Laboratory Investigation (2015) 95, 718–727

CDX2 homeoprotein is involved in the regulation of ST6GalNAc-I gene in intestinal metaplasia

Rita Pinto[1,2,3], Rita Barros[1,2], Isabel Pereira-Castro[2,8], Patricia Mesquita[4], Luis T da Costa[5], Eric P Bennett[6], Raquel Almeida[1,2,3,7] and Leonor David[1,2,3]

De novo expression of Sialyl-Tn (STn) antigen is one of the most common features of intestinal metaplasia (IM) and gastric carcinomas, and its biosynthesis has been mostly attributed to ST6GalNAc-I activity. However, the regulation of this glycosyltransferase expression is not elucidated. In IM lesions and in the intestine, CDX2 homeobox transcription factor is co-expressed with STn and ST6GalNAc-I. We therefore hypothesized that CDX2 might induce STn expression by positive regulation of *ST6GalNAc-I*. We showed that *ST6GalNAc-I* transcript levels and CDX2 have a coordinated expression upon Caco-2 *in vitro* differentiation, and overexpression of CDX2 in MKN45 gastric cells increases ST6GalNAc-I transcript levels. Nine putative CDX-binding sites in the *ST6GalNAc-I*-regulatory sequence were identified and analyzed by chromatin immunoprecipitation in Caco-2 cells and in IM. The results showed that CDX2 protein is recruited to all regions, being the most proximal sites preferentially occupied *in vivo*. Luciferase assays demonstrated that CDX2 is able to transactivate *ST6GalNAc-I*-regulatory region. The induction was stronger for the regions mapped in the neighbourhood of ATG start codon and site-directed mutagenesis of these sites confirmed their importance. In conclusion, we show that CDX2 transcriptionally regulates *ST6GalNAc-I* gene expression, specifically in the preneoplastic IM lesion.
Laboratory Investigation (2015) **95,** 718–727; doi:10.1038/labinvest.2015.52; published online 13 April 2015

ST6GalNAc-I is a α2,6-sialyltransferase present in several human tissues. The active enzyme, with a predicted 600 amino-acid length,[1] transfers a sialic acid residue to GalNAcα-O-Ser/Thr, generating sialyl-Tn (STn) antigen. The biosynthesis of STn has been mostly attributed to ST6GalNAc-I activity, both *in vitro* and *in vivo*.[1–4] It is rarely observed in normal tissues but frequently accompanies cancer progression, being one of the most common features of premalignant lesions of the gastrointestinal tract, namely intestinal metaplasia (IM),[5,6] and of carcinomas, namely gastric carcinomas.[5–7] However, to date, little is known about *ST6GalNAc-I* regulation that might explain aberrant expression of STn in preneoplastic and cancer.

Previous studies identified the homeobox transcription factor CDX2 as a direct regulator of MUC2 mucin expression,[8,9] the major carrier of STn in IM, in most gastric carcinomas[10] and in gastrointestinal mucinous adenocarcinomas.[11] Under normal conditions, CDX2 expression in adults is restricted to the intestine, being considered a 'master regulator' of intestinal terminal differentiation. However, CDX2 becomes ectopically expressed in human IM lesions of the stomach,[12,13] oesophagus[14] and gallbladder.[15] Both in the intestine and in IM foci in the stomach, CDX2 is co-expressed with MUC2,[12] ST6GalNAc-I[4] and STn.[5,6]

Besides MUC2, CDX2 controls the expression of multiple other intestinal proteins, such as LI-cadherin (LI-cad),[16] sucrase-isomaltase (SI)[17] and CDX2 itself.[18] A chromatin immunoprecipitation (ChIP)-seq analysis performed in the human intestinal cell line Caco-2 has further revealed direct binding of CDX2 to novel target genes and binding sites, namely to a 3′-end enhancer element of the HNF4α gene.[19] Caco-2 *in vitro* spontaneous differentiation into enterocyte-like cells is accompanied by an increase in CDX2 expression,[20] as well as alterations in the cellular glycosylation profile and enzyme expression/activity.[21–26] Moreover, the genetic regulation of glycosyltransferases—β3Gal-T5 and FUT2—by CDX2 has already been described in the intestinal cell lines.[24,27]

[1]Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal; [2]Institute of Molecular Pathology and Immunology, University of Porto, Porto, Portugal; [3]Faculty of Medicine, University of Porto, Porto, Portugal; [4]Instituto Nacional de Investigação Agrária e Veterinária, Quinta da Fonte Boa, Vale de Santarém, Portugal; [5]ICAAM, Universidade de Évora, Évora, Portugal; [6]Copenhagen Center for Glycomics, Departments of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark and [7]Department of Biology, Faculty of Sciences, University of Porto, Porto, Portugal
Correspondence: Professor L David, MD, PhD, Institute of Molecular Pathology and Immunology, University of Porto, Rua Dr. Roberto Frias, s/n, Porto 4200-465, Portugal.
E-mail: ldavid@ipatimup.pt
[8]Current address: Gene Regulation Group, IBMC—Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal.

Received 12 December 2014; revised 8 February 2015; accepted 7 March 2015

We therefore hypothesized that CDX2 homeobox transcription factor was involved in characteristic glycoproteome modifications of the preneoplastic lesion IM, namely the regulation of the expression of ST6GalNAc-I and its biosynthetic product STn. In order to evaluate the involvement of CDX2 on *ST6GalNAc-I* gene transactivation, both *in vitro* and *in vivo* gastrointestinal models were used.

MATERIALS AND METHODS

Human Tissue Samples

Frozen sections of gastric specimens with IM foci adjacent to carcinomas were obtained from patients undergoing surgery at Hospital S. João, Medical Faculty (Porto, Portugal) after ethical committee's approval and informed consent of the patients. Serial sections were cut and used for immunostaining.

Collection of samples for ChIP assays was performed in one case of IM under the same ethical approval process as above. Normal gastric mucosa and IM foci were selected upon frozen specimen's microscopic evaluation at the operating theater after removal of gastric cancer specimens.

Immunostaining

Immunohistochemical analysis of frozen sections was performed using anti-human CDX2 (CDX2-88 clone, Biogenex), anti-human MUC2 (PMH1 clone hybridoma supernatant, Copenhagen, Denmark), anti-human ST6GalNAc-I (2C3 clone hybridoma supernatant, Copenhagen, Denmark) and anti-human STn (TKH2 clone hybridoma supernatant, Copenhagen, Denmark). Negative controls were performed by omission of primary antibodies.

Briefly, for CDX2 and MUC2 immunostaining, the avidin–biotin–peroxidase complex (ABC) method was used. Frozen specimens were fixed in cold methanol for 5 min (for CDX2 staining) or acetone for 10 min (for MUC2 staining) and then treated with 3% hydrogen peroxide (H_2O_2) in methanol for 10 min to block endogenous peroxidase. Sections were then incubated for 45 min with normal rabbit serum in phosphate buffered saline (PBS) with 10% bovine serum albumin (BSA) to block nonspecific staining. Normal serum was then replaced by the primary antibodies (CDX2-88 1:50 and PMH1 undiluted hybridoma supernatant) in PBS with 5% BSA. Sections were incubated overnight at 4 °C. After washing, sections were incubated with biotin-labeled secondary rabbit anti-mouse antibody (Dako, Glostrup, Denmark) diluted 1:100 in PBS with 5% BSA for 30 min and with ABC Kit (Vector Labs, CA, USA) for 30 min. The reaction was developed with 0.05% 3,3′-diaminobenzidinetetrahydrochloride containing 0.01% H_2O_2, and sections were then counterstained with haematoxylin, dehydrated and mounted. Slides were examined using a Zeiss Optical Microscope.

For ST6GalNAc-I and STn immunofluorescence in frozen specimens, sections were fixed in cold acetone for 10 min and incubated for 20 min with normal rabbit serum in PBS with 10% BSA. Normal serum was then replaced by specific primary antibody (2C3 undiluted supernatant and TKH2 1:5) in PBS with 5% BSA. Sections were incubated overnight at 4 °C. After washing, sections were incubated with rabbit anti-mouse Ig FITC-labeled secondary antibody (Dako, Glostrup, Denmark) diluted 1:100 in PBS with 5% BSA for 45 min, protected from light. DAPI was used as a nuclear counterstain, and slides were mounted in Vectashield mounting media (Vector Labs). Samples were examined under a Leica DM2000 fluorescence microscope equipped with DAPI and FITC interference filters. Images were acquired using a Leica DFC340 FX camera and Leica Application Suit software.

Cell Culture

Human gastric carcinoma cell lines AGS (obtained from ATCC) and MKN45 (obtained from the Japanese Collection of Research Bioresources) were grown in RPMI 1640 medium with Glutamax (GIBCO) while human colon carcinoma cell line Caco-2 was maintained in Dulbecco's minimal essential medium (DMEM) (GIBCO). Media were supplemented with 10% inactive fetal bovine serum (FBS) (Invitrogen) and 1% antibiotics (100 U/ml penicillin and 100 μg/ml streptomycin) (GIBCO), and cells were maintained at 37 °C in a humidified 5% CO_2 incubator.

Time-Course, CDX2 Overexpression and CDX2 Knockdown Experiments

For the time-course experiments on Caco-2, 1×10^5 cells per well were seeded in six-well plates and harvested 2 days before reaching confluence, at confluence (Day 0) and at variable times after confluence.

For CDX2 overexpression in MKN45, cells seeded in six-well plates and at ≈80% confluence were transfected with 1 μg of a human CDX2 expression vector containing two FLAG tags at the N-terminus[18] or the corresponding empty vector in a ratio of 1:1.5 relatively to Lipofectamine 2000 reagent in Opti-MEM medium (GIBCO).

For CDX2 knockdown in AGS cells, cells seeded in six-well plates and at ≈80% confluence were transfected with a mix of three double-stranded small interfering RNA (siRNA) directed to different sequences of CDX2 mRNA (siCDX2) or scrambled controls in a total concentration of 50 nM. Target sequences and scrambled controls were chosen according to standard RNAi rules. The siRNA duplexes were used in a ratio of 1:5 relatively to Lipofectamine 2000 reagent in Opti-MEM medium. siRNA sequences are shown in Supplementary Table S1.

RNA Extraction, Reverse Transcriptase Assay and Quantitative PCR

Total RNA was extracted using TRI Reagent (Sigma, St Louis, MO, USA) and reverse transcription was performed with 3.5 μg total RNA. Specific transcripts were amplified using SYBR Green (Applied Biosystems, Foster City, CA, USA) in a fluorescence reader ABI Prism 7500 (Applied Biosystems). The sequences of the primers used are presented in Supplementary Table S1. 18S levels were used for normalization of target gene

abundance, and relative mRNA levels were calculated. Each experiment was carried out at least three times and in triplicate. The results are expressed as mean ± s.d. of the triplicates from a representative experiment.

Western Blotting of Cell Lysates

Whole-cell extracts were obtained by resuspension of cell pellets in RIPA buffer (50 mM Tris–HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.1% sodium dodecyl sulphate) in the presence of complete protease inhibitors cocktail (Roche, Indianopolis, IN, USA). Quantification of total protein was determined by bicinchoninic acid protein assay (Pierce, Rockford, IL, USA). Protein extracts (25 μg) were then analyzed by standard SDS-PAGE, transferred to a nitrocellulose membrane (Amersham Biosciences, Buckinghamshire, UK) and blotted o.n. at 4 °C with mouse monoclonal primary antibodies CDX2-88 1:500 (Biogenex) and anti-actin 1:8000 (Santa Cruz Biotechnology) in 5% BSA in PBS. Peroxidase-conjugated secondary antibodies (goat polyclonal anti-mouse-HRP 1:2000 for CDX2 and goat polyclonal anti-rabbit-HRP 1:2000 for actin, both from Santa Cruz Biotechnology) were used and developed with the ECL Detection Kit (BioRad, Hercules, USA). Quantification of the western blots was performed using the Bio-Rad Software Quantity One.

ChIP in an Intestinal Cell Line and in Human Intestinal Metaplasia

In vitro ChIP was carried out using 6 days postconfluence Caco-2 cells (Day +6) using Magna ChIP Kit (Upstate Biotechnology, Millipore, Temecula, USA) according to the manufacturer's recommendations. Briefly, cells were washed 1× with PBS, and DNA/protein crosslinking was induced with 1% formaldehyde for 10 min at room temperature and stopped with glycine. After washing 1× in PBS containing protease inhibitors, cells were lysed and sonicated in nuclear lysis buffer. The DNA fragment length range after sonication was checked, by running sonicated DNA on a gel. The solubilized chromatin was then incubated with anti-CDX2 antibody (BioGenex, The Hague, The Netherlands) or isotype control immunoglobulin G (IgG) together with Protein G-magnetic beads, overnight at 4 °C. Immunoprecipitated chromatin was then washed, and reversion of the crosslink was performed with proteinase K at 45 °C for 2 h followed by a 10-min incubation at 95 °C. DNA was then purified and analyzed by quantitative PCR using primer pairs covering the human ST6GalNAc-I promoter (Supplementary Table S1).

For *in vivo* ChIP, whole surgical specimens were briefly washed with water for removal of blood, and areas adjacent to the carcinoma were collected. The mucosa was separated from the muscle layer and washed in cold PBS, and the tissue segment was then cut into smaller pieces to facilitate the fixation in 1% formaldehyde and fast frozen in liquid nitrogen. Without thawing, the tissue was pulverized using a Biopulverizer (Biospec Products, Bartlesville, USA) to facilitate posterior lysis. Following this step, the Magna ChIP Kit was used as described. An adjacent/contiguous sample was collected for each of the fragment used for ChIP, frozen in OCT and stained with haematoxylin and eosin to check for the presence of IM.

Each experiment was carried out three times (for the Caco-2 cell line) or once (for each of the surgical specimens) in triplicate. The results are expressed as mean ± s.d. of the triplicates from a representative experiment.

Construction of Luciferase Reporter Plasmids

Three overlapping fragments upstream of human *ST6GalNAc-I* gene (−1.6, −0.8 and −0.3 kbp starting from the transcription start site) were amplified by PCR from purified genomic DNA samples obtained from human normal blood using BR1+8, BR4 and BR9 as forward primers (Supplementary Table S1) and a reverse primer containing an overhang immediately upstream of the transcription start site (5′-TATAGG TGGTGGGTCGGGTTC-3′). The amplified fragments were analysed on a 1% Tris-borate EDTA (TBE) agarose gel and gel-extracted using the QIAquick Gel Extraction Kit (Qiagen). The fragments were then cloned into pTA-Luc, an in house-developed luciferase TA-cloning vector. The search for putative transcription factor-binding sites in these regions was performed in TRANSFAC.[28] Reporter plasmids with mutated putative CDX2-binding sites were generated by site-directed mutagenesis (oligos listed in Supplementary Table S1) using the QuickChange protocol (Stratagene, Cedar Creek, USA). Mutations were created by substitution of A/T triplets for C/G triplets not prone to bind CDX2. All the constructed plasmids were verified by sequencing using the Luc-F and Luc-R plasmid primers (Supplementary Table S1), to ensure that they had the predicted sequences.

Luciferase Assays

Human gastric carcinoma cell lines AGS and MKN45 were seeded in 24-well plates and cultured under standard conditions. When ≈80% confluent, cells were transfected with a DNA mix in 300 μl of Opti-MEM media (Invitrogen), free of antibiotic and antimycotic, added to each well. In all, 0.5 μg of each reporter plasmid were used for cell transfection with Lipofectamine 2000 (Invitrogen, Carlsbad, USA). Cells were co-transfected with 1 μg of the human CDX2 expression vector or the corresponding empty vector and also of the 0.3 μg *Renilla* luciferase reporter vector. Ratio used was 1 mg DNA:1.5 ml Lipofectamine. Twenty-four hours posttransfection, Opti-MEM media was replaced by media supplemented with 10% FBS and 1% penicillin–streptomycin, and 48 h after transfection, luciferase activity was measured. Total extracts were prepared using the Dual-Luciferase Reporter Assay System Kit (Promega, Madison, USA) according to the manufacturer's instructions, and luciferase activity was measured in a 1450 Microbeta luminescence counter (Wallac, Perkin Elmer, Massachussets, USA). As controls, we used a luciferase vector with a CMV promoter and a promotorless

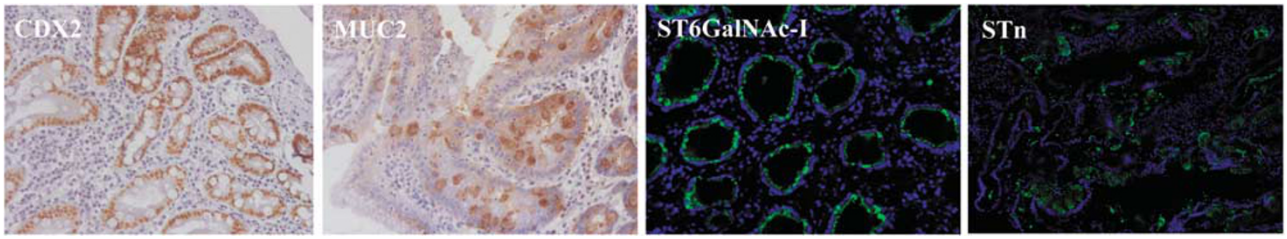

Figure 1 CDX2, MUC2, ST6GalNAc-I and STn immunohistochemical expression in intestinal metaplasia. Images of CDX2, MUC2 and STn staining were taken at ×200 magnification and the image of ST6GalNAc-I was taken at ×400 magnification.

luciferase construct, together with *Renilla* luciferase reporter vector to normalize for transfection efficiency. Each experiment was carried out at least twice and in triplicate for each of the fragments, and the results are expressed as mean ± s.d. of at least six values. Results are expressed as fold induction compared with values obtained for the empty vector.

Statistics

Data are expressed as means ± s.d. Statistical analysis was performed using Student's *t*-test. A $P < 0.05$ was considered as significantly different.

RESULTS

CDX2 Expression is Associated with ST6GalNAc-I and STn Expression in Intestinal Metaplasia

We observed that CDX2, MUC2, ST6GalNAc-I and STn are co-expressed in IM (Figure 1) and are absent in normal gastric mucosa, extending previous observations from our group.[4,12]

ST6GalNAc-I mRNA and CDX2 have a Coordinated Expression During in vitro Spontaneous Differentiation of Caco-2 Cells

Caco-2 cell line is derived from colon adenocarcinoma and frequently used as a model for intestinal differentiation as it undergoes spontaneous *in vitro* differentiation into enterocyte-like cells upon confluence. This *in vitro* differentiation is accompanied by an increase in CDX2 expression,[20] and it was previously shown that CDX2 regulates *β3Gal-T5* in this cellular model.[24] Therefore, we chose this cell line to assess the putative regulation of the *ST6GalNAc-I* gene by CDX2. Consistent with previous observations, CDX2 protein expression progressively increased during Caco-2 differentiation (Figure 2a, as a representative example of a total of three experiments), and *ST6GalNAc-I* transcript levels showed a concordant increase in time-course experiments (Figure 2b). The previously identified CDX2 target *β3Gal-T5* was used as a control. STn expression was also observed in Caco-2 cells (Supplementary Figure S2). However, STn expression levels are maintained along Caco-2 culture and do not accompany ST6GalNAc-I and CDX2 increase (data not shown). We could also not explore our previous findings of MUC2 as a carrier of STn, as Caco-2 cells mainly differentiate into enterocytic lineage and consequently express very low levels of MUC2.[29]

As activity and protein levels of several glycosyltransferases change during Caco-2 *in vitro* differentiation,[21–26] we decided to check whether or not the increase observed for *ST6GalNAc-I* mRNA expression is restricted or part of a more general glycosyltransferase induction. Therefore, the levels of other relevant glycosyltransferases were also evaluated by quantitative PCR (Supplementary Figure S1A). We observed that *GalNAc-T3* shows an expression profile similar to ST6GalNAc-I. The correspondent increase in the GalNAc-T3 protein expression can be also observed by immunohistochemistry (Supplementary Figure S1B).

CDX2 Regulates ST6GalNAc-I mRNA Expression in Gastric Carcinoma Cell Lines

In order to assess the putative involvement of CDX2 in *ST6GalNAc-I* expression regulation in the gastric setting, we overexpressed and downregulated CDX2 in gastric cell lines. Because of its low CDX2 expression levels, MKN45 cell line was transfected with a human CDX2 expression vector (Figure 3a), and an increase in *ST6GalNAc-I* mRNA levels was observed 72 h after transfection, whereas the levels of the CDX2 targets *β3Gal-T5* and *LI-cad* start to increase earlier (Figure 3b). Nevertheless, no increase in STn, the ST6GalNAc-I biological product, could be detected both by immunofluorescence and flow cytometry (data not shown).

Among gastric carcinoma cell lines, AGS expresses high levels of CDX2 and was therefore used as a model of CDX2 downregulation using siRNAs (Figure 3c). In this experiment, we observed that 48 h after transfection with a mix of siCDX2 there was a decrease of about 50% in *ST6GalNAc-I* mRNA levels as well as in the CDX2 known targets *β3Gal-T5* and *LI-cad*. However, 72 h after transfection, *ST6GalNAc-I* levels increased again, despite that low levels of CDX2 and other targets are maintained (Figure 3d). Again, STn expression levels were not altered by CDX2 downregulation (data not shown).

Identification of Active CDX2 cis-Elements within the ST6GalNAc-I Gene-Regulatory Region

The *ST6GalNAc-I* promoter has not been characterized thus far. As most gene-regulatory regions are located upstream of the start codon, we inspected *in silico* about 2000 bp upstream of the *ST6GalNAc-I* translation start site for CDX consensus-binding sequences, using the TRANSFAC database.[8,16,17,28]

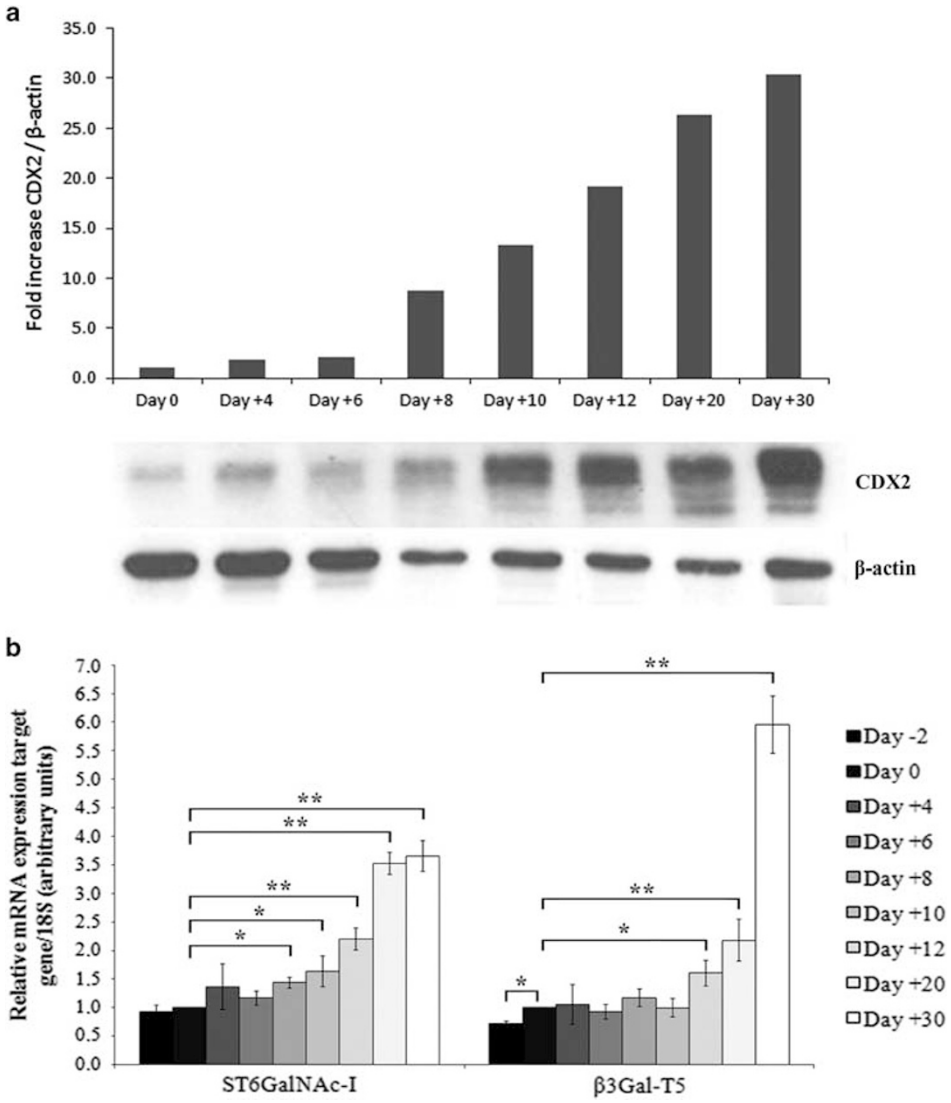

Figure 2 Alterations in CDX2 protein expression and *ST6GalNAc-I* mRNA levels during *in vitro* spontaneous differentiation of Caco-2 intestinal cell line. (**a**) Western blotting showing the increase in CDX2 protein expression. β-Actin was used as a loading control. Absence of Day − 2 protein expression results on western blotting is due to the low protein amount obtained after cell lysis. (**b**) *ST6GalNAc-I* mRNA expression levels increase and decrease concomitantly with CDX2 protein. *β3Gal-T5* mRNA expression was used as a control. Day 0 corresponds to the day where cells reach confluence, and the results at this point were referred to as 1. The results are expressed as mean ± s.d. of the triplicates from a representative experiment (*$P<0.05$ and **$P<0.01$).

We identified nine putative-binding regions (Figure 4a), which were analysed by ChIP assay in Caco-2 cells 6 days after confluence. All putative CDX-binding sites were considered except BR3, as no suitable primers could be designed to specifically amplify this region owing to its high A/T content. The results show that CDX2 protein was able to bind to the eight regions studied, when compared with an irrelevant IgG control (Figure 4b).

CDX2 is Bound to ST6GalNAc-I-Regulatory Region in Human Intestinal Metaplasia

In order to confirm the binding of CDX2 to the *ST6GalNAc-I*-regulatory region and its relevance *in vivo*, ChIP was performed against CDX2 in a surgical sample of human IM, where the association between expression of the two molecules had been observed.[4,12] As CDX2 is present in the stomach exclusively in the foci of IM, a whole mucosa fragment containing both normal and metaplastic glands was used. Again, all putative CDX2-binding sites were considered except BR3. Another set of primers encompassing the CDX-binding site on the SI promoter was used as a positive control (primer sequences in Supplementary Table S1). The results showed that CDX2 was bound to BR7 and BR9 regions, which are located close to the translation start site (Figure 5). As expected, CDX2 was also bound to SI promoter. A fragment of normal gastric mucosa

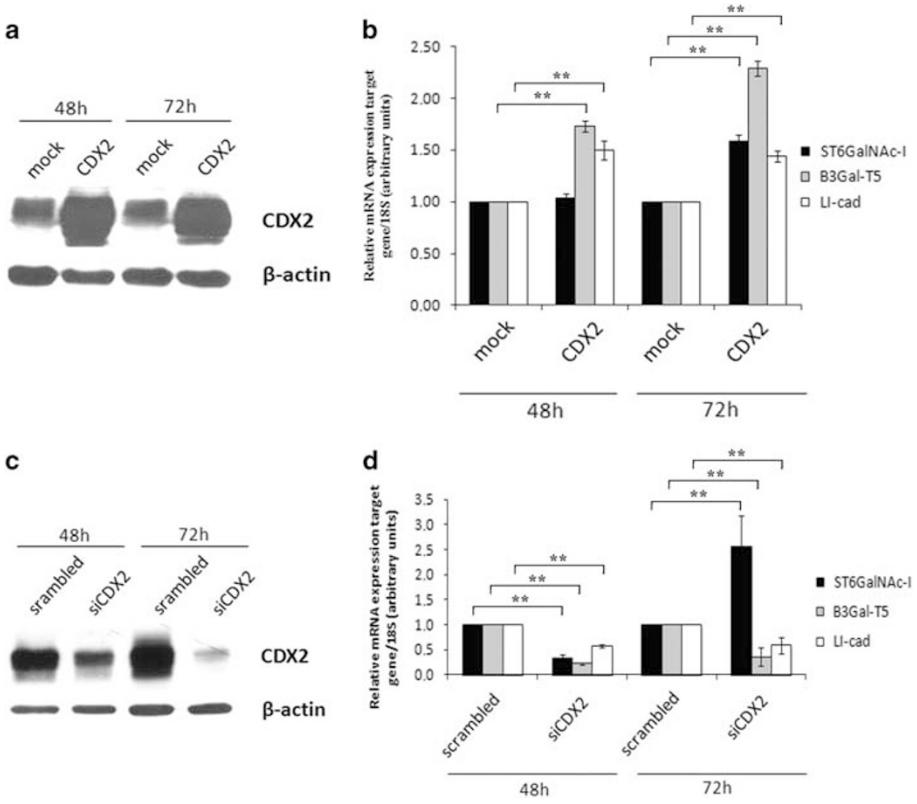

Figure 3 Effect of overexpression or downregulation of CDX2 in gastric cell lines on the *ST6GalNAc-I* transcript levels. (**a**) Western blotting showing CDX2-induced overexpression in MKN45 cell line. *β*-Actin was used as the loading control. (**b**) Fold increase of *ST6GalNAc-I* gene expression upon CDX2 overexpression. *β3Gal-T5* and *LI-cad* genes, regulated by CDX2, were used as a control. The values obtained with mock-treated cells were referred to as 1. mRNA levels were normalized with the corresponding 18S mRNA level. (**c**) Western blotting showing CDX2 knockdown in AGS cell line. *β*-Actin was used as the loading control. (**d**) Variation on *ST6GalNAc-I* gene expression upon CDX2 knockdown. *β3Gal-T5 and LI-cad* genes, regulated by CDX2, were used as a control. The values obtained for cells transfected with scrambled control were referred to as 1. mRNA levels were normalized with the corresponding 18S mRNA level. The results are expressed as mean ± s.d. of the triplicates from a representative experiment (***P* < 0.01).

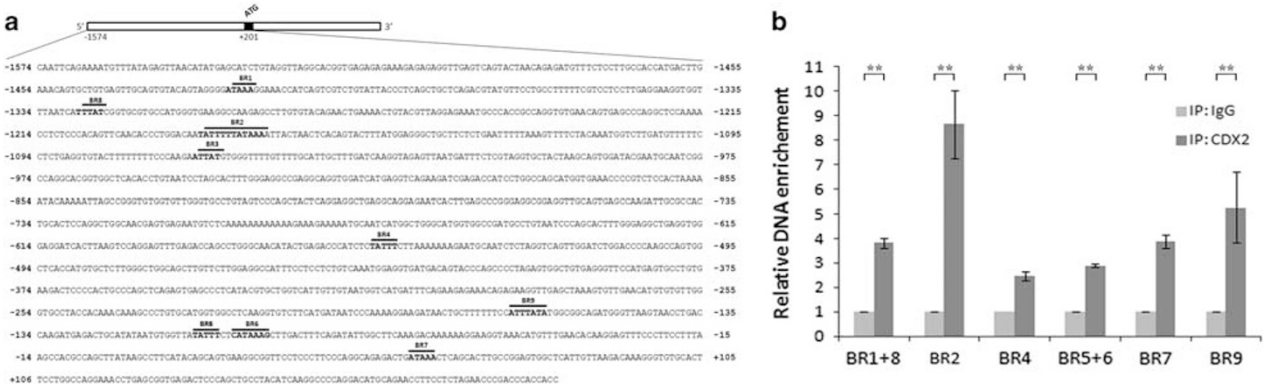

Figure 4 CDX2 binds to sequences upstream of the human *ST6GalNAc-I* gene. (**a**) Analysis of a 1774-bp fragment upstream the *ST6GalNAc-I* start codon (+201). Putative CDX-binding regions (BR) are identified with numbers. Each site was arbitrarily identified as BR1–BR9. (**b**) CDX2 binding to *ST6GalNAc-I* promoter in Caco-2 cells at Day +6 analysed by chromatin immunoprecipitation. Purified DNA was analysed by qPCR using specific primers covering six regions of the *ST6GalNAc-I* promoter. Fold enrichments are expressed as ratios of the IP:CDX2 signal to that of the IP:IgG signal and calculated by extrapolation from a standard curve of input DNA dilutions. The results are expressed as mean ± s.d. of the triplicates from a representative experiment (**P* < 0.05 and ***P* < 0.01).

was used as a negative control. As CDX2 is not present in the normal stomach, no signal was detected. We further tested a fragment of intestinal mucosa, where CDX2, ST6GalNAc-I and STn are expressed in the normal setting. ChIP analysis demonstrated for the first time that CDX2 binds to all the considered regions (Figure 5).

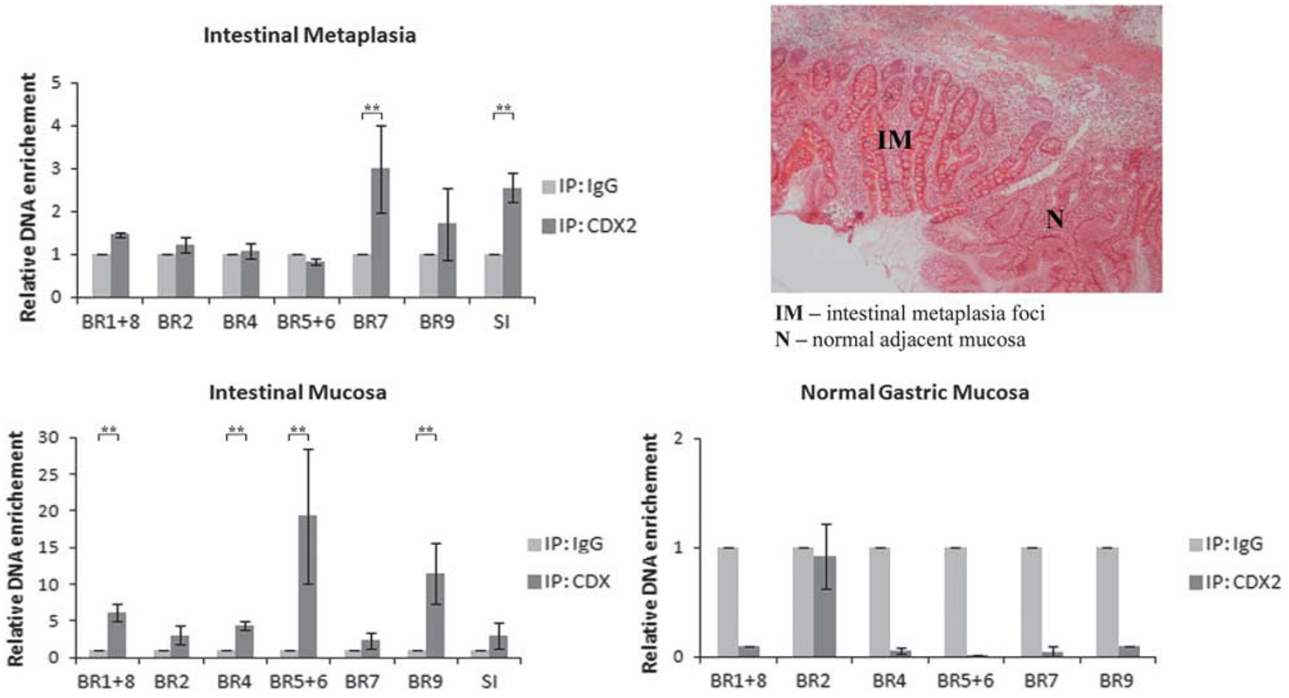

Figure 5 CDX2 binds to sequences upstream of the *ST6GalNAc-I* start codon in human intestinal metaplasia. ChIP was performed *in vivo* in a fragment containing IM foci (H&E image at 40x magnification) and in intestinal mucosa and normal stomach as controls for CDX2 binding. Purified DNA was analysed by qPCR using specific primers covering six regions of the *ST6GalNAc-I* promoter and another set of primers encompassing the Cdx2-binding site on the sucrase-isomaltase (SI) promoter as a positive control, except for normal gastric mucosa, where SI is absent. Fold enrichments are expressed as ratios of the IP:CDX2 signal to that of the IP:IgG signal and calculated by extrapolation from a standard curve of input DNA dilutions. The results are expressed as mean ± s.d. of the triplicates from a representative experiment (*$P<0.05$ and **$P<0.01$).

CDX2 Transactivates ST6GalNAc-I-Regulatory Region in Gastric Carcinoma Cell Lines

After observing that CDX2 binds to all putative CDX-binding sites located upstream of *ST6GalNac-I* coding sequence both in Caco-2 cells and intestinal mucosa, we assessed whether CDX2 was able to regulate *ST6GalNac-I* through its promoter region. Three different luciferase reporter plasmids containing ~1.6, ~0.8 and ~0.3 kb genomic fragments upstream of the human *ST6GalNAc-I* start codon were used for cell transfection and luciferase assays in two gastric carcinoma cell lines, AGS and MKN45, in combination with a CDX2 expression vector or the corresponding control vector. CDX2 promoter was used as a positive control, as it is known that this protein regulates its own promoter.[18] As shown in Figure 6, CDX2 was able to transactivate all the three *ST6GalNAc-I* promoter-tested constructs in both cell lines, the effect being more evident for the MKN45 cell line. A stronger induction was observed for the shorter fragment, suggesting that the most important regulatory sites are mapped in the neighbourhood of the start codon (within the 0.3pST6GalNac-I-Luc construct). We therefore used this smaller fragment to produce a series of mutants where A/T triplets were replaced by C/G triplets, hampering CDX2, to examine the functional role of the three CDX2-binding sites located within the 0.3pST6GalNac-I-Luc construct. The three most proximal sites (BR5, BR6 and BR7) were mutated independently or together by site-directed mutagenesis, in which two of them (BR5 and BR6) were always mutated together owing to their proximity. The wild-type and the three *ST6GalNAc-I* mutant constructs were transfected with a human CDX2 expression vector in AGS and MKN45 cells, and the promoter activity was evaluated by luciferase assay. The results (Figure 6) showed that each independent or combined mutation set led to a decrease in promoter activity when compared with the non-mutated fragment, and this decrease is more evident in the MKN45 cell line. Despite that mutation of the three most proximal sites does not completely abrogate the *ST6GalNac-I*-regulatory region functionality, there is a strong impairment in promoter regulation, namely in the MKN45 cell line, where the transactivation fold of the triple mutant is close to the value for the empty vector (here normalized to 1).

DISCUSSION

In the present study, we have demonstrated the role of an intestine-specific homeoprotein, CDX2, on the transcriptional regulation of *ST6GalNAc-I* gene. We showed that CDX2 is *de novo* expressed in IM,[12] together with ST6GalNAc-I glycosyltransferase and its biosynthetic product, STn. In contrast, in adjacent non-metaplastic gastric mucosa, CDX2 and STn are absent, while ST6GalNAc-I is very weakly expressed.[4] The expression of CDX2 and STn has been independently described in IM[5,12] and in a specific subset of

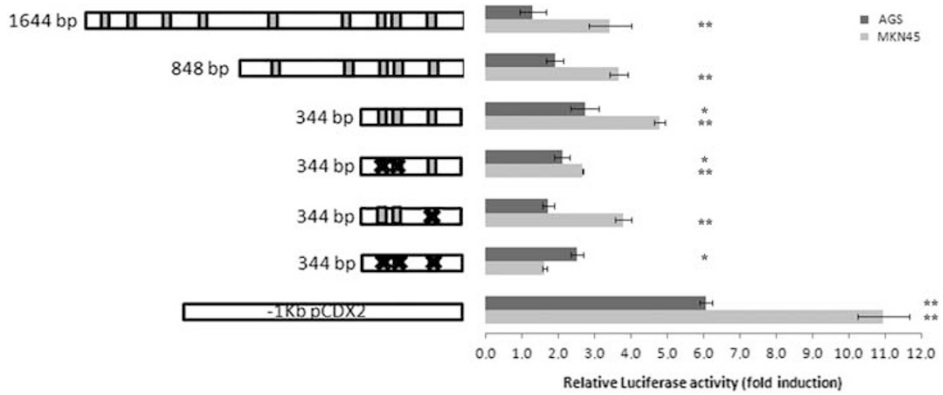

Figure 6 CDX2 transactivates *ST6GalNAc-I*. Three different luciferase reporter plasmids containing ~1.6, ~0.8 and ~0.3 kb genomic fragments upstream of the human *ST6GalNAc-I* start codon were tested independently in two gastric carcinoma cell lines, AGS and MKN45, in combination with a human CDX2 expression vector or the corresponding control empty vector. Another set of fragments comprising independent and combined mutations of the most proximal sites (BR5, BR6 and BR7) were also used. The CDX2-putative-binding regions are identified as small bars in the fragments, and the order corresponds to the one shown in Figure 4a. As a control, CDX2 promoter itself was used. The values obtained were corrected for transfection efficiency with *Renilla* luciferase activity, and results are expressed as fold induction compared with values obtained for the empty vector (here normalized to 1). The results are expressed as mean ± s.d. of at least six values from at least two experiments performed in triplicate for each fragment (*$P<0.05$ and **$P<0.01$).

carcinomas—the mucinous type.[11,30–32] Moreover, previous studies from our group showed that STn is mostly carried by MUC2 mucin in IM, gastric and mucinous carcinomas from several locations.[10,11] Taken together, these observations support the hypothesis that modulation of STn expression by CDX2 might result from direct transcriptional regulation of *ST6GalNAc-I* gene.

In favour of this possibility, we found nine TA-rich putative CDX-binding sites within a 2-kb region upstream the *ST6GalNAc-I* start codon. The human *ST6GalNAc-I* promoter has not been characterized, and we could not identify a TATA box, which is in agreement with the mouse *ST6GalNAc-I* gene.[33] Eight out of the nine putative CDX-binding sites found were shown to recruit the transcription factor both in Caco-2 cell line and in intestinal mucosa. The presence of many binding sites for a transcription factor in a gene promoter has been described and is, in fact, an indication of relevance of that transcription factor on the expression regulation of target genes.[34] CDX2 itself is autoregulated and at least six CDX2-binding sites are identified in a 1.2-Kb promoter.[18] However, it cannot be excluded that the binding of CDX2 to all the tested regions could be originated from a 'neighbourhood effect', if the PCR-analysed fragments were too long after DNA sonication. If the analysed fragments contain two putative sites, they will be amplified and generate false positivity even if CDX2 is only bound to one site. However, this hypothesis is not likely as the analysis of the sonicated DNA on a gel showed that most of the fragments were 100–200 bp in length, thus probably containing only one site.

The analysis of overlapping fragments containing a sequence of 1644 bp upstream the human *ST6GalNAc-I* start codon by luciferase assays, combined with site-directed mutagenesis of putative CDX2-binding sites, further showed that CDX2 is able to transactivate the *ST6GalNac-I*-regulatory region in gastric cell lines, with the effect being more evident for the MKN45 cell line. It is not surprising that AGS cells do not show a similar effect, because the endogenous levels of CDX2 are higher than in MKN45 cells, possibly saturating the system and thus limiting the effect on *ST6GalNac-I* transactivation upon CDX2 overexpression. The results also suggested that the most important regulatory sites are mapped in the vicinity of the start codon, as already observed for other genes regulated by CDX2.[18] Nevertheless, the contribution of more distant regulatory regions[19] cannot be discarded to explain *ST6GalNAc-I* regulation. Also in IM, where regulation of *ST6GalNac-I* by CDX2 holds strong promise as a biological explanation for STn aberrant expression, we demonstrated the preferential recruitment and activity of CDX2 to sites closer to the translation start site. Mutation of the three most proximal sites almost completely abrogate the *ST6GalNac-I*-regulatory region functionality in the MKN45 cell line. Taken together, these observations suggest that CDX2 has a role in the modulation of promoter activity, but other transcription factors may cooperate in *ST6GalNac-I* transactivation. Dynamic cooperation between CDX2 and other partners is not a rare phenomenon and has been demonstrated for other genes.[35–37]

The regulation of glycosyltransferases by CDX2 has been described,[24,27] in particular the transcriptional control of β3Gal-T5 in the Caco-2 intestinal cell line.[24] This cellular model expresses increasing amounts of CDX2[20] and also alters its glycosylation profile and enzyme expression/activity along enterocytic differentiation,[21–26] suggesting that CDX2 can act as a regulator of the glycosylation in this cell line.

Consistent with this hypothesis, we observed that *ST6GalNAc-I* transcript levels increase upon Caco-2 spontaneous *in vitro* differentiation. In addition to *ST6GalNAc-I*, and the already identified *β3Gal-T5*, we also observed an increase in *GalNAc-T3* transcripts in the Caco-2 cell line. Besides, also in the gastric cancer cell line MKN45 CDX2 showed a clear impact on *ST6GalNAc-I* regulation. The same impact is not detected in AGS cells upon CDX2 down-regulation. In this model, *ST6GalNAc-I* levels were not consistently decreased over time, leading to the supposition that *ST6GalNAc-I* regulation is not entirely dependent on CDX2, as described for other CDX2 targets.[35–37] The dramatic decrease of CDX2 levels 72 h posttransfection might activate compensatory mechanisms by alternative transcription factors acting on *ST6GalNAc-I* regulation.

However, we could not detect an increase in STn expression simultaneously to *ST6GalNAc-I* transcript levels, when CDX2 was overexpressed in MKN45 for 72 h (data not shown). The incomplete overlap between ST6GalNAc-I mRNA detection and STn might be explained by competition for the GalNAc substrate between ST6GalNAc-I and other glycosyltransferases, such as Core1 Gal-transferase. It is also possible that the kinetics of STn biosynthesis by ST6GalNAc-I is slower and takes more time to show an effect in STn expression. Alternative possibilities for aberrant STn expression in cancer contexts include mutations or promoter methylation of the molecular chaperone *Cosmc*,[38,39] aberrant localization of glycosyltransferases in the ER/Golgi,[40,41] or the activity of another sialyltransferase from the same sub-family as ST6GalNAc-I—ST6GalNAc-II[4]—have been demonstrated. This enzyme, ST6GalNAc-II, was also shown to have some specificity *in vitro* for GalNAc substrates *O*-linked to proteins (Tn antigen). Moreover, a lack of concordance observed between ST6GalNAc-I and STn expression in gastric carcinomas has been observed in colorectal carcinomas,[42] and it is known that cell lines and tissues lacking STn may still show low levels of *ST6GalNAc-I* mRNA,[43] indicating that alternative biosynthetic mechanisms may be at play.

Understanding how glycosylation is controlled is indispensable for a better understanding of cancer initiation and progression, as expression of specific *O*-glycan structures is closely related to cell differentiation. Our study identified a new mechanism of glycosylation control in gastric carcinogenesis by showing that CDX2 has a role in *ST6GalNAc-I* gene regulation in IM and some cell line models. For the first time, we demonstrated that a single homeobox protein transcriptionally regulates a glycoprotein profile associated with IM and gastric cancer.

Supplementary Information accompanies the paper on the Laboratory Investigation website (http://www.laboratoryinvestigation.org)

ACKNOWLEDGMENTS

We are grateful to Patrícia Marques and Susana Seixas for woman normal blood extracted DNA used for ST6GalNAc-I promoter fragment amplification; Jean-Noel Freund for the information about CDX2-putative-binding site search on human *ST6GalNAc-I* gene in the TRANSFAC database and Henrik Clausen for fruitful discussions. This work was supported by Fundação para a Ciência e a Tecnologia (FCT), (SFRH/BD/69983/2010) and Project NORTE (07-0124-FEDER-000024) co-financed by Programa Operacional Regional do Norte, under Quadro de Referência Estratégica Nacional (QREN). IPATIMUP—Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education—is partially supported by FCT.

DISCLOSURE/CONFLICT OF INTEREST

The authors declare no conflict of interest.

1. Ikehara Y, Kojima N, Kurosawa N *et al.* Cloning and expression of a human gene encoding an N-acetylgalactosamine-alpha2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. Glycobiology 1999;9:1213–1224.
2. Julien S, Krzewinski-Recchi MA, Harduin-Lepers A *et al.* Expression of sialyl-Tn antigen in breast cancer cells transfected with the human CMP-Neu5Ac: GalNAc alpha2,6-sialyltransferase (ST6GalNac I) cDNA. Glycoconj J 2001;18:883–893.
3. Marcos NT, Pinho S, Grandela C *et al.* Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. Cancer Res 2004;64:7050–7057.
4. Marcos NT, Bennett EP, Gomes J *et al.* ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues. Front Biosci (Elite Ed) 2011;3:1443–1455.
5. David L, Nesland JM, Clausen H *et al.* Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases. APMIS Suppl 1992;27:162–172.
6. Yonezawa S, Tachikawa T, Shin S *et al.* Sialosyl-Tn antigen. Its distribution in normal human tissues and expression in adenocarcinomas. Am J Clin Pathol 1992;98:167–174.
7. Victorzon M, Nordling S, Nilsson O *et al.* Sialyl Tn antigen is an independent predictor of outcome in patients with gastric cancer. Int J Cancer 1996;65:295–300.
8. Mesquita P, Jonckheere N, Almeida R *et al.* Human MUC2 mucin gene is transcriptionally regulated by Cdx homeodomain proteins in gastrointestinal carcinoma cell lines. J Biol Chem 2003;278:51549–51556.
9. Yamamoto H, Bai YQ, Yuasa Y. Homeodomain protein CDX2 regulates goblet-specific MUC2 gene expression. Biochem Biophys Res Commun 2003;300:813–818.
10. Conze T, Carvalho AS, Landegren U *et al.* MUC2 mucin is a major carrier of the cancer-associated sialyl-Tn antigen in intestinal metaplasia and gastric carcinomas. Glycobiology 2010;20:199–206.
11. Pinto R, Carvalho AS, Conze T *et al.* Identification of new cancer biomarkers based on aberrant mucin glycoforms by in situ proximity ligation. J Cell Mol Med 2012;16:1474–1484.
12. Almeida R, Silva E, Santos-Silva F *et al.* Expression of intestine-specific transcription factors, CDX1 and CDX2, in intestinal metaplasia and gastric carcinomas. J Pathol 2003;199:36–40.
13. Bai YQ, Yamamoto H, Akiyama Y *et al.* Ectopic expression of homeodomain protein CDX2 in intestinal metaplasia and carcinomas of the stomach. Cancer Lett 2002;176:47–55.
14. Eda A, Osawa H, Satoh K *et al.* Aberrant expression of CDX2 in Barrett's epithelium and inflammatory esophageal mucosa. J Gastroenterol 2003;38:14–22.
15. Osawa H, Kita H, Satoh K *et al.* Aberrant expression of CDX2 in the metaplastic epithelium and inflammatory mucosa of the gallbladder. Am J Surg Pathol 2004;28:1253–1254.
16. Hinoi T, Lucas PC, Kuick R *et al.* CDX2 regulates liver intestine-cadherin expression in normal and malignant colon epithelium and intestinal metaplasia. Gastroenterology 2002;123:1565–1577.
17. Suh E, Chen L, Taylor J *et al.* A homeodomain protein related to caudal regulates intestine-specific gene transcription. Mol Cell Biol 1994;14:7340–7351.
18. Barros R, da Costa LT, Pinto-de-Sousa J *et al.* CDX2 autoregulation in human intestinal metaplasia of the stomach: impact on the stability of the phenotype. Gut 2011;60:290–298.

19. Boyd M, Hansen M, Jensen TG *et al.* Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). J Biol Chem 2010;285: 25115–25125.
20. Pinto M, Robine-Leon S, Appay M *et al.* Enterocyte-like differentiation and polarization of the human colon carcinoma cell line Caco-2 in culture. Biol Cell 1983;47:323–330.
21. Amano J, Oshima M. Expression of the H type 1 blood group antigen during enterocytic differentiation of Caco-2 cells. J Biol Chem 1999;274:21209–21216.
22. Brockhausen I, Romero PA, Herscovics A. Glycosyltransferase changes upon differentiation of CaCo-2 human colonic adenocarcinoma cells. Cancer Res 1991;51:136–142.
23. Dall'Olio F, Malagolini N, Guerrini S *et al.* Differentiation -dependent expression of human beta-galactoside alpha 2,6-sialyltransferase mRNA in colon carcinoma CaCo-2 cells. Glycoconj J 1996;13:115–121.
24. Isshiki S, Kudo T, Nishihara S *et al.* Lewis type 1 antigen synthase (beta3Gal-T5) is transcriptionally regulated by homeoproteins. J Biol Chem 2003;278:36611–36620.
25. Youakim A, Herscovics A. Differentiation-associated decrease in the proportion of fucosylated polylactosaminoglycans of CaCo-2 human colonic adenocarcinoma cells. Biochem J 1987;247:299–306.
26. Youakim A, Romero PA, Yee K *et al.* Decrease in polylactosaminoglycans associated with lysosomal membrane glycoproteins during differentiation of CaCo-2 human colonic adenocarcinoma cells. Cancer Res 1989;49:6889–6895.
27. Sakuma K, Aoki M, Kannagi R. Transcription factors c-Myc and CDX2 mediate E-selectin ligand expression in colon cancer cells undergoing EGF/bFGF-induced epithelial-mesenchymal transition. Proc Natl Acad Sci USA 2012;109:7776–7781.
28. Matys V, Fricke E, Geffers R *et al.* TRANSFAC: transciptional regulation, from patterns to profiles. Nucleic Acids Res 2003;31:374–378.
29. Bu XD, Li N, Tian XQ *et al.* Caco-2 and LS174T cell lines provide different models for studying mucin expression in colon cancer. Tissue Cell 2011;43:201–206.
30. Fraggetta F, Pelosi G, Cafici A *et al.* CDX2 immunoreactivity in primary and metastatic ovarian mucinous tumours. Virchows Arch 2003;443: 782–786.
31. Rossi G, Murer B, Cavazza A *et al.* Primary mucinous (so-called colloid) carcinomas of the lung: a clinicopathologic and immunohistochemical study with special reference to CDX-2 homeobox gene and MUC2 expression. Am J Surg Pathol 2004;28:442–452.
32. Kaimaktchiev V, Terracciano L, Tornillo L *et al.* The homeobox intestinal differentiation factor CDX2 is selectively expressed in gastrointestinal adenocarcinomas. Mod Pathol 2004;17:1392–1399.
33. Kurosawa N, Takashima S, Kono M *et al.* Molecular cloning and genomic analysis of mouse GalNAc alpha2, 6-sialyltransferase (ST6GalNAc I). J Biochem 2000;127:845–854.
34. Hawkes HJ, Karlenius TC, Tonissen KF. Regulation of the human thioredoxin gene promoter and its key substrates: a study of functional and putative regulatory elements. Biochim Biophys Acta 2014;1840: 303–314.
35. Mitchelmore C, Troelsen JT, Spodsberg N *et al.* Interaction between the homeodomain proteins Cdx2 and HNF1α mediates expression of the lactase-phlorizin hydrolase gene. Biochem J 2000;346:529–535.
36. Shimakura J, Terada T, Shimada Y *et al.* The transcription factor Cdx2 regulates the intestine-specific expression of human peptide transporter 1 through functional interaction with Sp1. Biochem Pharmacol 2006;71:158–158.
37. Verzi MP, Shin H, He HH *et al.* Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. Dev Cell 2010;16: 713–726.
38. Ju T, Lanneau GS, Gautam T *et al.* Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc. Cancer Res 2008;68: 1636–1646.
39. Mi R, Song L, Wang Y *et al.* Epigenetic silencing of the chaperone *Cosmc* in human leukocytes expressing Tn antigen. J Biol Chem 2012;287:41523–41533.
40. Gill DJ, Chia J, Senewiratne J *et al.* Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes. J Cell Biol 2010;189:843–858.
41. Hassinen A, Pujol FM, Kokkonen N *et al.* Functional organization of Golgi N- and O-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells. J Biol Chem 2011;286:38329–38340.
42. Vázquez-Martín C, Cuevas E, Gil-Martín E *et al.* Correlation analysis between tumor-associated antigen sialyl-Tn expression and ST6GalNAc I activity in human colon adenocarcinoma. Oncology 2004;67:159–165.
43. Marcos NT, Cruz A, Silva F *et al.* Polypeptide GalNAc-transferases, ST6GalNAc-transferase I, and ST3Gal-transferase I expression in gastriccarcinoma cell lines. J Histochem Cytochem 2003;51:761–771.