were injected and analyzed by PGC nanoLC-ESI-MS/MS (Ultimate 3000 UHPLC system, Dionex, Sunnyvale, CA, USA) coupled to an amaZon ETD Speed ion trap equipped with a CaptiveSpray ionization source in negative ion mode (all Bruker Daltonics, Bremen, Germany). The instrument was set up to perform collision-induced dissociation fragmentation. An *m/z* range from 380 to 1850Da was used for data-dependent precursor scanning with the SPS target mass set to 900 *m/z*. The capillary voltage was set at 1.3 kV, and the three most intense signals were selected for collision-induced dissociation experiments (MS/MS scan range 50–2500*m/z*). MS as well as MS/MS data were recorded in the instrument's 'ultra scan mode'. Glycans were loaded onto a PGC precolumn (Hypercarb KAPPA 30 × 0.32 mm, 5 µm particle size) and separated on an analytical PGC column (Hypercarb PGC Column, Thermo, 100 mm × 75 µm, 3 µm particle size, both Thermo Fisher, Waltham, MA, USA). The samples were loaded onto the precolumn at a flow rate of 6 µl/min in 100% buffer A (10mM ammonium bicarbonate). The starting conditions for the analytical column were 2% buffer B (10 mM ammonium bicarbonate in 60% acetonitrile) at a flow rate of 0.8µl/min. The gradient conditions were as follows: increase of buffer B from 3% to 15.8% (6 min), further increase to 40.3% B (6–55 min), followed by a steeper increase to 90% B (55–60 min). The column

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

was held at 90% B for 6 min. At the same time, the precolumn was flushed with 90% Buffer C (10mM ammonium bicarbonate in 90% acetonitrile) at a flow rate of 6µl/min before re-equilibrating the precolumn as well as the analytical column in 98% buffer A for 5min. Glycans were automatically identified using an in-house established N-glycan spectral database and the spectral library tool integrated in Compass Data Analysis 4.1 (Bruker Daltonics). *N*-glycan structures not present in the database were assigned manually if sufficient data could be acquired. If no unambiguous identification of a structure was possible based on the MS/MS and LC data but just on MS data only, the respective portion of the N-glycan is labeled accordingly in the Supplementary Tables S1 and S2. Glycan compositions were determined using the GlycoMod tool available on the ExPASy server[47] (http://web. expasy.org/glycomod/; with a mass tolerance of 0.3 Da in the negative ion mode, reduced *N*-glycan oligosaccharides, respectively, and restricted to the monosaccharide residues hexose (Hex), *N*-acetyl hexosamine (HexNAc), deoxyhexose, *N*-acetyl neuraminic acid (NeuAc), *N*-glycolyl neuraminic acid (NeuGc) and sulfate.

The accessed glycan structures were then manually validated from tandem MS fragmentation spectra using the Glycoworkbench software tool (Eurocarb, Cambridge, UK).[48] Data reporting has been performed as proposed by the MIRAGE guidelines.[49,50]

Data processing and relative quantitation of PGC glycomics data

The peak area of each glycan composition and isoform was calculated using the respective extracted ion chromatograms using Compass QuantAnalysis (Bruker Daltonics). The integration of every extracted ion chromatogram was validated manually, in particular for peaks with very low abundance. The generated data were imported in CSV format, converted in Excel (Microsoft Office Excel 2013, Redmond, WA, USA) and expressed as a relative abundance value for each individual *N*-glycan structure, respectively, within each E-cadherin type. The *N*-glycans were classified into three categories (high mannose, neutral or sialylated).

Cell culture and transfection

MDCK (Madin Darby canine kidney), BT20 (breast tumor cell line), HT29 (human colon cancer cell line) and AGS cells (human adenocarcinoma epithelial cell line) were grown in monolayer culture and maintained at 37°C in an atmosphere of 5% CO_2 in RPMI 1640 GlutaMAX (or Dulbecco's modified Eagle's medium for BT20 cells), HEPES medium (Invitrogen, Paisley, UK) supplemented with 10% fetal bovine serum (Gibco, Invitrogen) and 1% penicillin–streptomycin (Gibco, Invitrogen). These cells were obtained from American Type Culture Collection (Manassas, VA, USA), which were tested and authenticated by the cell bank using their standard short tandem repeats–based techniques. Cells were also monitored by microscopy to maintain their original morphology and also tested for mycoplasma contamination.

Cells were transfected at 40–50% confluence with E-cadherin WT and E-cadherin *N*-glycan mutants using Lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions.

Westerns blotting, lectin blotting and immunoprecipitation

Cell protein lysates from different E-cadherin *N*-glycan mutants were subjected to 7.5% SDS–PAGE electrophoresis, transferred to nitrocellulose membranes and probed with the primary antibody against E-cadherin, β-catenin, p120-catenin and tubulin, as previously described.[15]

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

The expression of β1,6 GlcNAc-branched *N*-glycan structures and the bisecting GlcNAc *N*-glycans were detected by lectin blotting analysis, where membranes were incubated, respectively, with L-PHA (that specifically recognizes β1,6 GlcNAc-branched *N*-glycans, product of GnT-V) and E-PHA (that specifically recognizes bisecting GlcNAc *N*-glycans, product of GnT-III) lectins (1µg/ml; Vector Laboratories).[15] For the β1,6 GlcNAc-branched *N*-glycans analysis on E-cadherin immunoprecipitated, equal amounts of total protein (750–1500 µg) from each cell lysate were used and the membranes were probed with L-PHA lectin.[15]

PNGase F and Endo H digestion

Total cell lysates (30µg) were combined with denaturating buffer and incubated at 100 °C for 10 min. Samples were then digested for 3 h with 1 unit of *PNGase F* or *Endo H* (New England Biolabs, Hertfordshire, UK) at 37 °C. *Endo H* is an endoglycosidase that cleaves within the chitobiose core of high mannose and some hybrid oligosaccharide structures in *N*-linked glycoproteins. *PNGase F* is an amidase that cleaves between the innermost GlcNAc and asparagine residues of high-mannose, hybrid and complex oligosaccharides from *N*-linked glycoproteins. The deglycosylated proteins were loaded onto 7.5% SDS–PAGE and immunoblotting with anti-E-cadherin. For controls, the samples were incubated without the enzymes.

Calcium switch

AGS expressing E-cadherin WT, M1 and M234 were washed with phosphate-buffered saline (PBS), incubated with EGTA 2 mM for 10 min and supplemented with medium. Cells were then lysed after 2, 4 and 24 h with buffer containing PBS, 1% Triton X-100 and 1% NP40, and 30 µg of total cell lysate were subjected to 7.5% SDS–PAGE.

E-cadherin cis-dimer formation

Cells were plated in six-well plates to confluence. They were washed with PBS and incubated with 2.5 mM BS_3 (Sigma, St Louis, MO, USA) for 3h on ice. The resulting cells were then incubated with 10mM Tris (pH 7.5) for 15 min on ice. Cells were then lysed with buffer containing PBS, 1% Triton X-100 and 1% NP40, and 30 µg of total cell lysate were subjected to 6% SDS–PAGE. E-cadherin *cis*-dimer formation was detected by immunoblotting.

Immunofluorescence

AGS cells transfected with the different E-cadherin *N*-glycans mutants were plated on six-well plates with coverslips. Cells were fixed with methanol and blocked with bovine serum albumin 5% in PBS. Cells were incubated with mouse anti-E-cadherin monoclonal antibody.[15] Immunofluorescent images were obtained using a Zeiss Imager.Z1 AxioCam MRm (Carl Zeiss, Jena, Germany).

For the quantitative analyses of *in situ* immunofluorescence images, a 1D (one dimension) representative profile of protein level of expression and distribution was obtained. To undertake this analysis, original immunofluorescence images from cell populations were used to extract internuclear profile map from single cells and pairs of cells.[51]

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

Slow-aggregation assay

AGS cells transfected with E-cadherin WT and the different E-cadherin *N*-glycan mutants (5 × 104) were seeded onto an agar gel (0.66% w/v) in a 96-well plate. Aggregation formation was evaluated after 24 h using inverted phase-contrast microscope. Experiments were carried out in triplicate.

siRNA transfection

siRNA targeting *MGAT5* and control siRNA were purchased from Dharmacon (Lafayette, CO, USA). siMGAT5 (0–200nM) was transfected into AGS cells expressing Mock, E-cadherin WT, M1 and M234 with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Efficiency of *MGAT5* knockdown was optimal with 80 nM of siRNA after 48 h.

Statistical analysis

PLA analyses were performed using the SPSS statistical software (IBM, Armonk, NY, USA). Statistical associations between clinicopathological features and aberrant E-cadherin *N*-glycosylation mediated by GnT-V were assessed using the χ^2 test. Survival curves were estimated using the Kaplan–Meier method.

Statistical analyses were performed using the Graph Pad program (GraphPad Software, Inc., La Jolla, CA, USA). Student's tests were used to calculate significance in an interval of 95% confidence level. All statistics were compared with the E-cadherin WT group, and values of $P < 0.05$ were considered to be statistically significant (Student's t-test: $^{*}P \leqslant 0.05$; $^{**}P \leqslant 0.01$; $^{***}P \leqslant 0.001$).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

IPATIMUP integrates the I3S Research Unit, which is partially supported by FCT, the Portuguese Foundation for Science and Technology. This work is funded by FEDER funds through the Operational Programme for Competitiveness Factors — COMPETE — and National Funds through the FCT—Foundation for Science and Technology, under the projects: PTDC/CVT/111358/2009; EXPL/BIM-MEC/0149/2012; and PTDC/BBB-EBI/0786/2012. SC (SFRH/BD/77386/2011), AMD (SFRH/BI/52380/2013) BiotechHealth Doctoral Programme, and SSP (SFRH/BPD/63094/2009) thank FCT and the Luso-American Foundation (FLAD) for funding. JMS acknowledges FCT (UID/EEA/50009/2013). DK acknowledges support by the Max Planck Society and European Union (Seventh Framework Programme 'Glycoproteomics', grant number PCIG09-GA-2011-293847. CAR and DK acknowledge GastricGlycoExplorer project, grant number 316929). We thank Ola Soderberg and Gaëlle Cane (Department of Genetics and Pathology, University of Uppsala, Uppsala, Sweden) for providing the Streptavidin PLA probe and to Márcia Pereira and Sara Campos for the support in PLA technique and statistical analyses.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

REFERENCES

1. Takeichi M. Functional correlation between cell adhesive properties and some cell surface proteins. *J Cell Biol* 1977; **75**: 464–474.

2. Lewis JE, Jensen PJ, Johnson KR, Wheelock MJ. E-cadherin mediates adherens junction organization through protein kinase C. *J Cell Sci* 1994; **107**: 3615–3621.

3. Yap AS, Niessen CM, Gumbiner BM. The juxtamembrane region of the cadherin cytoplasmic tail supports lateral clustering, adhesive strengthening, and interac- tion with p120ctn. *J Cell Biol* 1998; **141**: 779–789.

4. Guilford P. E-cadherin downregulation in cancer: fuel on the fire? *Mol Med* Today 1999; **5**: 172–177.

5. van Roy F, Berx G. The cell-cell adhesion molecule E-cadherin. *Cell Mol Life Sci* 2008; **65**: 3756–3788.

6. Pinho SS, Seruca R, Gartner F, Yamaguchi Y, Gu J, Taniguchi N *et al*. Modulation of E-cadherin function and dysfunction by N-glycosylation. *Cell Mol Life Sci* 2011; **68**: 1011–1020.

7. Pinho SS, Carvalho S, Marcos-Pinto R, Magalhaes A, Oliveira C, Gu J *et al*. Gastric cancer: adding glycosylation to the equation. *Trends Mol Med* 2013; **19**: 664–676.

8. Shapiro L, Fannon AM, Kwong PD, Thompson A, Lehmann MS, Grubel G *et al*. Structural basis of cell-cell adhesion by cadherins. *Nature* 1995; **374**: 327–337.

9. Pertz O, Bozic D, Koch AW, Fauser C, Brancaccio A, Engel J. A new crystal structure, Ca2+ dependence and mutational analysis reveal molecular details of E-cadherin homoassociation. *EMBO J* 1999; **18**: 1738–1747.

10. Nelson WJ. Regulation of cell-cell adhesion by the cadherin-catenin complex. *Biochem Soc Trans* 2008; **36**: 149–155.

11. Paredes J, Figueiredo J, Albergaria A, Oliveira P, Carvalho J, Ribeiro AS *et al*. Epithelial E- and P-cadherins: role and clinical significance in cancer. *Biochim Biophys Acta* 2012; **1826**: 297–311.

12. Davis MA, Ireton RC, Reynolds AB. A core function for p120-catenin in cadherin turnover. *J Cell Biol* 2003; **163**: 525–534.

13. Hirohashi S, Kanai Y. Cell adhesion system and human cancer morphogenesis. *Cancer Sci* 2003; **94**: 575–581.

14. Zhou F, Su J, Fu L, Yang Y, Zhang L, Wang L *et al*. Unglycosylation at Asn-633 made extracellular domain of E-cadherin folded incorrectly and arrested in endoplasmic reticulum, then sequentially degraded by ERAD. *Glycoconj J* 2008; **25**: 727–740.

15. Pinho SS, Figueiredo J, Cabral J, Carvalho S, Dourado J, Magalhaes A *et al*. E-cadherin and adherens-junctions stability in gastric carcinoma: functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and V. *Biochim Biophys Acta* 2013; **1830**: 2690–2700.

16. Vester-Christensen MB, Halim A, Joshi HJ, Steentoft C, Bennett EP, Levery SB *et al*. Mining the O-mannose glycoproteome reveals cadherins as major O-mannosy- lated glycoproteins. *Proc Natl Acad Sci USA* 2013; **110**: 21018–21023.

17. Lommel M, Winterhalter PR, Willer T, Dahlhoff M, Schneider MR, Bartels MF *et al*. Protein O-mannosylation is crucial for E-cadherin-mediated cell adhesion. *Proc Natl Acad Sci USA* 2013; **110**: 21024–21029.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

18. Pinho SS, Osorio H, Nita-Lazar M, Gomes J, Lopes C, Gartner F *et al*. Role of E-cadherin N-glycosylation profile in a mammary tumor model. *Biochem Biophys Res Commun* 2009; **379**: 1091–1096.

19. Gu J, Sato Y, Kariya Y, Isaji T, Taniguchi N, Fukuda T. A Mutual regulation between cell–cell adhesion and N-glycosylation: implication of the bisecting GlcNAc for biological functions. J Proteome Res 2009; **8**: 431–435.

20. Pinho SS, Reis CA, Paredes J, Magalhaes AM, Ferreira AC, Figueiredo J *et al*. The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. *Hum Mol Genet* 2009; **18**: 2599–2608.

21. de-Freitas-Junior JC, Carvalho S, Dias AM, Oliveira P, Cabral J, Seruca R *et al*. Insulin/IGF-I signaling pathways enhances tumor cell invasion through bisecting GlcNAc N-glycans modulation. an interplay with E-cadherin. *PLoS One* 2013; **8**: e81579.

22. Yoshimura M, Ihara Y, Matsuzawa Y, Taniguchi N. Aberrant glycosylation of E-cadherin enhances cell-cell binding to suppress metastasis. *J Biol Chem* 1996; **271**: 13811–13815.

23. Kitada T, Miyoshi E, Noda K, Higashiyama S, Ihara H, Matsuura N *et al*. The addition of bisecting N-acetylglucosamine residues to E-cadherin down-regulates the tyrosine phosphorylation of beta-catenin. *J Biol Chem* 2001; **276**: 475–480.

24. Langer MD, Guo H, Shashikanth N, Pierce JM, Leckband DE. N-glycosylation alters cadherin-mediated intercellular binding kinetics. *J Cell Sci* 2012; **125**: 2478–2485.

25. Varelas X, Bouchie MP, Kukuruzinska MA. Protein N-glycosylation in oral cancer: dysregulated cellular networks among DPAGT1, E-cadherin adhesion and canonical Wnt signaling. *Glycobiology* 2014; **24**: 579–591.

26. Nita-Lazar M, Noonan V, Rebustini I, Walker J, Menko AS, Kukuruzinska MA. Overexpression of DPAGT1 leads to aberrant N-glycosylation of E-cadherin and cellular discohesion in oral cancer. *Cancer Res* 2009; **69**: 5673–5680.

27. Soderberg O, Gullberg M, Jarvius M, Ridderstrale K, Leuchowius KJ, Jarvius J *et al*. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods* 2006; **3**: 995–1000.

28. Przybylo M, Hoja-Lukowicz D, Litynska A, Laidler P. Different glycosylation of cadherins from human bladder non-malignant and cancer cell lines. *Cancer Cell Int* 2002; **2**: 6.

29. Jones J, Krag SS, Betenbaugh MJ. Controlling N-linked glycan site occupancy. *Biochim Biophys Acta* 20051726. 121–137.

30. Gavel Y, von Heijne G. Sequence differences between glycosylated and non- glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Eng* 1990; **3**: 433–442.

31. Bause E. Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. *Biochem J* 1983; **209**: 331–336.

32. Oda T, Kanai Y, Oyama T, Yoshiura K, Shimoyama Y, Birchmeier W *et al*. E-cadherin gene mutations in human gastric carcinoma cell lines. *Proc Natl Acad Sci USA* 1994; **91**: 1858–1862.

33. Jensen PH, Karlsson NG, Kolarich D, Packer NH. Structural analysis of N- and O-glycans released from glycoproteins. *Nat Protoc* 2012; **7**: 1299–1310.

34. Deshpande N, Jensen PH, Packer NH, Kolarich D. GlycoSpectrumScan: fishing glycopeptides from MS spectra of protease digests of human colostrum sIgA. *J Proteome Res* 2010; **9**: 1063–1075.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

35. Thaysen-Andersen M, Packer NH. Site-specific glycoproteomics confirms that protein structure dictates formation of N-glycan type, core fucosylation and branching. *Glycobiology* 2012; **22**: 1440–1452.

36. Liwosz A, Lei T, Kukuruzinska MA. N-glycosylation affects the molecular organization and stability of E-cadherin junctions. *J Biol Chem* 2006; **281**: 23138–23149.

37. Gu J, Isaji T, Xu Q, Kariya Y, Gu W, Fukuda T *et al*. Potential roles of N-glycosylation in cell adhesion. *Glycoconj J* 2012; **29**: 599–607.

38. Guo HB, Johnson H, Randolph M, Pierce M. Regulation of homotypic cell-cell adhesion by branched N-glycosylation of N-cadherin extracellular EC2 and EC3 domains. *J Biol Chem* 2009; **284**: 34986–34997.

39. Sato Y, Isaji T, Tajiri M, Yoshida-Yamamoto S, Yoshinaka T, Somehara T et al. An N-glycosylation site on the beta-propeller domain of the integrin alpha5 subunit plays key roles in both its function and site-specific modification by beta1,4- N-acetylglucosaminyltransferase III. *J Biol Chem* 2009; **284**: 11873–11881.

40. Pruitt KD, Tatusova T, Brown GR, Maglott DR. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res* 2012; **40**: D130–D135.

41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; **215**: 403–410.

42. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007; **23**: 2947–2948.

43. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 1993; **234**: 779–815.

44. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983; **22**: 2577–2637.

45. Chou PY, Fasman GD. Prediction of beta-turns. Biophys J 1979; 26: 367–383.

46. Rose PW, Bi C, Bluhm WF, Christie CH, Dimitropoulos D, Dutta S et al. The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res* 2013; **41**: D475–D482.

47. Cooper CA, Gasteiger E, Packer NH. GlycoMod--a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* 2001; **1**: 340–349.

48. Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res* 2008; **7**: 1650–1659.

49. Kolarich D, Rapp E, Struwe WB, Haslam SM, Zaia J, McBride R *et al*. The minimum information required for a glycomics experiment (MIRAGE) project: improving the standards for reporting mass-spectrometry-based glycoanalytic data. *Mol Cell Proteomics* 2013; **12**: 991–995.

50. York WS, Agravat S, Aoki-Kinoshita KF, McBride R, Campbell MP, Costello CE *et al*. MIRAGE: the minimum information required for a glycomics experiment. *Glycobiology* 2014; **24**: 402–406.

51. Sanches JM, Figueiredo J, Fonseca M, Durães C, Melo S, Esménio S *et al*. Quantification of mutant E-cadherin using bioimaging analysis of *in situ* fluorescence microscopy. A new approach to CDH1 missense variants. *Eur J Hum Genet* 2014; e-pub ahead of print 12 November 2014; doi:10.1038/ejhg.2014.240.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

Table 1. Relationship between aberrant glycosylation of E-cadherin mediated by GnT-V and prognostic variables of diffuse gastric cancer patients

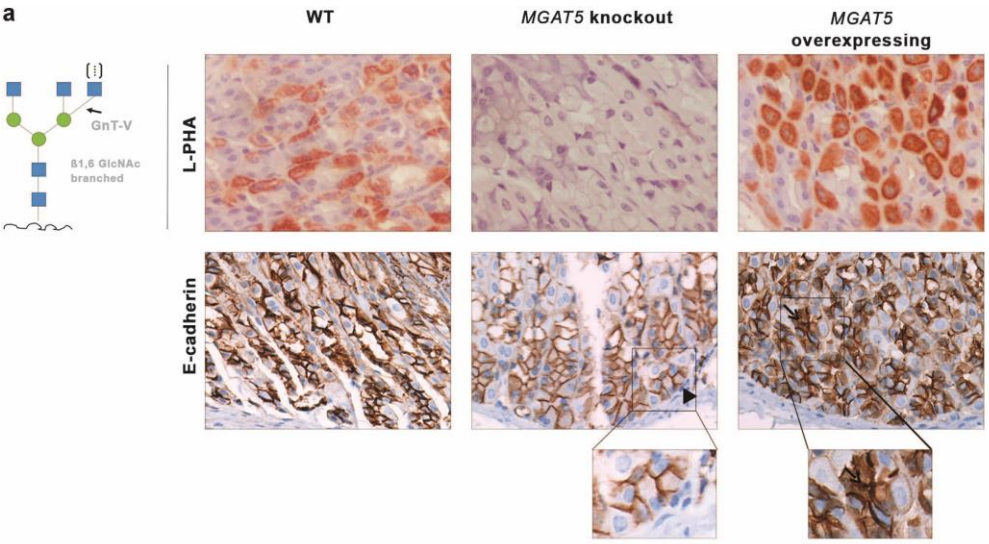
Variable	*PLA E-cadherin/L-PHA*		
	Negative	*Positive*	P
	No. of patients (%)		
Depth of invasion			
pT1	2 (25%)	6 (75%)	0.298
pT2	0 (0%)	4 (100%)	
pT3–4	3 (42.9%)	4 (57.1%)	
Lymph node involvement			
Absent (pN0)	2 (28.6%)	5 (71.4%)	0.865
Present (pN+)	3 (25%)	9 (75%)	
Lymphatic permeation			
Negative	2 (22.2%)	7 (77.8%)	0.701
Positive	3 (30%)	7 (70%)	
Venous invasion			
Negative	2 (18.2%)	9 (81.8%)	0.345
Positive	3 (37.5%)	5 (62.5%)	
Status			
Alive	5 (38.5 %)	8 (61.5%)	0.077
Dead	0 (0%)	6 (100%)	

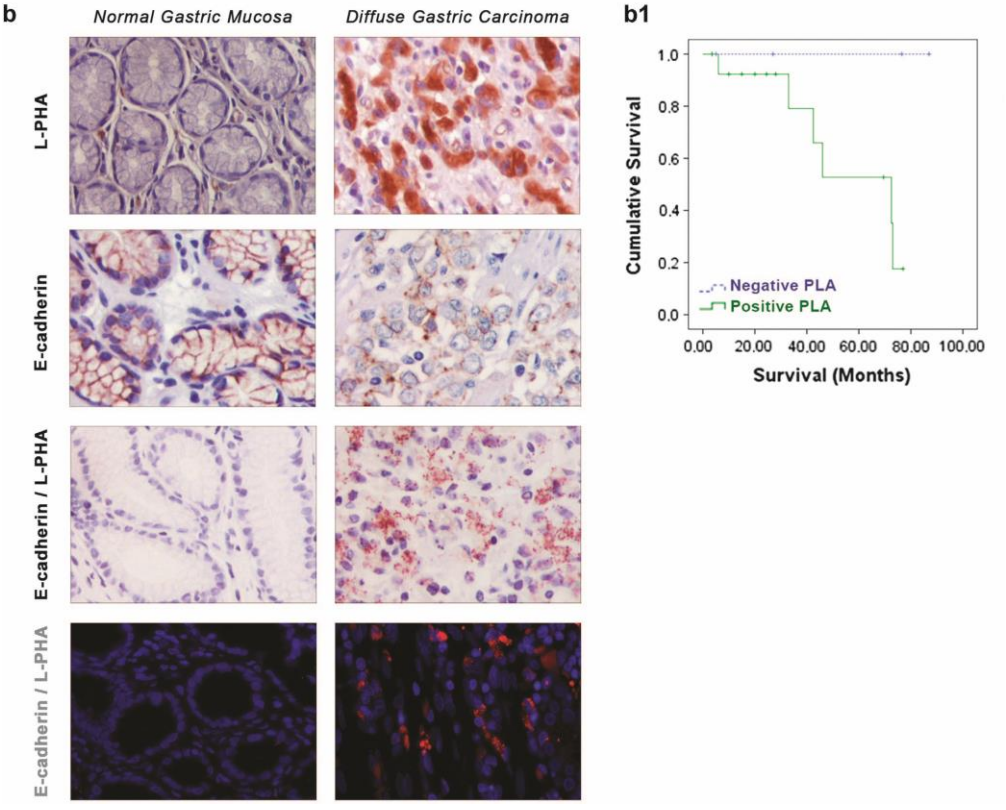
Abbreviations: Gnt-V, *N*-acetylglucosaminyltransferase V; L-PHA, *biotinylated Phaseolus vulgaris leucoagglutinin*; PLA, Proximity Ligation Assay.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

FIGURES





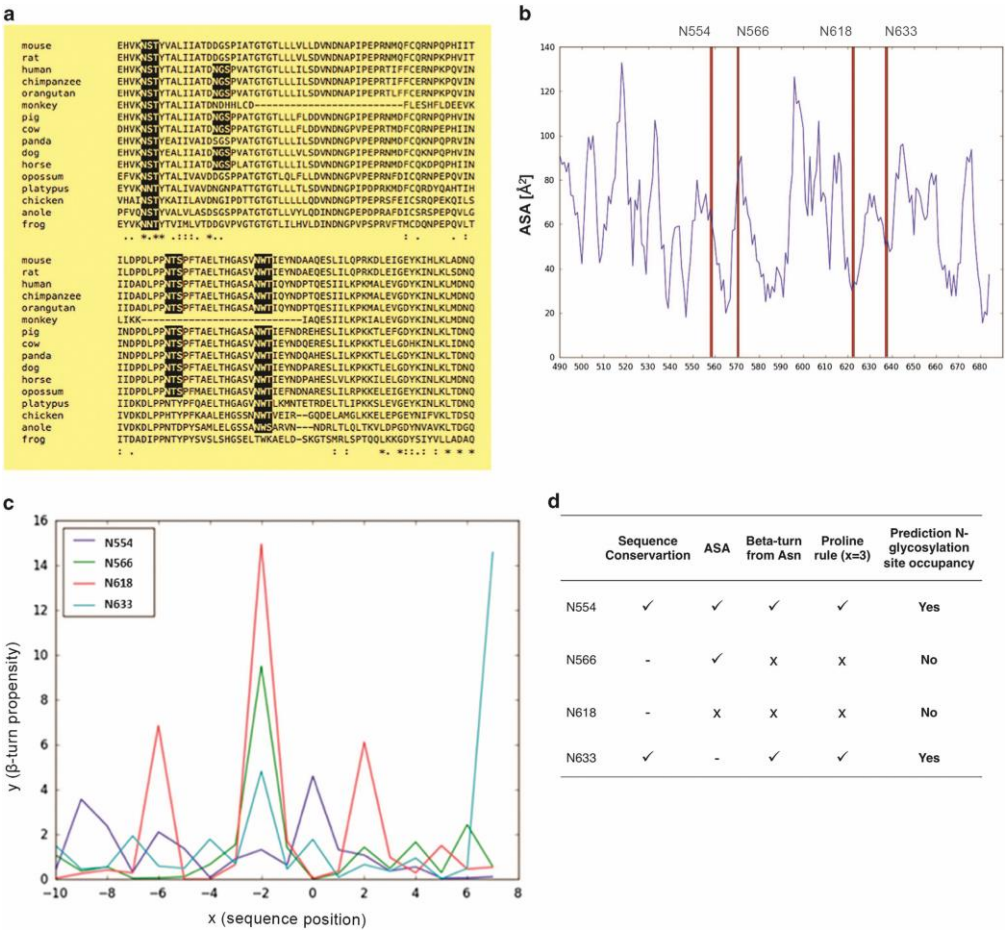
INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

Figure 1. Evaluation of the expression of E-cadherin and β1,6 GlcNAc-branched structures in the gastric mucosa of *MGAT5* knockout and *MGAT5* overexpressing and in human normal gastric mucosa versus gastric carcinoma. **(a)** L-PHA histochemistry detecting the β1,6 GlcNAc- branched *N*-glycans catalyzed by GnT-V showed a moderate expression of the β1,6-branched structures in the gastric mucosa of WT mice. No positive L-PHA reactivity was observed in the gastric mucosa of *MGAT5* KO mice, whereas a clear overexpression of β1,6 GlcNAc-branched structures was detected in *MGAT5* transgenic gastric mucosa. *MGAT5* KO mice showed a normal E-cadherin expression at the basolateral cell surface (arrowhead). In *MGAT5* transgenic mice, immunoexpression of E-cadherin was displayed aberrantly in the cytoplasm of mucous neck cells (not shown) and in deep glands of the gastric mucosa (arrow). **(b)** In situ PLA showed weak/absence PLA signal in normal gastric mucosa and a marked positivity of PLA signal (brown dots for brightfield and red dots for immunofluorescence) in neoplastic cells from gastric carcinoma, demonstrating a profound modification of E-cadherin with the β1,6 GlcNAc-branched structures in human gastric cancer (×40, original magnification). **(b1)** Survival rates of patients with diffuse gastric cancer (DGC) accordingly with PLA signal (positive versus negative). Kaplan–Meier curves demonstrate the probability of overall survival for patients with DGC accordingly with the presence/absence of GnT-V-mediated aberrant glycosylation of E-cadherin (P = 0.093).

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt



d

	Sequence Conservartion	ASA	Beta-turn from Asn	Proline rule (x=3)	Prediction N-glycosylation site occupancy
N554	✓	✓	✓	✓	Yes
N566	-	✓	x	x	No
N618	-	x	x	x	No
N633	✓	-	✓	✓	Yes

Figure 2. Prediction of E-cadherin *N-glycan* site occupancy based on bioinformatics. **(a)** Species homology of the E-cadherin peptide sequence. The potential *N*-glycosylation sites Asn-554 and Asn-633 are well conserved. **(b)** Calculation of ASA of the four Asn residues. The potential *N*-glycosylation site Asn-618 displays a low ASA value. **(c)** Evaluation of β-turn propensity of the region adjacent to the Asn-X-Ser/Thr acceptor sequence. Horizontal axis (*x*) indicates sequence position around the sequon (*x*=0: Asn), and vertical axis (*y*) shows the β-turn propensity. For *x* = 0, high values of β-turn propensity mean high possibility for *N*-glycosylation to occur. **(d)** Prediction of E-cadherin *N*-glycan's occupancy. Asn-554 and Asn-633 are the two putative *N*-glycosylation sites more likely to be occupied.

INSTITUTO DE INVESTIGAÇÃO E INOVAÇÃO EM SAÚDE
UNIVERSIDADE DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

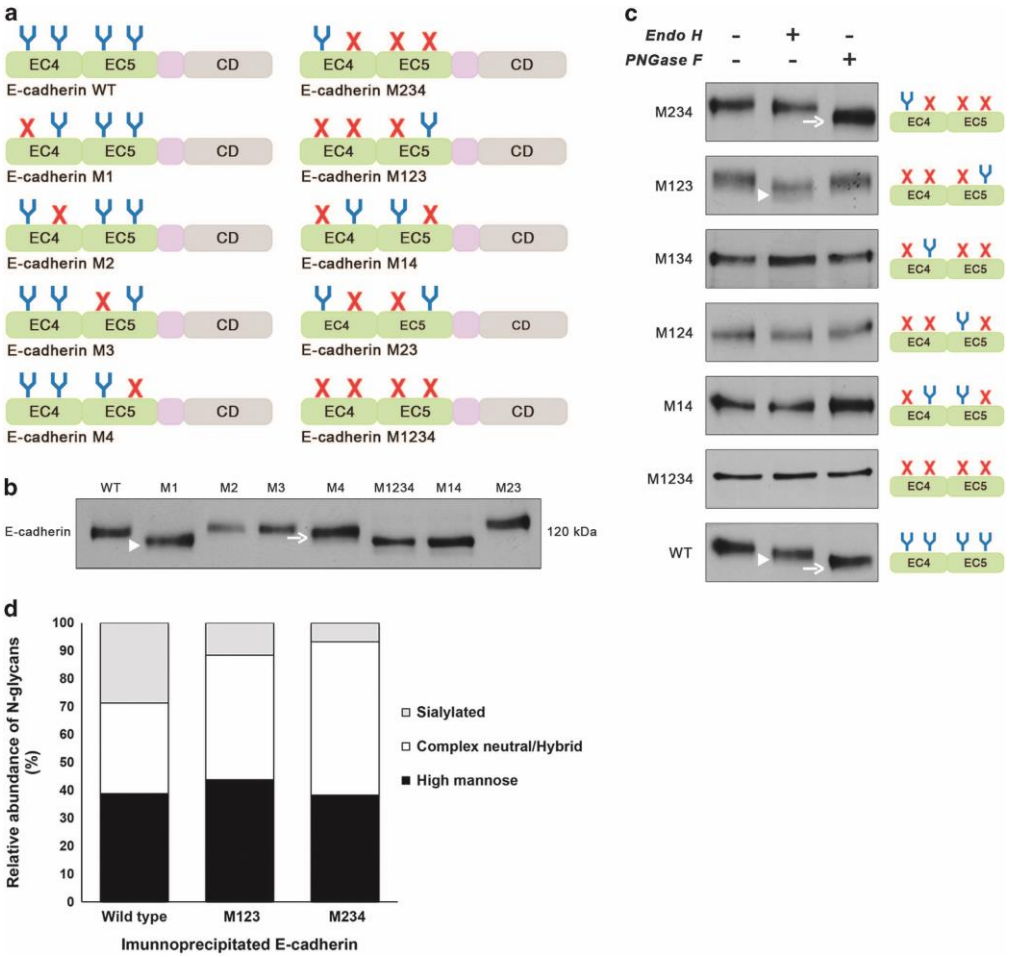


Figure 3. *N*-glycosylation sites Asn-554 and Asn-633 are *N*-glycosylated being occupied with complex-type *N*-glycans and high-mannose/hybrid- type *N*-glycans, respectively. **(a)** Schematic representation of different E-cadherin *N*-glycan mutants lacking selected putative *N*-glycosylation sites by replacing asparagine residue with glutamine in each *N*-glycosylation consensus sequence (NXS/T) either individually or in different combinations. EC_4, subdomain EC_4; EC_5, subdomain EC_5; CD, cytoplasmic domain; E-cadherin WT, E-cadherin wild type; E-cadherin M1, lacking Asn-554; M2, lacking Asn-566; M3, lacking Asn-618; M4, lacking Asn-633; M14, lacking Asn-554 and Asn-633; M23, lacking Asn-566 and Asn-618; M234, lacking Asn-566, Asn-618, and Asn-633; M123, lacking Asn-554, Asn-566, Asn-618; M1234, lacking Asn-554, Asn-566, Asn-618, and Asn-633. **(b)** Evaluation of the mobility shift of E-cadherin *N*-glycan mutants in AGS cells compared with WT. E-cadherin *N*-glycan mutant M1 and M4 showed an increased mobility compared with E-cadherin WT. No detectable changes were observed regarding cadherin M2 and M3. E-cadherin M1 (arrowhead) displayed a higher mobility shift than E-cadherin M4 (arrow). **(c)** Characterization of *N*-glycosylation profile of E-cadherin *N*-glycan mutants in AGS cells. Total cell lysates from selected transfectants were treated with *Endo H* and *PNGase F* and analyzed for mobility shifts by western blotting. E-cadherin WT was sensitive to Endo H (arrowhead) and PNGase F (arrow) being modified with high-mannose-, hybrid- and complex-type *N*-glycans. E-cadherin M234 was mostly sensitive to PNGase F (arrow) and resistant to Endo H while E-cadherin M123 was predominantly Endo H-sensitive (arrowhead). See also Supplementary Figure S3. **(d)** *N*-glycomics analysis using PGC nanoLC ESI MS/MS analyses showed E-cadherin WT to carry high-mannose, complex neutral and complex sialylated *N*-glycans in similar amounts. E-cadherin WT and M234 showed higher levels of sialylated and complex-type *N*-glycans, respectively. The data presented is referred to one biological replicate, which according to the well-established methodology, can be considered reproducible and significant. See also Supplementary Figure S2 and Supplementary Tables S1 and S2.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

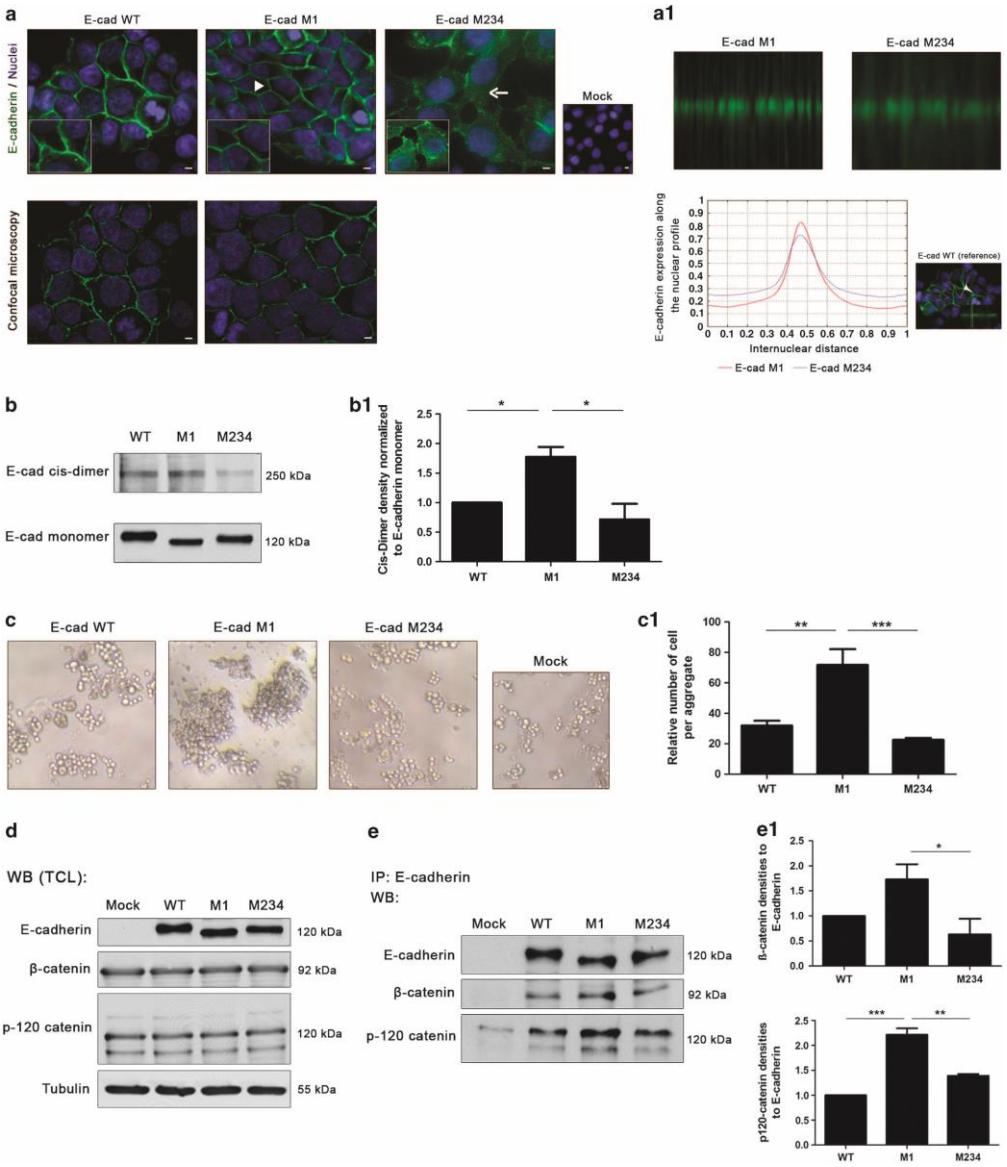


Figure 4. *N*-glycosylation at Asn-554 affects E-cadherin biological functions. **(a)** Immunofluorescence analysis showing that AGS cells expressing E-cadherin WT displayed an epithelial morphology and a membrane localization of E-cadherin (arrowhead). E-cadherin M1 showed a correct localization of E-cadherin at the cell membrane (arrowhead) with a more focused membrane staining than the WT (confocal microscopy). N-glycan mutant with Asn-554 (site 1) occupied with *N*-glycan structures (M234) is characterized by an incomplete localization of E-cadherin at the cell–cell contacts together with some cytoplasmic staining (arrow). White size bars ~5μm. **(a1)** Evaluation of the internuclear profiles of E-cadherin M1 and E-cadherin M234. Mutation at site 1 (E-cadherin M1) induces an increased membrane localization of E-cadherin compared with M234, which shows a decreased E-cadherin membrane expression. **(b)** Evaluation of *cis*-dimer-formation capacity of E-cadherin using BS_3. Mutation of Asn-554 (M1) leads to a significant increase of *cis*-dimerization of E-cadherin. **(b1)** Bar graphs. Amounts of E-cadherin *cis*-dimer were determined from the ratio of densities of E-cadherin *cis*-dimer/E-cadherin monomer. Results are described as mean ± s. d. of three independent experiments. The E-cadherin *cis*-dimer formation in E-cadherin M1 and M234 are expressed as the fold increase, compared with the E-cadherin WT, which was taken as 1 (Student's t-test: $^{*}P \leqslant 0.05$; $^{**}P \leqslant 0.01$; $^{***}P \leqslant 0.001$). **(c)** *In vitro* cell–cell aggregation assay. Mutation at Asn-554 (M1) resulted in a significant increase of cell–cell aggregation than E-cadherin WT and

INSTITUTO DE INVESTIGAÇÃO E INOVAÇÃO EM SAÚDE
UNIVERSIDADE DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

M234 showing larger cell aggregates. **(c1)** Bar graphs. Quantification of cell–cell aggregation by measuring the relative number of cells per aggregate (three or more cells). Results are described as the mean ± s.d. of three independent experiments. Student's t-test: $^{*}P \leqslant 0.05$; $^{**}P \leqslant 0.01$; $^{***}P \leqslant 0.001$). **(d)** Evaluation of the β- and p120-catenin total protein expression levels in the total protein lysates from AGS Mock, AGS E-cadherin WT and E-cadherin *N*-glycan mutants M1 and M234 cells. No significant differences were observed among the E-cadherin WT and E-cadherin *N*-glycan mutants. **(e)** E-cadherin immunoprecipitation followed by β- and p120-catenin Western-blot. The results showed that mutation of site 1 (Asn-554), M1, is associated with an increased interaction between E-cadherin and the β- and p120-catenin. A decreased interaction between E-cadherin and β-catenin (of about 1.0-fold) in M234 mutant compared with E-cadherin M1 was also verified. **(e1)** Bar graphs. Amounts of association were determined from the ratios of densities of β- or p120-catenin after normalization to E-cadherin. Results are described as mean ± s.d. of three independent experiments. The E-cadherin/catenin interaction levels in E-cadherin M1 and M234 are expressed as the fold increase, compared with E-cadherin WT, which was considered as 1 (Student's t-test: $^{*}P \leqslant 0.05$; $^{**}P \leqslant 0.01$; $^{***}P \leqslant 0.001$). See also Supplementary Figure S4.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

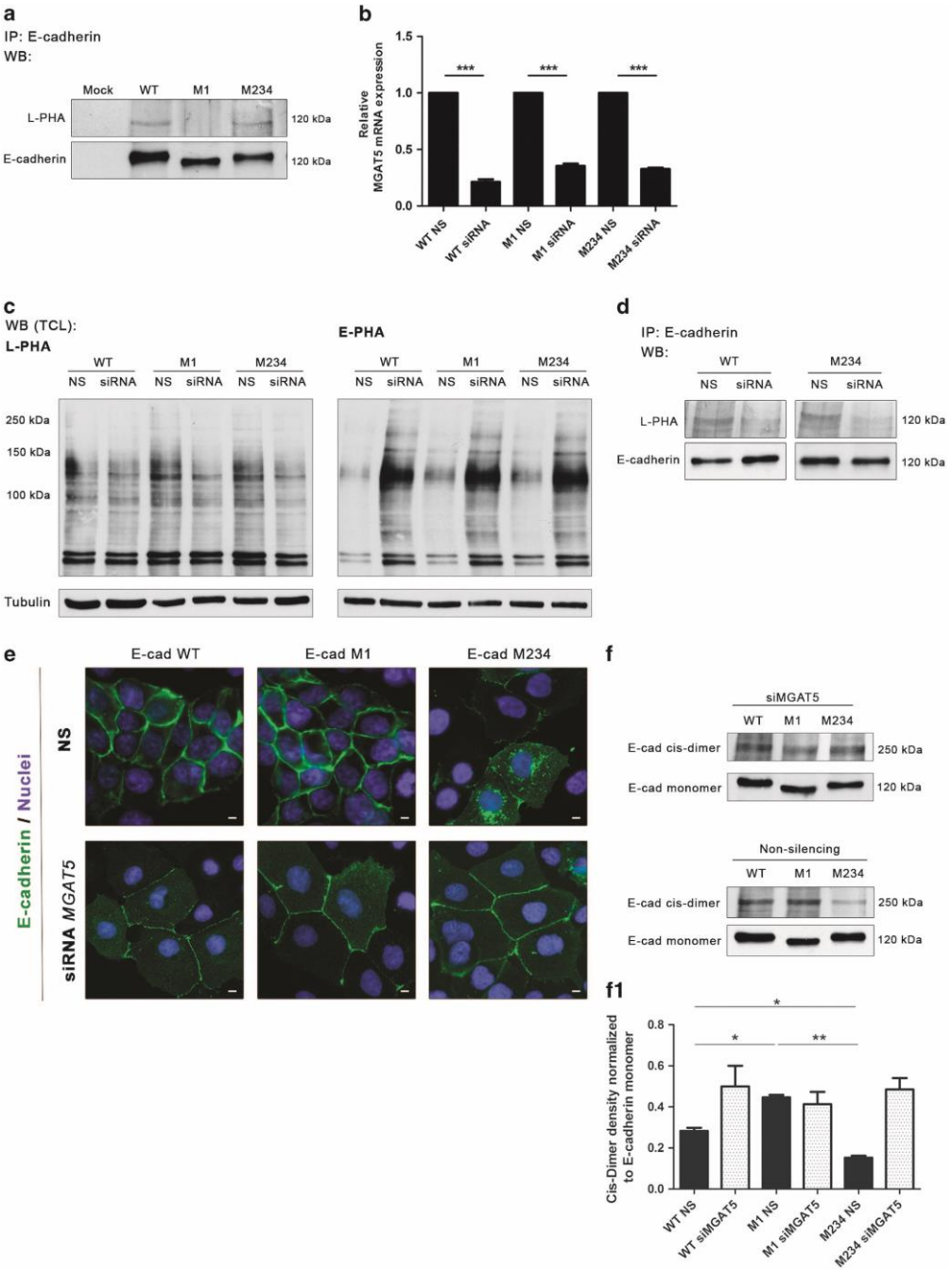


Figure 5. Knockdown of GnT-V induces the recovery of E-cadherin expression and functions. **(a)** Immunoprecipitation of E-cadherin M234 showed an increased L-PHA reactivity compared with the M1 mutant, revealing a significant increase in β1,6 GlcNAc-branched N-glycan structures modification in comparison with E-cadherin M1. **(b)** *MGAT5* mRNA expression of AGS cells expressing E-cadherin WT, M1 and M234 after MGAT5 silencing. Around 70–80% of *MGAT5* silencing was observed in WT, M1 and M234 mutants. NS, non-silencing. The relative mRNA expression of siRNA cells expressing E-cadherin WT, M1 and M234 are expressed as the fold increase, compared with the respective non- silencing cells, which was taken as 1 (Student's t-test: $^{*}P \leqslant 0.05$; $^{**}P \leqslant 0.01$; $^{***}P \leqslant 0.001$). **(c)** Lectin blot analysis of GnT-V and GnT-III products on total cell lysate from AGS cells expressing E-cadherin WT, M1 and M234 after MGAT5 knockdown. A decreased L-PHA reactivity after *MGAT5* knockdown and a competitive increase of E-PHA reactivity were observed. **(d)** *MGAT5* silencing resulted in a

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

remarkable decreased L-PHA reactivity for the E-cadherin band from the E-cadherin WT and E-cadherin M234, suggesting that Asn-554 is occupied with β1,6 GlcNAc- branched *N*-glycans. **(e)** Immunofluorescence analysis of AGS cells expressing E-cadherin WT, M1 and M234 after *MGAT5* knockdown. After *MGAT5* silencing, E-cadherin M234 showed an increased localization at the cell–cell membrane. White size bars ~ 5 μm. **(f)** Evaluation of the impact of *MGAT5* silencing on *cis*-dimer formation of E-cadherin. Knockdown of *MGAT5* leads to a similar ratio of E-cadherin *cis*-dimer formation in AGS cells expressing E-cadherin M1 and M234. **(f1)** Bar graphs. Amounts of E-cadherin *cis*-dimer were determined from the ratio of densities of E-cadherin *cis*-dimer/E-cadherin monomer. Results are reported as the mean ± s.d. of two independent experiment. See also Supplementary Figure S6.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt

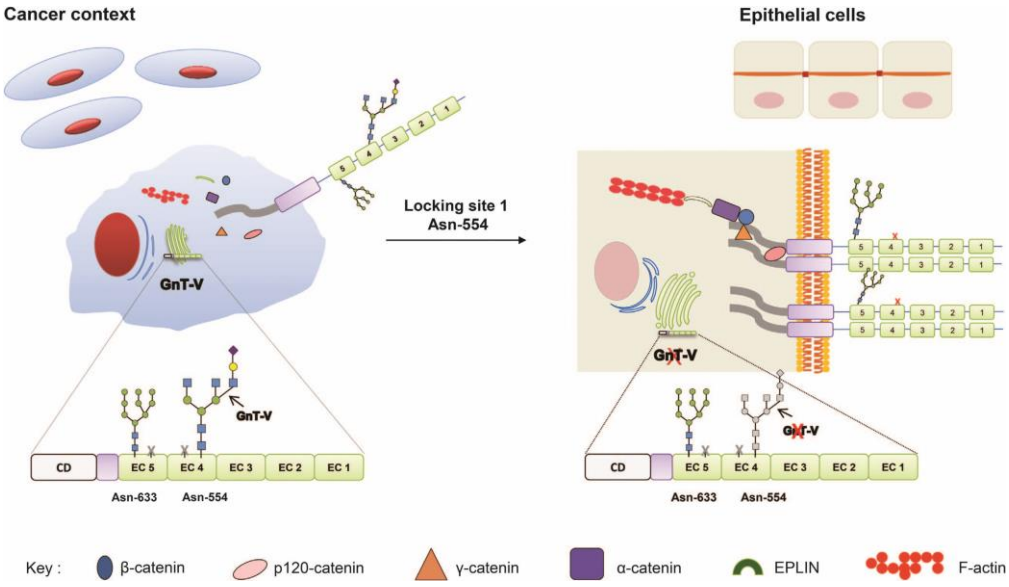


Figure 6. Proposing model for the site-specific E-cadherin *N*-glycosylation in cancer context. The site-specific occupancy of E-cadherin Asn-554 with β1, 6 GlcNAc-branched structures catalyzed by GnT-V impairs its biological functions leading to a decreased E-cadherin *cis*-dimerization capability, a decreased cellular aggregation and an impairment of the molecular assembly of adherens junctions. Locking this site-specific modification either by Asn-554 mutation or by GnT-V silencing recovers the biological functions of E-cadherin by increasing E-cadherin cis-dimerization, aggregation and the stability of the E-cadherin–catenin complex.

INSTITUTO
DE INVESTIGAÇÃO
E INOVAÇÃO
EM SAÚDE
UNIVERSIDADE
DO PORTO

Rua Alfredo Allen, 208
4200-135 Porto
Portugal
+351 220 408 800
info@i3s.up.pt
www.i3s.up.pt