Published in final edited form as:

Cancer Immunol Immunother. 2013 March ; 62(3): 423–435. doi:10.1007/s00262-012-1325-2.

Human Mucin MUC1 RNA Undergoes Different Types of Alternative Splicing Resulting in Multiple Isoforms

Lixin Zhang^{1,*}, Anda Vlad^{1,2}, Christine Milcarek¹, and Olivera J Finn^{1,*}

¹Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, U.S.A

²Department of Obstetrics and Gynecology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, U.S.A

Abstract

MUC1 is a transmembrane mucin with important functions in normal and transformed cells carried out by the extracellular domain or the cytoplasmic tail. A characteristic feature of the MUC1 extracellular domain is the variable number of tandem repeats (VNTR) region. Alternative splicing may regulate MUC1 expression and possibly function. We developed an RT-PCR method for efficient isolation of MUC1 mRNA isoforms that allowed us to evaluate the extent of alternative splicing of MUC1 and elucidate some of the rules that govern this process. We cloned and analyzed 21, 24, and 36 isoforms from human tumor cell lines HeLa, MCF7, and Jurkat respectively, and 16 from normal activated human T cells. Among the 78 MUC1 isoforms we isolated, 76 are new and different cells showed varied MUC1 expression patterns. The VNTR region of exon 2 was recognized as an intron with a fixed 5' splice site but variable 3' splice sites. We also report that the 3506 A/G SNP in exon 2 can regulate 3' splice sites selection in intron 1 and produce different MUC1 short isoform proteins. Furthermore, the SNP A to G mutation was also observed in vivo, during de novo tumor formation in MUC1^{+/-}Kras^{G12D/+}Pten^{loxP/loxP} mice. No specific functions have been associated with previously reported short isoforms. We now report that one new G SNP-associated isoform MUC1/Y-LSP, but not the A SNP-associated isoform MUC1/Y, inhibits tumor growth in immunocompetent but not immunocompromised mice.

Keywords

MUC1; alternative splicing; isoforms; tumor

INTRODUCTION

Alternative splicing of eukaryotic pre-mRNAs can enhance the proteome diversity and contribute to genome evolution and gene regulation [1–4]. Up to 95% of human pre-mRNAs that contain more than one exon are alternatively spliced to yield multiple mRNA isoforms [5–7]. Aside from its role in physiological adaptation of cells [8], alternative splicing has been shown to occur in human diseases, including cancer [9–18]. Detection of splice variants is considered more reliable than transcription-level differences in discriminating tumor from normal tissue [19,20].

*To whom correspondence should be addressed: Mailing address: E1001 Biomedical Science Tower, 200 Lothrop St, Pittsburgh, PA 15261, Tel: 412-648-9816; Fax: 412-648-7042; Liz36@pitt.edu; ojfinn@pitt.edu.

Conflict of interest: The authors declare that they have no conflict of interest.

Supplementary data available on-line.

The mucin family of epithelial proteins currently comprises 20 members (MUC1-MUC20) [21]. MUC1, the first identified and most studied mucin is a high molecular weight transmembrane glycoprotein that is overexpressed and hypoglycosylated in most adenocarcinomas, including breast, lung, colon, pancreas, stomach, prostate and ovarian tumors [22]. Though it is encoded as a single protein, MUC1 is expressed as a type I transmembrane heterodimer with a large extracellular α -subunit non-covalently bound to a smaller transmembrane β -subunit [23]. The *MUC1* gene is located on the human chromosome 1q21 and contains seven exons. The extracellular fragment encoded by exons 1–4 is dominated by the VNTR domain located in exon 2. Exon 2 consists of tandemly repeated (between 20 and 200 times depending on genotype) units of 60 base pairs encoding a proline-rich sequence PDTRPAPGSTAPPAHGVTS. The rigid structure of this protein domain, which we previously analyzed by 1D and 2D NMR [24,25], can extend 300–500 nm above the cell surface, towering over other cell surface molecules. While the MUC1 extracellular domain plays important roles in cell-cell and cell-matrix interactions [26,27] it has also been shown to regulate chemotaxis of cells of the innate immune system [28]. The smaller β -subunit of MUC1 (~ 20 kDa) encoded by exons 4–7, contains a short extracellular portion, a transmembrane region and a short cytoplasmic tail with seven phosphorylation sites that support the role of MUC1 as a signaling molecule [21,29].

MUC1 was characterized first by us [30] and later by others as a tumor antigen and a target for immunotherapy, as well as a molecule with immunosuppressive [28,31] and tumor promoting functions [32,33]. MUC1 can promote malignant transformation by regulating the function of β -catenin [34] and p53 [35], and by repressing a microRNA (miR-322) via a feed-back loop [36]. All of these functions have been ascribed to the whole molecule (MUC1/TR) or the truncated form consisting of the transmembrane domain and the cytoplasmic tail. The well-known MUC1 isoforms are the full length form encoded by all 7 exons with the VNTR region (MUC1/TR) and the short isoforms lacking VNTR, such as MUC1/X, MUC1/Y, MUC1/Z, and MUC1/ZD [37–40]. The overall prevalence, mechanism of generation, and individual contributions, if any, of MUC1 isoforms to the known functions of MUC1 are not defined. We report here results from cloning and sequencing of a large number of MUC1 mRNA short isoforms from tumor cells and normal cells using our optimized RT-PCR-based method that allowed us to compare the expression pattern of MUC1 isoforms in different cells and to identify specific alternative splicing events that led to their generation. We discuss possible biological roles and, in the case of MUC1Y-LSP we isolated, show that it can interfere with tumor growth *in vivo*.

MATERIALS AND METHODS

Ethics Statement

Mice were maintained or bred in specific pathogen-free conditions at the University of Pittsburgh and treated under the guidelines in accordance with the protocol No. 0905822B-3 approved by the Institutional Animal Care and Use Committee.

Cell culture and flow cytometry

All cell lines were purchased from the American Type Culture Collection (ATCC) unless indicated. Human tumor cell lines HeLa (cervical cancer), MCF7 (breast cancer), Jurkat (T lymphoma), and DM6 (melanoma) were cultured in RPMI 1640 (GIBCOL BRL) medium supplemented with 10% fetal bovine serum (GIBCOL BRL), 10000U/L penicillin (Sigma), 10000 g/L streptomycin (Sigma) and 2mmol/L L-glutamine (GIBCOL BRL). Human PBMCs from healthy donors were obtained from the Central Blood Bank of Pittsburgh and cultured in RPMI 1640 supplemented with 10% human AB serum, 10000U/L penicillin, 10000 g/L streptomycin and 2mmol/L L-glutamine, 20000U/L IL-2, 10mg/L

Phytohemagglutinin (PHA), at 37°C, 6% CO₂ for 3 days. RMA mouse T lymphoma cell line[18] was cultured in DMEM (GIBCOL BRL) supplemented with 10% fetal bovine serum (GIBCOL BRL), 10000U/L penicillin (Sigma), 10000 g/L streptomycin (Sigma) and 2mmol/L L-glutamine (GIBCOL BRL), 1% NEAA (GIBCOL BRL), 1mM sodium pyruvate, 0.1 mM beta- mercaptoethanol (Gibco). A total of 2×10^5 cells were used for flow cytometry analysis of MUC1 expression. In brief, cells were suspended in staining buffer (1% BSA in PBS, pH7.4) and FITC mouse anti-Human CD227(MUC1) (BD) or FITC mouse IgG1(BD) were added and incubated on ice for 30min; wash 2x with staining buffer and samples were run on BD LSRII flow cytometer.

Primers

The following primers were used for the PCR reactions: P1S: 5'-GCG CCT GCC TGA ATC TGT TC-3', P1A: 5'-CCC ACA TGA GCT TCC ACA CAC-3', P2S: 5'-CTC CCC ACC CAT TTC ACC AC-3', P2A: 5'-GGC GAC GTG CCC CTA CAA GT-3'. pcDNA-S: 5'-CCA CTG CTT ACT GGC TTA TCG A-3'. TR-S: 5'-GCC TGA ATC TGT TCT GCC C-3'. TR-A: 5'-CCT GAA GCT GGT TCC GTG-3'. MUCGA: 5'-GCT TGC ATG ACC AGA ACC C-3', MUCAA: 5'-GCT TGC ATG ACC AGA ACC T-3', -actin-S: 5' ACA CTG TGC CCA TCT ACG ACC 3', -actin-A: 5' AGG GGC CGG ACT CGT CAT AGA 3', GAPDH-S: 5'-GAA GGT GAA GGT CGG AGT C-3' GAPDH-A: 5'-GGA AGA TGG TGA TGG GAT TTC-3'. All primers were ordered from Integrated DNA Technologies, Inc (Coralville, Iowa).

Amplification and cloning of MUC1 isoforms

Trizol™ reagent (Invitrogen) or AllPrep DNA/RNA Mini Kit (Qiagen) was used to extract total RNA and genomic DNA from HeLa, MCF7, Jurkat cells, and PHA-activated human T cells according to the protocol provided by the manufacturer. Reverse transcription was performed using SuperScript™ First-Strand Synthesis System (Invitrogen) with Oligo (dT)₁₂₋₁₈ primers according to the instruction manual.

The method takes advantage of the first and last exons being constitutive exons included in most isoforms. Two pairs of primers binding to the first and last exons were designed. One pair (P1S-P1A) for the first PCR amplification of different possible alternatively spliced variants from reverse transcribed products, and the second pair (P2S-P2A) for eliminating the false positives of the first round using colony PCR. The PCR conditions were as following: First PCR: 10 × ExTaq (TaKaRa Bio Inc., Japan) buffer 5μL, dNTP (2.5mM) 5μL, P1S (10 μM) and P1A (10μM) primer mixture 5μL, cDNA 2μL, ExTaq 0.5μL, H₂O 32.5μL, 94°C, 5min; (94°C, 30sec.; 65°C, 30sec.; 72°C, 2min;) × 30–35; 72°C, 7min. PCR for β-actin: 10× ExTaq buffer 2 μL, dNTP (2.5mM) 1.8 μL, β-actin-S (10μM) 1 μL, β-actin-A (10μM) 1μL, cDNA 1μL, ExTaq 0.5μL, H₂O 12.7μL, 94°C, 5min, (94°C, 30sec.; 65°C, 30sec.; 72°C, 1 min;) × 35; 72°C, 7min. PCR for GAPDH: 10 × ExTaq buffer 2μL, dNTP (2.5mM) 1.8μL, GAPDH-S (10μM) 1μL, GAPDH-A (10μM) 1μL, cDNA 1μL, ExTaq 0.5μL, H₂O 12.7μL, 94°C, 5min, (94°C, 30 sec.; 65°C, 30 sec.; 72°C, 1 min;) × 35; 72°C, 7 min.

After 30 or 35-rounds of PCR, bands were resolved on a 1% TAE-agarose gel (Supplemental Fig 1). Individual bands were recovered and purified with Promega Wizard PCR Preps DNA Purification System or QIAquick Gel Extraction Kit, cloned into pGEM-T (Promega) or pCR®4-TOPO® vector (Invitrogen) according to the instructions provided, and transformed into DH5α competent cells. Positive colonies were selected in Luria—Bertani (LB) 1% agar plates containing 100 g/ml ampicillin at 37°C after 16–20 hrs.

To eliminate colonies that contained non-specific PCR products, 20 colonies from each transformation were picked for colony PCR analysis. 10× ExTaq buffer 2μL, dNTP (2.5mM) 2μL, P2-S (10μM) and P2-A (10μM) mixture 2μL, ExTaq 0.1μL and H₂O 13.9μL were mixed together. The transformed *E. coli* colonies were sampled and pipetted into the PCR mixtures. The PCR condition was as follow: 94°C, 5min., (94°C, 30sec.; 65°C, 30sec.; 72°C, 1 min;) × 25; 72°C, 7min. All PCR products were run on 1% or 2% TAE-agarose gel (Supplemental Fig 2).

Sequencing of PCR products

Transformants confirmed by colony PCR were amplified in LB medium with 100 g/ml ampicillin at 37°C and recombinant plasmids were purified with QIAprep Spin Miniprep Kit. DNA fragments were amplified from genomic DNA(100ng) from different cell lines using TR-S and TR-A primers: 10 × ExTaq buffer 5μL, dNTP (2.5mM) 5μL, TR-S (10μM) and TR-A (10μM) primer mixture 5μL, cDNA 2μL, ExTaq 0.5μL, H₂O 32.5μL, 94°C, 5min; (94°C, 40sec.; 58°C, 40sec.; 72°C, 1min;) × 35; 72°C, 7min. PCR fragments were purified with QIAprep Kit. All DNA sequencing was done at the DNA Sequencing Facility of University of Pittsburgh. The sequences were analyzed by searching BLASTn using MUC1 gene (GenBank No. M61170) as a template.

Cloning of MUC1 short isoforms generated by MUC1/TR

pcDNA3 plasmids (Invitrogen) containing MUC1/42TR, were transfected into HeLa cell lines using Lipofectamine™2000 (Invitrogen) according to the manufacturer's instruction manual. Twenty-four hours after transfection, total RNA were isolated and RT-PCRs were done using primers pcDNA-S and P2A. Bands were recovered and cloned into pCR®4-TOPO® vector (Invitrogen) for sequence analysis as showed previously.

PCR analysis of MUC1 3506 A/G SNP

P2S+MUCGA and P2S+MUCAA primers were used for amplifying 3506 G and 3506 A specific fragments, respectively. PCR condition: 94°C, 5min., (94°C, 30sec.; 62°C, 30sec.; 72°C, 1min;) × 38; 72°C, 5min. Distilled water was used as PCR negative control.

Western blot analysis

DM6 or RMA cell lines were stably transfected with pcDNA3-MUC1/Y, pcDNA3-MUC1/Y-LSP and pcDNA3- MUC1/22TR, pcDNA3-MUC1/S1, and pcDNA3-MUC1/S2 using Lipofectamine 2000 or electroporation. Cells were washed with PBS twice and lysed with 1x SDS loading buffer. Proteins were separated on 4–20% gradient polyacrylamide/SDS gels and transferred onto PVDF membranes at 36V overnight. The membranes were blocked with 5% skimmed milk for 1 hour at room temperature, incubated with anti-MUC1 monoclonal antibody CT-2 which recognizes a 17-aa long epitope SSLSYTNPAVAATSANL (Fig 5b) in the MUC1 cytoplasmic tail (gift from Dr. Sandra Gendler, Mayo Clinic, Scottsdale) or mouse anti-human β-actin antibody (Sigma) for 1 hour at room temperature. The signals were detected using SuperSignal West Femto Maximum Sensitivity Substrate (PIERCE) and images were captured on the Kodak Image Station 2000.

De novo tumor formation in conditional mice

We generated triple transgenic MUC1^{+/-}Kras^{G12D/+}Pten^{loxP/loxP} mice by crossing MUC1^{+/-} transgenic mice (purchased from Dr. S. Gendler, Mayo Clinic, Scottsdale, AZ) that carry 1–2 copies of MUC1 gene and express human MUC1 under the endogenous promoter [41] with the previously described Kras^{G12D/+}Pten^{loxP/loxP} mice [42]. Genotyping was performed as previously described [43,44].

To generate tumors, Cre recombinase encoding adenovirus [Ad5CMVCre (AdCre)] obtained from the University of Iowa Gene Transfer Vector Core, was administered either under the ovarian bursa (IB) or intraperitoneally (IP) as described by us and others [45,42]. For IB injections, the mice received 5 l of 2.5×10^7 plaque-forming units (p.f.u.) of Ad5CMVCre delivered to the ovarian surface epithelium (OSE) of left ovary only, during survival surgery. The contra-lateral ovary served as control. For the intraperitoneal experiments we administered 100 l of AdCre (5×10^8 p.f.u.) per animal.

Mice were sacrificed when moribund (hunched appearance, ruffled fur, unable to reach for food or water). Tumors from the primary and metastatic sites were harvested and processed for DNA extraction as well as for paraffin embedding (histology).

Tumor growth *in vivo*

1×10^5 transfected or untransfected RMA cells were s.c. injected into the thighs of C57BL/6 (6–8 weeks), NOD.CB17-*Prkdc*^{scid}/J (6–8 weeks) (The Jackson Laboratory, Bar Harbor, Maine) or MUC1.Tg mice (6–8 weeks, Mayo Clinic, Scottsdale, AZ). Mice were sacrificed within 2–3 weeks and tumors were measured.

In vivo tumor cell killing assay

In brief, 1×10^7 RMA, RMA-Y-LSP, RMA-Y and RMA-22TR cells in PBS containing 1% FBS were labeled with 1 μ M (low) or 10 μ M (high) CFSE (Invitrogen) at 37 °C for 15 minutes, then transferred on ice and in 10 ml cold complete DMEM for additional 10 minutes. RMA-low and RMA-Y-LSP-high, RMA-low and RMA-Y-high, RMA-low and RMA-22TR-high were mixed together at 1:1 ratio. A total of 1×10^6 mixed cells were injected i.p. into MUC1.Tg mice. Surviving cells were recovered after 44 hours and flow analysis was run to analyze CFSE-low versus CFSE-high cell populations on LSR II (BD Biosciences).

RESULTS

Tumor cell lines express a large variety of MUC1 isoforms

To select candidate cell lines for our molecular studies on MUC1 isoforms, we stained several cell lines for MUC1 expression. Flow cytometry analysis using an antibody binding to the VNTR showed that almost all epithelial MCF7 (99.6%) and HeLa (98.1%) cells and a portion of activated T cells (24.2%) expressed MUC1/TR on the cell surface, while very few if any of the T cell leukemia Jurkat cells (1.5%) were MUC1 positive. However, in Jurkat cells, MUC1/TR seems to be confined to the cytoplasm, as intracellular staining revealed that 84.6% of cells were MUC1/TR+ (Fig. 1). These different levels and patterns of expression provided a good setting in which to explore the extent and diversity of alternative splicing of MUC1.

Total RNAs were isolated from the tumor cell lines for RT-PCR amplification and cloning of MUC1 short isoforms. Analysis of cloned and sequenced cDNA of MUC1 from the three tumor cell lines yielded a total of 68 isoforms, 21 from HeLa (Fig. 2a), 24 from MCF7 (Fig. 2b), and 36 from Jurkat cells (Fig. 2c). Among the 68 isoforms, the previously reported MUC1/Y and MUC1/Z were also included. These two isoforms, as well as MUC1/Y-LSP which is one of the new isoforms we isolated, were seen in all three tumor lines. Of the 68 total, the open reading frames (ORFs) of 31 isoforms are in frame with the full length MUC1/TR. Twenty one of these have the potential to encode transmembrane forms while 10 could encode only secreted proteins, having spliced out exon 5 that encodes the transmembrane domain.

Premature termination codons (PTCs) were observed in 37 isoforms (38.1% in HeLa, 12.5% in MCF-7 and 75% in Jurkat), occurring more than 50 nucleotides upstream of the final splice junction, suggesting that these isoforms may be the potential target for nonsense-mediated mRNA decay (NMD), a surveillance mechanism that selectively degrades mRNAs that would encode premature transcription termination codons [46].

MUC1 isoforms are generated through different alternative splicing events

We examined the sequence of the cloned MUC1 isoforms and identified that they could be generated through different types of alternative splicing, including exon skipping, alternative 5' splice sites, alternative 3' splice sites and intron retention. Only exon 1 (E1) did not seem to be alternatively spliced. The most frequent alternative splicing event was exon skipping that involves exons E3, E4 and E5, with E4 having the highest skipping rate. E6 skipping occurred only in one isoform (MUC1/M16) derived from MCF7. Intron retention involving intron 1 (full) and intron 6 (partial) was detected in Jurkat cells. Only one isoform (MUC1/M11) from MCF7 showed partial intron 6 retention.

Both conventional and cryptic splice sites are used in alternative splicing of the MUC1 gene. To produce the full-length MUC1/TR containing all 7 exons, the conventional splice sites in the six introns are used. In intron 1, the cryptic splice sites used are 5' site 3222 and 3096 and 3' splice sites 3252, 3261, 3471, 3507 and 3534. Usage of the 3' splice site 4945 located in exon 5 is always associated with exon skipping, including E4 skipping, E3–E4 double-skipping, and E3–E4 plus part of E2 skipping. This can be observed in HeLa and MCF7, but not in Jurkat cells. In intron 6, different 5' and 3' cryptic splice sites usage can cause partial intron retention, mostly observed in Jurkat cells yielding isoforms MUC1/J1, J10, J12, J13 and J14. The VNTR region in E2 was spliced out in all MUC1 short isoforms using a single cryptic 5' splice site at position 3600 and multiple cryptic 3' splice sites at positions 4205, 4235, 4259, 4278, 4192 and 4120. These results suggest that tumor cells can produce vastly different forms of MUC1 simultaneously by multiple alternative splicing pathways.

Differences and similarities in the expression of MUC1 isoforms in normal T cells versus epithelial tumor cells

We [47] and others [48,49] have shown that primary T cells express MUC1 upon activation. Here, we used PHA activated T cells to test if MUC1 mRNA isoforms are made by normal cells and what, if any, differences there might be in the specific alternative splicing mechanisms between normal and tumor cells. cDNAs from PHA activated T cells from a healthy donor were used to isolate MUC1 mRNA isoforms using the same PCR method as for tumor cells. In all, 16 new isoforms were cloned (Fig 2c). Our analysis revealed that the same mechanisms seen in tumor cells, exon skipping, alternative 5' splice sites, and alternative 3' splice sites, were also used by primary T cells to produce these isoforms, with the exception of complete intron retention. The VNTR region is spliced out as in tumor cells with the same 5' splice site, some common 3' splice sites (4205, 4259, 4278, 4120, 4192) and one unique 3' splice site (3934) whose product (MUC1/T1) has a part of a degenerate VNTR (dTR) region. The 3' splice sites at positions 4120 and 4192 and the 3096 5' splice site were only used by Jurkat (T cell lymphoma) and normal T cells. Normal T cells share other isoforms with Jurkat but not with HeLa and MCF7 cells, including MUC1/G-LSP, T4, T11, T12, suggesting that these isoforms may be T cell specific. Eight of them (50%) have PTCs (premature termination codons) and may not produce protein products. Of the 78 MUC1 short isoforms (submitted to Genbank) isolated from the tumor cell lines and activated T cells, 43 (55.1%) have PTCs. When compared with those from tumor cells, MUC1 isoforms from T cells have a lower rate of exon skipping (12.5% vs 71.4% in HeLa, 78.3% in MCF7, and 37.1% in Jurkat).

Overall, our results, plotted as a Venn Diagram, show that while many of the isoforms are shared between either all cells or all tumor cells, 57.1% in HeLa, 58.3% in MCF7, 75.0% in Jurkat, and 62.5% in the activated T cells are unique to that cell type, suggesting potential regulation of cell type-specific MUC1 RNA processing and potential functions (Fig. 3).

The VNTR-containing region of exon 2 is recognized as an intron

A common feature of all the short MUC1 isoforms is that the VNTR region is spliced out, implying that this region could be recognized by the spliceosome as an intron. To elucidate whether the VNTR region can be spliced out in an already spliced mature MUC1 mRNA containing all seven exons, we transfected the pcDNA3 expression vector containing MUC1 cDNA with 42 TRs (pcDNA3-MUC1/42TR) into HeLa cells. In order to isolate mRNA products from transcripts produced by the plasmid encoded MUC1 and not from the endogenous MUC1, we used a pair of primers combining one that binds to the MUC1 sequence and the other to the vector sequence. If the VNTR region is recognized as an intron, we expected the transcripts of MUC1/42TR to be spliced to generate short isoforms devoid of VNTRs. The results (Fig. 4) showed that short MUC1 mRNA isoforms MUC1/Y and MUC1/G were among the products generated by pcDNA3-MUC1/42TR, indicating that mature MUC1/TR mRNAs may be further alternatively spliced after being produced from the MUC1 pre-mRNA and thus could serve as another source for generating short isoforms in addition to the MUC1 pre-mRNA.

3506 A/G SNP can regulate production of MUC1 short isoforms and can change during cell transformation

In exon 2 of the MUC1 gene (GenBank No. M61170), there is an A/G SNP (ref SNP ID: rs4072037) at nucleotide 3506 that was previously reported to dictate the 3' splice site selection of 3498 (when it is A) versus 3471 (when it is G) in the full length MUC1/TR in tumor [50] as well as MUC1/TR and MUC1 devoid of TR in normal cells [51], resulting in a 27 bp length difference in the signal peptide-encoding region. We wanted to see if this rule also held in the generation of MUC1 short isoforms devoid of TR region in the cells we used and whether there were any exceptions. We sequenced the MUC1 gene fragment containing the SNP in all the cell lines and found that it is A/G in HeLa, MCF7 and Jurkat cells and G/G in normal T cells (Fig 5a). Of the 68 short isoforms isolated from the three tumor cell lines, some have A at 3506 while others have G (19% in HeLa, 41.7% in MCF7 and 47.2% in Jurkat), indicating that both MUC1 alleles are transcribed and alternatively spliced at their appropriate 3' splice sites. The exceptions were seven short isoforms that showed no 3' splice site use and thus complete intron 1 retention, and two other short isoforms, MUC1/J3 and MUC1/J7, that used other 3' splice sites. MUC1/J3 that has a G at position 3506 but was spliced at 3' splice site 3250 and MUC1/J7 with an A at position 3506 used the 3' splice site 3259. These results indicated that MUC1 3506 A/G SNP could produce different MUC1 isoforms with different signal peptide coding sequences by using at least four different 3' splice sites. Intron 1 retention could happen under both 3506 A and 3506 G.

We then wanted to see how the 3506 A/G SNP affects the MUC1 short isoforms expression at the protein level. Four MUC1 short isoforms, including two A allele associated isoforms MUC1/Y, MUC1/S1 and two G allele associated isoforms MUC1/Y-LSP, MUC1/S2 (Fig 5b), were transfected into DM6 cell line. Western-blotting results showed that MUC1/Y-LSP protein is slightly larger than MUC1/Y as predicted, but MUC1/S1 and MUC1/S2 exhibited a different protein pattern than predicted (Fig 5b), with MUC1/S1 having one band while MUC1/S2 had two bands (Fig 5c). Obviously, a different signal peptide coding sequence in the short MUC1 isoforms could generate different mature protein products.

To examine the occurrence of 3506 A/G SNP in de novo tumors in vivo, we employed triple transgenic mice that express human MUC1 as self [41] and carry conditional (loxP) engineered loci in the Kras and Pten genes [42]. The MUC1^{+/+}-Kras^{G12D/+}Pten^{loxP/loxP} (or briefly MUC1KrasPten) mice mirror the tissue MUC1 expression seen in humans and progress to MUC1 positive tumors of the ovaries and peritoneal cavity. Injection of Cre-encoding adenovirus under the ovarian bursa triggered development of endometrioid ovarian tumors in mouse 2347 (Fig 6a), while intraperitoneal (IP) AdCre injection triggered peritoneal carcinomatosis in one male (2949) and one female (2688) mouse (Fig 6a). Our SNP analysis of the tumor DNA using 3506 SNP A/G specific PCR showed that in 2 of 3 MUC1KrasPten mice that are normally 3506 A/A, the tumors carry both 3506 A and 3506 G, indicating that A to G mutation happened during the in vivo tumor formation.

MUC1/Y-LSP isoform inhibits tumor growth

We next sought to test whether the short isoforms MUC1/Y and MUC1/Y-LSP, when expressed in a MUC1 negative tumor cell line, might affect its growth *in vivo*. To address this, we transfected MUC1/Y, MUC1/Y-LSP and full length MUC1/TR with 22 tandem repeats (MUC1/22TR) into the mouse tumor cell line RMA and later selected the MUC1 positive, stable transfectants. When injected subcutaneously, RMA-MUC1/22TR and RMA-MUC1/Y cells grew progressively in two immunocompetent mouse strains, C57BL/6 and C57BL/6MUC1.Tg, while RMA-MUC1/Y-LSP cells did not (Fig. 7a). Though the growth rate of the RMA, RMA-MUC1/Y, RMA-MUC1/Y-LSP, and RMA-MUC1/22TR cells are similar in vitro (data not shown), the *in vivo* CFSE assay showed that after intraperitoneal injection into MUC1.Tg mice, most RMA-MUC1/Y-LSP cells were killed over a period of 44 hours while the parental RMA, RMA-Y and RMA-22TR survived, though in different ratios (Fig 7b), indicating that the expression of MUC1/Y-LSP targeted the tumor cells for destruction through a so far unknown mechanism. One mechanism we considered was the killing of tumor cells by NK cells. To test this possibility we injected the RMA-MUC1/Y-LSP tumor cells into immunocompromised NOD. CB17-*Prkdc*^{scid}/J mice (3/3) (Fig. 7a) that have NK cells but lack T and B cells. Our results show that all the mice injected developed tumors, ruling out NK-mediated lysis as the cause for the lack of *in vivo* tumor growth in immunocompetent mice.

DISCUSSION

Alternative splicing has an important role in the regulation of expression of genes associated with many diseases, including cancer. High-throughput approaches such as microarray chips [52], digital polony exon profiling technology [53] and deep sequencing [5–7] have been used for predicting or detecting splicing events. However, none of these methods yields splicing products for biochemical and functional studies. We adapted an RT-PCR based method for a comprehensive study of alternative splicing and resulting isoforms of a biologically important molecule, MUC1. This method should also be appropriate for isolating isoforms of many other genes of interest, with one caveat: the primers used are complementary to the start exon (E1) and the end exon (E7) and may miss rare isoforms with alternative start and end sites.

The MUC1 isoforms we describe here are generated by different alternative splicing events including single, double, triple and even quadruple exon-skipping, alternative 5' splice site selection, alternative 3' splice site selection, and full or partial intron retention. Different combinations of these patterns can be seen in different isoforms and in different tumor cells.

The VNTR region in exon 2 is a characteristic feature of the MUC1 gene. In all the alternatively spliced MUC1 isoforms this region was spliced out using the same 5' splice site and variable 3' splice sites (Fig. 2a–d). Exon 2 is considerably larger than any typical

eukaryotic exons, which normally range from 50 to 250 bp in length [54]. With an average of 80–100 TRs (60 bp per TR), MUC1 exon 2 can be over 6200bp, longer than many other genes in their entirety. The VNTR sequence contains the core splicing signals such as the 5' splice site, 3' splice site and the branch point site. In addition, it contains some cis-regulatory elements that have also been described, like the GGxxxxGGG motif which can act as an intron splice enhancer (ISE) [55]. The MUC1 tandem repeat sequence has this motif at the junction of the repeats. N tandem repeats would yield N-1 GGxxxxGGG motifs, the cumulative effect of which will be much stronger than a single enhancer. There are also 3 Nova splicing factor binding sites (YCAY, with Y being any pyrimidine) in each repeat, which when bound by Nova may block exon inclusion [56]. All these characteristics make the VNTR region of exon 2 look and function like an intron, possibly via the formation of tandem repeats with complementary hairpin loops pairing during transcription and the sequences flanking the loops being spliced together by specific spliceosomes [57–59].

In contrast to the splicing-out of the VNTR region, the cytoplasmic tail region remains mainly intact in almost all the MUC1 short isoforms, preserving important signaling functions [60] or interfering with the function of the full length molecule by competing for ligand binding.

Exon 2 is also significantly involved in isoform generation because of the A/G SNP (rs4072037) at position 3506. Though it is a silent SNP that does not change the amino acid encoded, it regulates the 3' splice site selection, in most cases, position 3498 being associated with A and 3471 site associated with G. Association of position 3250 with G and position 3259 with A were reported by us. The A/G SNP can produce different MUC1 short isoform protein products. Both A and G alleles were transcribed and alternatively spliced in the three tumor cell lines although the frequencies of isoforms generated from different alleles were quite different. Many more cell lines would need to be examined to determine if the observed difference was significant. There are reports that the A/A allele is associated with a higher risk of gastric cancer [61,62].

One surprising observation was that during de novo tumor formation in MUC1KrasPten mice, A to G mutation was observed. As 3506 SNP has been shown to regulate 3' splice site selection in intron 1, this change could alter the MUC1 isoform repertoire in cells. We plan to address this in future studies on archival samples of tumor and matching normal tissues, where the outcome of the disease is already known.

The extracellular domain of MUC1, in particular the VNTR region, undergoes profound changes in posttranslational modification during transformation from normal to tumor cells and this may result in changes in the interactions with the extracellular environment and the types of signals transmitted through the MUC1 cytoplasmic tail. Splicing out of that region, as we saw in all short isoforms, may serve to attenuate the consequences of those changes. It will be important to show convincingly that at least some of the short isoforms have a biological function. Until now, this has been difficult to show for many alternatively spliced genes in cancer [63]. We have begun this by cloning and expressing MUC1/Y-LSP and showing that it can inhibit tumor growth indirectly by stimulating a so far unidentified immune mechanism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health (CA56103 and CA73743 to O.J.F. and CA163462 to A.M.V), Department of Defense (OC093429 to AMV), National Science Foundation (MCB-0842725 to C.M.) and the Nathan Arenson Fund (O.F.).

References

1. Modrek B, Lee C. A genomic view of alternative splicing. *Nat Genet.* 2002; 30 (1):13–19. [PubMed: 11753382]
2. Johnson JM, Castle J, Garrett-Engle P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R, Shoemaker DD. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science.* 2003; 302 (5653):2141–2144. [PubMed: 14684825]
3. Xing Y, Lee C. Alternative splicing and RNA selection pressure--evolutionary consequences for eukaryotic genomes. *Nat Rev Genet.* 2006; 7 (7):499–509. [PubMed: 16770337]
4. Blencowe BJ. Alternative splicing: new insights from global analyses. *Cell.* 2006; 126 (1):37–47. [PubMed: 16839875]
5. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet.* 2008; 40 (12):1413–1415. [PubMed: 18978789]
6. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature.* 2008; 456 (7221):470–476. [PubMed: 18978772]
7. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. *Nature.* 2010; 463 (7280):457–463. [PubMed: 20110989]
8. Stamm S. Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. *Hum Mol Genet.* 2002; 11 (20):2409–2416. [PubMed: 12351576]
9. Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. *Genes Dev.* 2003; 17 (4):419–437. [PubMed: 12600935]
10. Garcia-Blanco MA, Baraniak AP, Lasda EL. Alternative splicing in disease and therapy. *Nat Biotechnol.* 2004; 22 (5):535–546. [PubMed: 15122293]
11. Pagani F, Baralle FE. Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet.* 2004; 5 (5):389–396. [PubMed: 15168696]
12. Venables JP. Aberrant and alternative splicing in cancer. *Cancer Res.* 2004; 64 (21):7647–7654. [PubMed: 15520162]
13. Pajares MJ, Ezponda T, Catena R, Calvo A, Pio R, Montuenga LM. Alternative splicing: an emerging topic in molecular and clinical oncology. *Lancet Oncol.* 2007; 8 (4):349–357. [PubMed: 17395108]
14. Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet.* 2007; 8 (10):749–761. [PubMed: 17726481]
15. Kim E, Goren A, Ast G. Insights into the connection between cancer and alternative splicing. *Trends Genet.* 2008; 24 (1):7–10. [PubMed: 18054115]
16. Krawczak M, Reiss J, Cooper DN. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum Genet.* 1992; 90 (1–2): 41–54. [PubMed: 1427786]
17. Lopez-Bigas N, Audit B, Ouzounis C, Parra G, Guigo R. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett.* 2005; 579 (9):1900–1903. [PubMed: 15792793]
18. Venables JP, Klinck R, Koh C, Gervais-Bird J, Bramard A, Inkel L, Durand M, Couture S, Froehlich U, Lapointe E, Lucier JF, Thibault P, Rancourt C, Tremblay K, Prinos P, Chabot B, Elela SA. Cancer-associated regulation of alternative splicing. *Nat Struct Mol Biol.* 2009; 16 (6): 670–676. [PubMed: 19448617]
19. Li HR, Wang-Rodriguez J, Nair TM, Yeakley JM, Kwon YS, Bibikova M, Zheng C, Zhou L, Zhang K, Downs T, Fu XD, Fan JB. Two-dimensional transcriptome profiling: identification of

- messenger RNA isoform signatures in prostate cancer from archived paraffin-embedded cancer specimens. *Cancer Res.* 2006; 66 (8):4079–4088. [PubMed: 16618727]
20. Zhang M, Gish W. Improved spliced alignment from an information theoretic approach. *Bioinformatics.* 2006; 22 (1):13–20. [PubMed: 16267086]
 21. Theodoropoulos G, Carraway KL. Molecular signaling in the regulation of mucins. *J Cell Biochem.* 2007; 102 (5):1103–1116. [PubMed: 17957706]
 22. Vlad AM, Kettel JC, Alajez NM, Carlos CA, Finn OJ. MUC1 immunobiology: from discovery to clinical applications. *Adv Immunol.* 2004; 82:249–293. [PubMed: 14975259]
 23. Ligtenberg MJ, Kruijsaar L, Buijs F, van Meijer M, Litvinov SV, Hilken J. Cell-associated episialin is a complex containing two proteins derived from a common precursor. *J Biol Chem.* 1992; 267 (9):6171–6177. [PubMed: 1556125]
 24. Fontenot JD, Finn OJ, Dales N, Andrews PC, Montelaro RC. Synthesis of large multideterminant peptide immunogens using a poly-proline beta-turn helix motif. *Pept Res.* 1993; 6 (6):330–336. [PubMed: 8292850]
 25. Fontenot JD, Tjandra N, Bu D, Ho C, Montelaro RC, Finn OJ. Biophysical characterization of one-, two-, and three-tandem repeats of human mucin (muc-1) protein core. *Cancer Res.* 1993; 53 (22):5386–5394. [PubMed: 8221676]
 26. Fukuda M. Roles of mucin-type O-glycans in cell adhesion. *Biochim Biophys Acta.* 2002; 1573 (3):394–405. [PubMed: 12417424]
 27. Hilken J, Ligtenberg MJ, Vos HL, Litvinov SV. Cell membrane-associated mucins and their adhesion-modulating property. *Trends Biochem Sci.* 1992; 17 (9):359–363. [PubMed: 1412714]
 28. Carlos CA, Dong HF, Howard OM, Oppenheim JJ, Hanisch FG, Finn OJ. Human tumor antigen MUC1 is chemotactic for immature dendritic cells and elicits maturation but does not promote Th1 type immunity. *J Immunol.* 2005; 175 (3):1628–1635. [PubMed: 16034102]
 29. Kufe DW. Mucins in cancer: function, prognosis and therapy. *Nat Rev Cancer.* 2009; 9 (12):874–885. [PubMed: 19935676]
 30. Barnd DL, Lan MS, Metzgar RS, Finn OJ. Specific, major histocompatibility complex-unrestricted recognition of tumor-associated mucins by human cytotoxic T cells. *Proc Natl Acad Sci U S A.* 1989; 86 (18):7159–7163. [PubMed: 2674949]
 31. Saeland E, van Vliet SJ, Backstrom M, van den Berg VC, Geijtenbeek TB, Meijer GA, van Kooyk Y. The C-type lectin MGL expressed by dendritic cells detects glycan changes on MUC1 in colon carcinoma. *Cancer Immunol Immunother.* 2007; 56 (8):1225–1236. [PubMed: 17195076]
 32. Ren J, Raina D, Chen W, Li G, Huang L, Kufe D. MUC1 oncoprotein functions in activation of fibroblast growth factor receptor signaling. *Mol Cancer Res.* 2006; 4 (11):873–883. [PubMed: 17114345]
 33. Mahanta S, Fessler SP, Park J, Bamdad C. A minimal fragment of MUC1 mediates growth of cancer cells. *PLoS ONE.* 2008; 3 (4):e2054. [PubMed: 18446242]
 34. Ren J, Li Y, Kufe D. Protein kinase C delta regulates function of the DF3/MUC1 carcinoma antigen in beta-catenin signaling. *J Biol Chem.* 2002; 277 (20):17616–17622. [PubMed: 11877440]
 35. Wei X, Xu H, Kufe D. Human MUC1 oncoprotein regulates p53-responsive gene transcription in the genotoxic stress response. *Cancer Cell.* 2005; 7 (2):167–178. [PubMed: 15710329]
 36. Ramasamy S, Duraisamy S, Barbashov S, Kawano T, Kharbanda S, Kufe D. The MUC1 and galectin-3 oncoproteins function in a microRNA-dependent regulatory loop. *Mol Cell.* 2007; 27 (6):992–1004. [PubMed: 17889671]
 37. Levitin F, Baruch A, Weiss M, Stiegman K, Hartmann ML, Yoeli-Lerner M, Ziv R, Zrihan-Licht S, Shina S, Gat A, Lifschitz B, Simha M, Stadler Y, Cholostoy A, Gil B, Greaves D, Keydar I, Zaretsky J, Smorodinsky N, Wreschner DH. A novel protein derived from the MUC1 gene by alternative splicing and frameshifting. *J Biol Chem.* 2005; 280 (11):10655–10663. [PubMed: 15623537]
 38. Baruch A, Hartmann M, Zrihan-Licht S, Greenstein S, Burstein M, Keydar I, Weiss M, Smorodinsky N, Wreschner DH. Preferential expression of novel MUC1 tumor antigen isoforms in human epithelial tumors and their tumor-potentiating function. *Int J Cancer.* 1997; 71 (5):741–749. [PubMed: 9180140]

39. Rubinstein DB, Karmely M, Ziv R, Benhar I, Leitner O, Baron S, Katz BZ, Wreschner DH. MUC1/X protein immunization enhances cDNA immunization in generating anti-MUC1 alpha/beta junction antibodies that target malignant cells. *Cancer Res.* 2006; 66 (23):11247–11253. [PubMed: 17145869]
40. Oosterkamp HM, Scheiner L, Stefanova MC, Lloyd KO, Finstad CL. Comparison of MUC-1 mucin expression in epithelial and non-epithelial cancer cell lines and demonstration of a new short variant form (MUC-1/Z). *Int J Cancer.* 1997; 72 (1):87–94. [PubMed: 9212228]
41. Rowse GJ, Tempero RM, VanLith ML, Hollingsworth MA, Gendler SJ. Tolerance and immunity to MUC1 in a human MUC1 transgenic murine model. *Cancer Res.* 1998; 58 (2):315–321. [PubMed: 9443411]
42. Dinulescu DM, Ince TA, Quade BJ, Shafer SA, Crowley D, Jacks T. Role of K-ras and Pten in the development of mouse models of endometriosis and endometrioid ovarian cancer. *Nat Med.* 2005; 11 (1):63–70. [PubMed: 15619626]
43. Budiu RA, Mantia-Smaldone G, Elishaev E, Chu T, Thaller J, McCabe K, Lenzner D, Edwards RP, Vlad AM. Soluble MUC1 and serum MUC1-specific antibodies are potential prognostic biomarkers for platinum-resistant ovarian cancer. *Cancer Immunol Immunother.* 2011; 60 (7):975–984. [PubMed: 21461842]
44. Lesche R, Groszer M, Gao J, Wang Y, Messing A, Sun H, Liu X, Wu H. Cre/loxP-mediated inactivation of the murine Pten tumor suppressor gene. *Genesis.* 2002; 32 (2):148–149. [PubMed: 11857804]
45. Budiu RA, Diaconu I, Chrissluis R, Dricu A, Edwards RP, Vlad AM. A conditional mouse model for human MUC1-positive endometriosis shows the presence of anti-MUC1 antibodies and Foxp3+ regulatory T cells. *Disease models & mechanisms.* 2009; 2 (11–12):593–603. [PubMed: 19841240]
46. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet.* 2002; 3 (4):285–298. [PubMed: 11967553]
47. Correa I, Plunkett T, Vlad A, Mungul A, Candelora-Kettel J, Burchell JM, Taylor-Papadimitriou J, Finn OJ. Form and pattern of MUC1 expression on T cells activated in vivo or in vitro suggests a function in T-cell migration. *Immunology.* 2003; 108 (1):32–41. [PubMed: 12519300]
48. Agrawal B, Krantz MJ, Parker J, Longenecker BM. Expression of MUC1 mucin on activated human T cells: implications for a role of MUC1 in normal immune regulation. *Cancer Res.* 1998; 58 (18):4079–4081. [PubMed: 9751614]
49. Fattorossi A, Battaglia A, Malinconico P, Stoler A, Andreocci L, Parente D, Coscarella A, Maggiano N, Perillo A, Pierelli L, Scambia G. Constitutive and inducible expression of the epithelial antigen MUC1 (CD227) in human T cells. *Exp Cell Res.* 2002; 280 (1):107–118. [PubMed: 12372344]
50. Ligtenberg MJ, Gennissen AM, Vos HL, Hilkens J. A single nucleotide polymorphism in an exon dictates allele dependent differential splicing of episialin mRNA. *Nucleic Acids Res.* 1991; 19 (2):297–301. [PubMed: 2014168]
51. Ng W, Loh AX, Teixeira AS, Pereira SP, Swallow DM. Genetic regulation of MUC1 alternative splicing in human tissues. *Br J Cancer.* 2008; 99 (6):978–985. [PubMed: 19238635]
52. Roberts GC, Smith CW. Alternative splicing: combinatorial output from the genome. *Curr Opin Chem Biol.* 2002; 6 (3):375–383. [PubMed: 12023119]
53. Zhu J, Shendure J, Mitra RD, Church GM. Single molecule profiling of alternative pre-mRNA splicing. *Science.* 2003; 301 (5634):836–838. [PubMed: 12907803]
54. Zhang XH, Chasin LA. Computational definition of sequence motifs governing constitutive exon splicing. *Genes Dev.* 2004; 18 (11):1241–1250. [PubMed: 15145827]
55. McCarthy EM, Phillips JA 3rd. Characterization of an intron splice enhancer that regulates alternative splicing of human GH pre-mRNA. *Hum Mol Genet.* 1998; 7 (9):1491–1496. [PubMed: 9700205]
56. Ule J, Stefani G, Mele A, Ruggiu M, Wang X, Taneri B, Gaasterland T, Blencowe BJ, Darnell RB. An RNA map predicting Nova-dependent splicing regulation. *Nature.* 2006; 444 (7119):580–586. [PubMed: 17065982]

57. Lian Y, Garner HR. Evidence for the regulation of alternative splicing via complementary DNA sequence repeats. *Bioinformatics*. 2005; 21 (8):1358–1364. [PubMed: 15673565]
58. Eperon LP, Estibeiro JP, Eperon IC. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. *Nature*. 1986; 324 (6094):280–282. [PubMed: 2946960]
59. Thompson-Jager S, Domdey H. Yeast pre-mRNA splicing requires a minimum distance between the 5' splice site and the internal branch acceptor site. *Mol Cell Biol*. 1987; 7 (11):4010–4016. [PubMed: 3323885]
60. Carson DD. The cytoplasmic tail of MUC1: a very busy place. *Sci Signal*. 2008; 1(27):pe35. [PubMed: 18612140]
61. Jia Y, Persson C, Hou L, Zheng Z, Yeager M, Lissowska J, Chanock SJ, Chow WH, Ye W. A comprehensive analysis of common genetic variation in MUC1, MUC5AC, MUC6 genes and risk of stomach cancer. *Cancer Causes Control*. 2010; 21 (2):313–321. [PubMed: 19924550]
62. Xu Q, Yuan Y, Sun LP, Gong YH, Xu Y, Yu XW, Dong NN, Lin GD, Smith PN, Li RW. Risk of gastric cancer is associated with the MUC1 568 A/G polymorphism. *Int J Oncol*. 2009; 35 (6): 1313–1320. [PubMed: 19885554]
63. David CJ, Manley JL. Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev*. 2010; 24 (21):2343–2364. [PubMed: 21041405]

Fig 1. MUC1 expression in tumor cell lines and activated T cells

Expression of MUC1/TR protein containing the VNTR region was analyzed by flow cytometry using cell surface (MCF7, HeLa, Jurkat, and activated T cells) and intra-cellular (Jurkat-intra) staining with an anti-VNTR antibody as described in Materials and Methods. Isotype control antibody staining is shown in the upper left corner in each figure. Percentages indicate MUC1+ cells.

Fig 2. Schematic representation of MUC1 isoforms isolated from tumor cell lines and activated T cells

MUC1 short isoforms were isolated from (a) HeLa, (b) MCF7, (c) Jurkat, and (d) activated T cells. MUC1 gene (M61170) is shown atop with seven exons (different color of blocks) and introns (black lines). The VNTRs in exon 2 are shown as vertical red lines. MUC1 short isoforms are ordered according to the length of the cDNA with the longest presented first. Names of the isoforms are shown in the left column. The cDNA column gives the total isoform length and ORF (open reading frame) column the ORF length. Numbers in red mark isoforms that have pre-mature termination codons (PTCs). Asterisks (*) mean there were more than one colony carrying the isoform.

Fig 3. Common and cell-specific MUC1 short isoforms

The diagram shows the common as well as cell-specific MUC1 short isoforms isolated from tumor cell lines and the activated T cells.

Fig 4. MUC1 VNTR region is recognized as an intron

Plasmids containing MUC1 cDNA with 42 TRs (pcDNA3-MUC1/42TR) were transfected into HeLa cells and total RNAs were isolated from HeLa/42TR cells and HeLa cells (negative control) after culturing for 24 hrs. After RT-PCR, there were multiple bands as well as smears due to the alternative splicing of MUC1/42TR mRNA and incomplete amplification of the TR region caused by the high GC content. Bands A, B, and C were recovered and cloned into pCR 4-TOPO vector for sequence analysis. MUC1/Y and MUC1/G, two MUC1 short isoforms devoid of the VNTR region, were products generated from MUC1/42TR through alternative splicing.

Fig 5. 3506 A/G SNP in MUC1 exon 2

(a) PCR products from genomic DNA containing the 3506 A/G SNP of HeLa, MCF7, Jurkat and activated primary T cells were sequenced. (b) Schematic diagram of MUC1/TR protein and 3506 A/G SNP associated short isoform proteins. It was predicted that mature MUC1 isoform proteins with G SNPs (MUC1/Y-LSP, MUC1, S2) should have 9 additional a.a. on their N-terminals than those with A SNPs (MUC1/Y, MUC1/S1). MUC1/TR contains 2 subunits (α and β) noncovalently bound together. In the short isoforms, due to the disruption or deletion of the SEA domain, the proteins have only one unit. CT2 antibody binding site is shown. (c) Western blotting of DM6 cells stably transfected with different MUC1 short isoforms.

Fig 6. 3506 A to G mutation in spontaneous tumors generated in $MUC1^{+/-}Kras^{G12D/+}Pten^{loxP/lox}$ mice

(a) H&E staining of tumors in the left ovary (mouse 2347, primary ovarian tumor) and on peritoneal walls (mice 2688 and 2949, primary peritoneal carcinomatosis). (b) PCR amplification of SNP A and G specific fragment from genomic tumor DNAs. DNAs from T cells (3506 G/G), mouse tail DNA (3506 A/A), and HeLa cells (3506 A/G) were used as control templates.

Fig 7. MUC1/Y-LSP has tumor growth inhibition activity *in vivo*

(a) Tumor formation of RMA, RMA-Y, RMA-Y-LSP, and RMA-22TR in different mouse models. (b) CFSE assays in MUC1.Tg mice (3 each) using RMA vs RMA-Y-LSP, RMA vs RMA-Y and RMA vs RMA-22TR, respectively. The histograms on the left of the dashed line are CFSE-labeled cells before injection. Those on the right of the dashed line are cells 44 hours after injection into 3 mice, respectively.