



Published in final edited form as:

J Immunol. 2004 January 1; 172(1): 651–660.

A Novel Mechanism of Alternative Promoter and Splicing Regulates the Epitope Generation of Tumor Antigen CML66-L¹

Yan Yan^{*}, Leuyen Phan^{*}, Fan Yang^{*}, Moshe Talpaz[§], Yu Yang^{*}, Zeyu Xiong^{*}, Bernard Ng^{*}, Nikolai A. Timchenko[‡], Catherine J. Wu[¶], Jerome Ritz[¶], Hong Wang^{*}, and Xiao-Feng Yang^{*,†,§,2}

^{*}Department of Medicine, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030

[†]Department of Immunology, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030

[‡]Department of Pathology, Baylor College of Medicine, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030

[§]Department of Bioimmunotherapy, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030

[¶]Center for Hematologic Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115

Abstract

This report describes the difference in the epitope generation of two isoforms of self-tumor Ag CML66 and the regulation mechanism. We identified a new CML66 short isoform, termed CML66-S. The previously identified long CML66 is referred to as CML66-L. CML66-S shares the C terminus with CML66-L but has its unique N terminus. CML66-S is predominantly expressed in testis, but is also expressed in very low levels in tumor cells, whereas CML66-L is expressed in tumor cells and testis. Differential expression of CML66-L and CML66-S in tumor cells resulted from regulation at transcription, although alternative splicing also participated in the generation of the isoforms. In addition, Ab titers to a CML66-L peptide were significantly higher than that to CML66-S peptide in the sera from patients with tumors. Finally, the Abs to full-length CML66-L in the sera from patients with tumors were correlated with the Abs in the sera from these patients to CML66-L-38, which is a fusion protein with a CML66-L-specific N terminus. This suggests that the CML66-L isoform is mainly responsible for the epitope generation. Our studies have identified the alternative promoter in combination with alternative splicing as a novel mechanism for regulation of the epitope generation of a self-tumor Ag.

Cancer-testis Ags (CT Ags)³ are unmutated self-Ags and encoded by an expanding group of genes including *MAGE* (1) and *NY-ESO-1* (2), whose mRNAs are predominantly expressed

¹This work was partially supported by a career development fund at Baylor College of Medicine; National Institutes of Health Grants K08 AI054514, P30 DK56238 (Gulf Coast Digestive Disease Center), P50 CA58204 (Specialized Program of Research Excellence in prostate cancer), and P20 CA103698 (Baylor Cancer Center); and funds from the Kostas Fund for Dermatomyositis, the Caroline Weiss Law Foundation for Molecular Medicine, the Myositis Association of America, American Heart Association Texas affiliate, Leukemia and Lymphoma Society, and the Myeloproliferative Disorders Foundation. X.-F.Y. is a Chao's Family Scholar of Medicine.

Copyright © 2004 by The American Association of Immunologists, Inc.

²Address correspondence and reprint requests to Dr. Xiao-Feng Yang, Section of Immunology, Allergy, and Rheumatology, Department of Medicine, Biology of Inflammation Center, Baylor College of Medicine, One Baylor Plaza, BCM 285, Suite 672E, Houston, TX 77030-3411. xyang1@bcm.tmc.edu.

³Abbreviations used in this paper: CT Ag, cancer-testis Ag; CML, chronic myelogenous leukemia; UTR, untranslated region; TNT, transcription and translation; ORF, open reading frame.

in 10–40% of a wide range of different cancers and testis (3) but generally not in most other normal tissues (4). Because testis is an immune-privileged site that does not express MHC molecules (5), CT Ags can be practically regarded as tumor specific and highly desirable as targets for immunotherapy (4).

Recently, using sera from patients with chronic myelogenous leukemia (CML) to conduct serological analysis of tumor Ags by recombinant cDNA expression cloning (6–8), we identified a novel, broadly immunogenic tumor Ag, CML66 (9, 10). CML66 encodes a 583-aa protein with a molecular mass of 66 kDa and shows no significant homology to other known genes. CML66 is expressed in a variety of tumor cells. CML66 expression in normal tissues is mostly restricted to testis, suggesting that CML66 may be similar to the Ags in CT Ag group. CML66-specific IgG Ab is found in sera from 18 to 38% of patients with lung cancer, melanoma, and prostate cancer. More importantly, the development of high-titer IgG Ab specific for CML66 correlates with immune-induced remission of CML in a patient who received infusion of normal donor lymphocytes for treatment of relapse. Our findings suggest that CML66 is broadly immunogenic in a wide variety of malignancies and may be a target for Ag-specific immunotherapy (10).

The mechanisms for selective expression of CT Ags in testis and cancers are generally ascribed to demethylation events in cancer (11). However, the mechanism for the immunogenicity of the CT Ags up-regulated in tumor cells remains largely unknown. Elucidation of the mechanism for immunogenicity of CT Ags is very important for future development of Ag-specific immunotherapy. To this end, we attempted to use novel Ag CML66 as a model to explore the mechanism of the immunogenicity of CT Ags. In this report, we identified a CML66 short isoform with 554 aa, termed CML66-S. We refer to the previously identified CML66 of 583 aa as CML66-L. CML66-S shares the C-terminal 544 aa with CML66-L but has its unique 5'-untranslated region (UTR) and its own N terminus. Our results suggest that CML66-L predominantly expressed in tumor cells, rather than newly identified CML66-S that is mainly expressed in testis, is the dominant immunogenic isoform. Up-regulation of CML66-L, but not CML66-S, in tumor cells is the mechanism for its immunogenicity in patients with tumors. Our studies have identified the alternative promoter in combination with alternative splicing as a novel mechanism for regulation of the epitope generation of a self-tumor Ag.

Materials and Methods

Serum samples

Serum samples were obtained from normal donors and patients with CML receiving IFN- α therapy enrolled in an Institutional Review Board-approved clinical trial (M. D. Anderson Cancer Center and Baylor College of Medicine). Serum samples from patients with hormone-refractory advanced prostate cancer were generously provided by Dr. P. Kantoff (Harvard) (10).

Human testis cDNA library screening and 5'-RACE

A human testis cDNA library (1×10^6 phages) (BD Clontech, Palo Alto, CA) was screened with a 0.8-kb ^{32}P -labeled CML66 common-region probe (10). Positive clones were converted into plasmid pTriplEx and sequenced in both strands (Lone Star, Houston, TX). 5'-RACE was conducted by PCR with the human testis Marathon-Ready cDNAs (BD Clontech) as templates. Two antisense primers specific for the 5'-UTR of CML66-S cDNA, SU1 (5'-ATGCTCCAACCCTGGAGATAAAGCAT-3') and the nested primer SU3 (5'-ACGATGGACATGGGACCCTTCAATAC-3'), were synthesized at Sigma-Genosys (The Woodlands, TX). The primer SU1 was paired with the adaptor primer AP1 provided in the

kit, and the nested primer SU3 was paired with the adaptor primer AP2, also from the kit. PCR was conducted in a two-temperature touch-down format, with three sets of cycling parameters as follows: 94°C for 30 s and 72°C for 1 min for 5 cycles; 94°C for 30 s and 70°C for 1 min for 5 cycles; and finally, 94°C for 30 s and 68°C for 1 min for 25 cycles. Specific PCR products were subcloned into the pCR-TOPO vector kit (Invitrogen, Carlsbad, CA) and sequenced at Lone Star.

In vitro transcription and translation (TNT)

Bluescript plasmid containing cDNA encoding CML66-L, plasmid pTriplEx containing cDNA encoding CML66-S, and plasmid pTriplEx without cDNA insert (as a negative control) were used (1 µg each) in the in vitro TNT, which were then performed using a TNT T7/T3-coupled system according to the manufacturer's protocol (Promega, Madison, WI) (12).

RT-PCR and PCR cloning

RNA was prepared from tumor cell lines by using RNAzole (Tel-Test, Friendswood, TX). RNA from human testis was purchased from BD Clontech. RT-PCR and PCR cloning were performed as described (10). A sense primer (25S) specific for the 5' upstream CML66-S isoform (5'-CGGAGAATTCGATGCTTTATCTCCAGGGTTGGAGC-3') (the underline indicates the sequence for cloning purposes) and a second antisense primer (25H) (5'-CCCAGGTAGAAGATGAGAAATGGATA-3') specific for the CML66 common region were used for PCR. The PCR with 35 cycles was used for amplification of CML66-S fragment due to its very low abundance. The amplified PCR products (290 bp) were confirmed by using Southern blot hybridization with a ³²P-labeled CML66 internal probe (25EF) that was not overlapped with either primer 25S or primer 25H. The hybridization signals in each lane were analyzed with a Kodak documentation system (Eastman Kodak, Rochester, NY) and normalized by *GAPDH* (housekeeping gene control) density signals amplified in the same cDNA preparation. A sense primer (25K) specific for the CML66-L-specific 5' end (5'-CGGAGAATTCGGCACGAGTCCCAGTCTCTGTGCGA-3') was paired with the antisense primer 25H specific for the CML66 common region to amplify the CML66-L transcript as confirmed previously (10). The PCRs with 30 cycles were used in the amplification of CML66-L fragment and that of *GAPDH* due to their relatively high abundance.

Matched tumor/normal expression array analysis

A matched tumor/normal expression array was purchased from BD Clontech, which includes SMART-amplified cDNA from 64 tumor and corresponding normal tissues from individual patients. The array was hybridized, respectively, with three ³²P-labeled probes, including CML66-L-specific cDNA probe, CML66-S-specific cDNA probe, and human ubiquitin cDNA probe (housekeeping gene control). The signals in the array were completely removed between two hybridizations. Probes were prepared using the Strip-EZ DNA labeling kit (Ambion, Austin, TX) and [α -³²P]dATP (ICN, Irvine, CA). The hybridization signals were detected with Kodak Bio-Max films (Eastman Kodak) and were subjected to densitometric analysis using the Kodak System EDAS-290. After normalization with signal density of the ubiquitin in each sample, CML66-L signal densities from normal samples were used to calculate the mean ratios (X) and 3 SD, which were used as the reference for signal variations between normal samples (13). Variations presumably resulted from the manufacturer procedures of the array, gene expression differences between individuals, and gene expression differences between tissues (14). The ratios of CML66-L signal in tumor sample/CML66-L signal in normal sample were calculated as CML66-L expression level for each matched tumor/normal pair sample. The expression ratios of

CML66-L in tumor over CML66-L expression in matched normal control larger than the upper limit of the variations among the normal samples (mean + 3 SD) was considered as the positive in the expression of CML66-L in the tumor sample.

Mapping of the exon encoding for CML66-S-specific 5'-end cDNA

Long PCR was used to map the relative position of the exon in human genomic DNA encoding for CML66-S-specific cDNA. Two sense primers, primer 25K (5'-GGCACGAGTCCCAGTCTCTGTGCGA-3') specific for the 5' end of CML66-L and primer 25SA (5'-GTATTGAAGGGTCCCAT GTCCATCG-3') specific for the 5' end of CML66-S, were synthesized. In addition, two antisense primers, primer SU1 (5'-ATGCTCCAACCCTG GAGATAAAGCAT-3') specific for the 5' end of CML66-S and primer 25M1 (5'-CCTAAGGCAGTGTCCAGCATTACTG-3') specific for the exons encoding CML66 common region cDNA, were synthesized. A long PCR was performed by using a λ genomic clone of CML66 (10) as the DNA template with a genomic DNA PCR kit (BD Clontech).

Construction of luciferase report vector, transfection, and promoter function assay

The 1.782-kb DNA sequence, termed the CML66-L promoter (the promoter 66L), was generated by a high-fidelity PCR (BD Clontech) using sense primer Pro5U (5'-GCCGGGCCGGTACCGGCTAATGCAGTGTGGGTTT-3'), antisense primer Pro3D (5'-GCCGGGCCGGTACCGTTG CATGCTGGGCTATGTA-3'), and human genomic DNAs (BD Clontech) as the template. The underlined sequences were designed for subcloning purposes. The pGL-3 luciferase reporter enhancer vector designed for the examination of promoter activity (Promega) was used. The promoter CML66-L was subcloned into the *KpnI* sites of pGL-3 luciferase reporter enhancer vector in both sense and antisense orientations and confirmed by DNA sequencing. The 2.6-kb DNA sequence (*KpnI* fragment) upstream of the 5' end of CML66-S, termed promoter CML66-S, was subcloned from a CML66 genomic DNA clone (10) into the *KpnI* site of pGL-3 luciferase enhancer vector in both sense and antisense orientations. 66L promoter-pGL-3 and 66S promoter-pGL-3 vectors as well as three control vectors, including 1) pGL-3 enhancer without any inserts, 2) pGL-3 Basic vector, and 3) pGL-3 Control vector (not shown), were transfected into HeLa cells, HCCT hepatic carcinoma cells (not shown), and NTERA-2 cl.D1 human testicular embryonal carcinoma cells (American Type Culture Collection, Manassas, VA), respectively. The transfections were performed with the reagent Lipofectamine 2000 (Invitrogen) according to the manufacturer's instruction. Luciferase activity in transfected cells was measured with the Dual Luciferase system (Promega) on a TD20/20 luminometer (Turner Designs, Sunnyvale, CA). Cotransfected *Renilla* luciferase control vector (Promega) provides the expression of *Renilla* luciferase as an internal control. The ratio of the firefly luciferase activity vs the *Renilla* luciferase activity is used as relative luciferase activity.

Peptide synthesis

A CML66-L-specific peptide (NH₂-PCYQLELDAAVAEVK-COOH) and a CML66-S-specific peptide (NH₂-LYLQWSMPAVAEVK-COOH) were synthesized at Sigma-Genosys.

Generation of GST fusion protein

A cDNA fragment encoding the 38 aa in CML66-L-specific N terminus was generated by PCR using the Advantage HF (high-fidelity) PCR kit (BD Clontech), sense primer 385 (5'-CCGGGAATTCAGAGGTGGCGGCTAATTGCTCCCTA-3'), and antisense primer 383 (5'-CCGGGAATTCCCGGGTGCGTCAAGCTCCAGCTGGTAACA-3'). Similarly, a cDNA fragment encoding full-length CML66-S protein was generated by high-fidelity PCR

using the Advantage HF PCR kit, sense primer 66S-5 specific for CML66-S-specific 5'-end sequence of CML66-S open reading frame (ORF) (5'-TGCTGGGAATTCACTTTATCTCCAGGGTTGGAGCATG-3'), and antisense primer 66S-3 specific for the 3'-end sequence of CML66 common region (5'-TGCTGGGAATTCCCGGGTTAATTCTCTGTATTACTTTTAT-3'). The underlined sequences were designed for subcloning into the *EcoRI* of GST vector pGEX-3X (Amersham Biosciences, Piscataway, NJ). CML66-L-38 fragment and full-length CML66-S ORF cDNA were fused in frame, respectively, to GST and were confirmed by DNA sequencing. The GST, GST fusion proteins were purified with Glutathione Sepharose beads according to the manufacturer's protocols (Amersham Biosciences) (10).

Mass spectrometric determination of the molecular mass of GST-CML66-S

After being purified using Glutathione Sepharose 4B affinity chromatography (Amersham Biosciences), purified GST-CML66-S was further subjected to size exclusion chromatography on a gel filtration column (Bio-Silect SEC-125 HPLC; Bio-Rad, Hercules, CA). After verification with Western blot using anti-GST Ab, GST-CML66-S containing HPLC fractions (200 μ l each) were pooled and freeze-dried on a Virtis (Gardiner, NY) lyophilizer. The molecular mass of GST-CML66-S was determined by mass spectrometry on the matrix-assisted laser desorption ionization-mass spectrometry/time of flight BiFlex III instrument (Bruker Daltonics, Billerica, MA) at Dr. T. Marriott's laboratory (Mass Spectrometry Center, Department of Chemistry, Rice University, Houston, TX). Briefly, mass spectrometry was performed by mixing 5 mM sample in water (1:2) with sinapinic acid matrix made up in 50:50 acetonitrile:water with 0.1% trifluoroacetic acid added. A volume of 1.5 μ l of the sample plus matrix solution was spotted onto the target plate and allowed to evaporate to dryness. The matrix-assisted laser desorption ionization-time of flight-mass spectrometry was obtained on the BiFlex III with a pulsed nitrogen laser (337 nm), in linear positive-ion mode, using a 19-kV acceleration voltage. BSA and GST were used as standards.

Detection of CML66-specific Ab in patient sera by ELISA

ELISA plates (VWR Scientific, Atlanta, GA) were coated with 50 μ l of purified proteins (GST, GST-CML66-L, or GST-CML66-L-38) or CML66-L peptide and CML66-S peptide at 5 μ g/ml in coating buffer (PBS plus 0.05% sodium azide) overnight at 4°C. ELISAs were performed as previously described (10). OD₄₀₅ was immediately read on the Spectra Max Plus plate reader (Molecular Devices, Sunnyvale, CA). In addition to using GST-CML66-L full length and GST-CML66-L-38 as coating Ags, each sera sample was also tested against GST alone as a control. The ratio of the OD₄₀₅ GST-CML66-L full length/OD₄₀₅ GST or OD₄₀₅ GST-CML66-L-38/OD₄₀₅ GST was used to determine the specific reactivity above background, as the results previously reported (9, 10, 15). An OD₄₀₅ ratio higher than 3 SD above the mean ratio of OD₄₀₅ (test Ag)/OD₄₀₅ GST of normal donor sera was interpreted as positive. The specificity of positive samples also was confirmed by demonstrating inhibition of reactivity by preincubation of patient sera with purified recombinant GST-CML66-L or GST-CML66-L-38, respectively. Similarly, the OD₄₀₅ ratio of the OD₄₀₅ CML66-L peptide/OD₄₀₅ CML66-S peptide was used to determine the specific reactivity against CML66-L. An OD₄₀₅ ratio of OD₄₀₅ CML66-L peptide/OD₄₀₅ CML66-S peptide higher than 3 SD above the mean ratio of OD₄₀₅ was interpreted as positive.

Results

Predominant expression of novel CML66 short form, CML66-S, in normal testis

In our previous studies (10), we observed a double-band structure of CML66 in human testis RNA by Northern blot, one with a size of 2.4 kb and one with a size of 2.1 kb. The observed

single band of β -actin, as an RNA quality control, in human testis in the same Northern blot suggested that there are two CML66 transcripts with different sizes, a long CML66 transcript and a short CML66 transcript, in a good quality of RNAs from human testis (10). In contrast, only the longer transcript of CML66 with a size of 2.4 kb was found in tumor cell lines, such as erythroleukemia cell line K562, colorectal adenocarcinoma cell SW480, and six other tumor cells shown previously (10), suggesting that these two transcripts of CML66 have different expression patterns, and the shorter CML66 transcript is only expressed in human testis (10). To elucidate the structure of the shorter CML66 transcript, we screened a human testis cDNA library and performed 5'-RACE to clone the complete sequence of the 5'-UTR of the shorter CML66. Through screening of human testis cDNA library, in addition to CML66 identified in the previous studies (10), we identified a novel short form of CML66 with the replacement of 349 bp at the 5' end of the previously published CML66 sequence by a novel sequence segment of 123 bp. The results of 5'-RACE with different testis cDNA resources further confirmed the existence in testis of the short-form transcript of CML66 identified in the screening of testis cDNA library. This form is termed CML66 short form (CML66-S) (Fig. 1A). Thus, we refer to the previously identified CML66 as CML66 long form (CML66-L) (10). The identification of CML66-S in testis by molecular cloning further verified the previous observation of two CML66 transcripts in testis Northern blot analysis.

Although the sequence around the start codon in CML66S did confirm an adequate Kozak consensus sequence (A/GNNATGC/T) rather than the dominant Kozak consensus (A/GNNATGG) (16), in correlation with this analysis, the results from the in vitro TNT assay, CML66-S cDNA can be transcribed and translated into a protein with a calculated molecular mass of 63.5 kDa in an SDS-PAGE, slightly smaller in mass than 66 kDa of CML66-L (Fig. 1B), suggesting that CML66-S transcript is capable of being translated into a protein. By 5'-RACE of CML66-S, we identified a complete 5'-UTR of CML66-S. The cDNA encoding full-length CML66-S protein was completely sequenced in both sense and antisense strands. Thus, CML66-S is 2098 bp long, including its specific 98-bp 5'-UTR and its specific N-terminal 10 aa, but CML66-S shares the C-terminal 544-aa common region and the 335-bp 3'-UTR with CML66-L (GenBank accession no. AF521133). To further verify the molecular mass of CML66-S calculated based on the DNA sequence analysis, by using a CML66-S-specific sense primer specific for the beginning of CML66-S ORF, an antisense primer specific for the cDNA sequence encoding CML66 common C terminus and high-fidelity PCR, a full-length GST-CML66-S fusion protein was constructed. After purification of GST-CML66-S fusion protein with Glutathione Sepharose beads affinity chromatography followed by a size exclusion chromatography on HPLC, we measured the molecular mass of purified GST-CML66-S fusion protein by mass spectrometry. As shown in Fig. 1C, GST-CML66-S has a molecular mass of 90,178 Da, suggesting that CML66-S has a molecular mass of 63,297 Da after subtraction of 26,881 Da of GST mass, which verified the calculated molecular mass of CML66-S and the apparent molecular mass results from the SDS-PAGE analysis of in vitro-translated CML66-S. By database mining, we identified a mouse homolog of CML66-L with a transcript of 2601 bp that encodes 582 aa residues (GenBank accession no. AF521132). Mouse CML66-L is highly homologous to human CML66-L, with a 90.4% amino acid identity, suggesting that CML66-L is highly conserved through evolution. However, a database search did not reveal mouse CML66-S.

The long PCR results suggested that the exon encoding CML66-S-specific 5'-cDNA (the exon 2) was located 4 kb downstream from the exon encoding the CML66-L-specific 5' end (the exon 1) (not shown). There was a 7- to 8-kb intron region between the CML66-S-specific exon (the exon 2) and the downstream CML66 common exon (the exon 3) (not shown). The identities of the long PCR products were confirmed by DNA sequencing (GenBank accession no. AF521134). Corresponding with our finding, the most recent

GenBank search showed that the *CML66* gene has a total of 11 exons and spans a 91.6-kb-long region (GenBank accession no. AC021237) (Fig. 1D). CML66-L is encoded by all the exons except exon 2 (exons 1 and 3–11). In contrast, CML66-S is encoded by all the exons except exon 1 (exons 2–11). All of the intron-exon junctions confirmed the GT-AG rule (not shown) (17). This cross-examination with different sources of human genomic DNAs confirmed the results of CML66-S exon position between the exon 1 and CML66 common exon 3 and also confirmed no DNA rearrangement or deletion in the CML66 region. These results suggested that, in addition to potential transcriptional mechanism explored in the latter section of this study, generation of CML66-S and CML66-L is partially modulated by alternative splicing, either the exon 2 splicing to the isoform-shared exons 3–11 to generate CML66-S transcript, or the exon 1 splicing to the isoform-shared exons 3–11 to generate the CML66-L transcript, as shown in Fig. 1E. The alternative promoter and alternative splicing of the first exon was also found to be the mechanism of gene regulation used in other genes in the generation of isoforms including myosin L chain (18, 19).

Predominant expression of CML66-L but not CML66-S in tumor cells

Previously, Northern blot analysis with CML66 common-region probe showed that the expression of CML66-S was limited to human testis but not tumor cells or any other normal tissues (10). By using semiquantitative RT-PCR, as shown in Fig. 2A, the expression of CML66-S was found mostly in testis but only detected in very low levels, even after 35 cycles, in tumor cell lines, including lung cancer line A549, breast cancer line MCF7, lymphoid leukemia line Jurkat, and erythroleukemia line K562, which were verified by Southern blot hybridization with a CML66-specific cDNA probe with no overlap with PCR primers (not shown). In contrast, CML66-L transcripts and the control housekeeping gene *GAPDH* were detected in all four tumor cell lines and testis with 30 cycles, which corresponded to the previous report (10). This result suggested that CML66-S was specifically expressed in testis but expressed at barely detectable levels in tumor cells.

Previous results on CML66-L expression were derived from Northern blot using RNAs purified from tumor cell lines (10). To examine CML66-L expression in tumor cells from patients, hybridization analyses of the matched tumor/normal expression array were performed with ³²P-labeled CML66-L-specific probe, CML66-S-specific probe, and ubiquitin probe, respectively. These analyses showed that CML66-L was expressed in a variety of tumors, including kidney cancer, breast cancer, uterine cancer, ovarian cancer, colon cancer, lung cancer, gastric cancer, and intestinal cancer, with positive rates from 20 to 66%, shown in Fig. 2B. In contrast, CML66-S had no detectable expression in these tumor samples and their normal control tissues (not shown).

Predominant expression of CML66-L in tumor cells resulted from tumor-specific transcription of CML66-L promoter

By performing primer extension experiments with an antisense primer (not shown) annealing to the CML66-L-specific 5' end, we found that there was a dominant fragment with a size around 280 bp from the start codon of CML66-L (not shown), which suggests that the CML66-L promoter may locate upstream from position –280. These results corresponded well with the following promoter mapping analysis. In the human *CML66* locus genomic sequence (GenBank accession no. AC021237) upstream of CML66-L 5'-UTR, with a promoter prediction software TSSW human poly II recognition using the TRANSFAC database (Baylor website; <http://mbr.bcm.tmc.edu/databases.html>), computer analyses indicated that in the sequence upstream of CML66-L start codon (+1), there are a TATA box structure (–289) and several transcription factor binding sites (20), including c-ETS-1 (–320), p53 (–387), NF-κB (–446), AP2 (–456), and SP1 (–501), suggesting that the DNA portion may be a promoter responsible for CML66-L transcription, termed CML66-L

promoter. In addition, the detailed 5'-RACE of CML66-S showed that three CML66-S 5'-RACE clones had a 98-bp sequence, whereas the other three clones had the same sequence but 12 bp shorter. Staggered starts in the 5' end of CML66-S 5'-RACE clones suggest that the transcription of the CML66-S-specific transcripts can be driven by an alternative promoter upstream of CML66-S, followed by alternative splicing of the 3' end of CML66-S-specific transcripts into the 5' end of CML66 transcripts encoded by CML66 common exons 3–11. This conclusion was further supported by the finding that there are no sequences sharing between CML66-L 5'-UTR and the CML66-S 5'-UTR, suggesting that CML66-S-specific transcription was driven by a CML66-S-specific promoter rather than CML66-L promoter. Similar search for a potential promoter for CML66-S transcription in the sequence of the intron 1 (GenBank accession no. AF521134) showed there were no any TATA box structures in 2.6 kb, even in the whole intron 1 (4 kb), which was confirmed with another promoter software PROSCAN (National Institutes of Health website; <http://bimas.dcr.t.nih.gov/molbio/index.html>). However, it is noteworthy that several transcription factor-binding sites were found in the sequence of the CML66-S promoter, including NF- κ B (-341), TATA box binding protein (-321), Pit-1 (-319), PU.1 (-200), C-Myb (-52), and GATA-1 (-18). Although there is no TATA box-based promoter, the existence of several transcription factor binding sites in the CML66-S promoter implies that it may have TATA box-less promoter activity as demonstrated in other genes (21). We hypothesized that predominant expression of CML66-L, but not CML66-S, in tumor cells may partially result from the differential promoter activities of the two promoters. To test this hypothesis, we adopted a luciferase reporter assay to examine the promoter activities of CML66-L promoter and CML66-S promoter. As shown in Fig. 3, C and D, the representative results from the experiments repeated four times showed that the 1.8-kb CML66-L promoter in the sense orientation had strong promoter activities compared with the antisense-oriented CML66-L promoter control and the no-insert vector control in both HeLa cells and HCCT cells (not shown). The strength of sense-oriented CML66-L promoter was in a level equivalent to that of SV40 promoter in pGL-3 Control vector (the positive control, not shown) (22, 23). In contrast, CML66-S promoter in the sense orientation had much low promoter activities in comparison to that of CML66-L promoter in HeLa cells and HCCT cells. The ratio of CML66-L promoter activities vs CML66-S promoter activities was 38:1 in HeLa cells, and similar results were obtained in HCCT cells (not shown). These results corresponded well with the sequence analysis regarding the TATA box and transcription factor binding site structure of CML66-L promoter and CML66-S promoter. However, it is notable that the CML66-S promoter activities (mean \pm 2 SD, 2.28 ± 0.88 in HeLa cells) were statistically higher than that of CML66-S ($p < 0.05$) in the antisense orientation control (mean \pm 2 SD, 0.70 ± 0.29) and no insert vector control (mean \pm 2 SD, 0.74 ± 0.35) in these tumor cells. These results suggest that CML66-S promoter had very low promoter activities in both tumor cells. The CML66-S promoter may be functional in testis, as suggested by RT-PCR results in Fig. 2A. To test whether CML66-S promoter has a higher activity in the testis-derived cell line than that in HeLa cells and HCCT cells, we performed a luciferase reporter assay to examine the promoter activities of CML66-L promoter and CML66-S promoter in NTERA-2 cl.D1 human testicular embryonal carcinoma cells (American Type Culture Collection). We found that CML66-S promoter activities in NTERA-2 cl.D1 testicular carcinoma cells were seven times higher than that in HeLa cells and HCCT cells (Fig. 3D). The ratio of CML66-L promoter activities vs CML66-S promoter activities in NTERA-2 cl.D1 testicular carcinoma cells (Fig. 3D) was decreased to 6.7:1 from 38:1 in HeLa cells and HCCT cells (C). The results further confirmed the functional activities of CML66-S promoter in testis. The difference between CML66-L promoter activity and CML66-S promoter activity in NTERA-2 cl.D1 testicular carcinoma cells may result from their structural and functional differences and the malignancy of the testicular carcinoma cells. The transcription results of CML66-L promoter and CML66-S promoter in tumor cells

corresponded well with the expression data of both CML66-L and CML66-S transcripts detected with semiquantitative RT-PCR and cDNA array in tumor cells (Fig. 2). These results suggest that the predominant expression of CML66-L, but not CML66-S, in tumor cells mainly results from the difference of RNA transcriptions driven, respectively, by CML66-L promoter and CML66-S promoter in tumor cells.

Epitope generation of CML66 contributed mostly by CML66-L but not CML66-S

CML66-L may be the dominant Ag epitope generation form in comparison to CML66-S in patients with tumors. To directly test this issue, 15-mer CML66-L-specific peptide was synthesized with the N-terminal 9 aa unique to the N terminus of CML66-L and the 6 aa from the CML66 common region (Fig. 4A). In comparison, the 15-mer CML66-S-specific peptide that has the 9 aa from the N terminus of CML66-S and the 6 aa from CML66 common region in its C terminus, which are the same as those in CML66-L peptide, was also synthesized. The OD₄₀₅ ratio of the OD₄₀₅ CML66-L peptide/OD₄₀₅ CML66-S peptide was used to determine the specific reactivity against CML66-L. An OD₄₀₅ ratio of OD₄₀₅ CML66-L peptide/OD₄₀₅ CML66-S peptide higher than 3 SD above the mean ratio of OD₄₀₅ was interpreted as positive. As shown in Fig. 4B, the ELISA results demonstrated that there were no significant differences in IgG Ab responses to CML66-L peptide and CML66-S peptide in the sera of normal donors. The mean OD₄₀₅ ratio plus 3 SD of CML66-L peptide/CML66-S peptide was 1.094. In contrast, the OD₄₀₅ ratios of CML66-L peptide/CML66-S peptide in the sera from patients with tumors (four with CML and three with prostate cancer) were significantly higher than that of the mean plus 3 SD in normal donors ($p < 0.01$). In addition to the OD₄₀₅ ratios of CML66-L peptide/CML66-S peptide (Fig. 4B), the mean OD₄₀₅ plus 3 SD of CML66-S peptide in the sera of normal donors was 0.129. The OD₄₀₅ readings in the sera of patients with tumors were all < 0.126 , suggesting that sera from patients with tumors had no significant IgG Ab responses to CML66-S peptide. Our previous ELISA data with the full-length CML66-L showed that these patients with tumor had positive IgG Ab responses to CML66 (not shown) (10). The results suggest that IgG Ab responses to CML66-L peptide in patients with tumors were significantly higher than that to the CML66-S peptide, and suggest that CML66-L is the dominant immunogenic form.

Furthermore, we hypothesized that if CML66-L is the dominant immunogenic form, specific Ab responses to CML66-L should be correlated with the specific Abs to the CML66-L-specific N-terminal region. To test this hypothesis, we constructed a CML66-L-specific GST fusion protein by fusing the cDNA encoding the N-terminal 38 aa of CML66-L onto the C terminus of GST. In addition to the DNA sequences of cDNA inserts in GST fusion protein vectors, the authenticity of GST-CML66-L full length constructed previously (10) and GST-CML66-L-38 were confirmed by Western blot with anti-GST Ab (not shown) and N-terminal protein sequencing (not shown). Among eight patients, including four patients with CML and four patients with prostate cancer whose specific Abs to CML66-L full length were statistically higher, all eight patients also had statistically higher Ab titers to CML66-L-38. The OD₄₀₅ ratios of GST-CML66-L full length/GST and the OD₄₀₅ ratios of GST-CML66-L-38/GST for all of the eight positive samples (A–H) were listed in Fig. 4D. These results demonstrated that specific Ab responses to CML66-L were well correlated with the specific Abs to CML66-L-38. Preincubation of patient sera with purified recombinant GST-CML66-L full length or GST-CML66-L-38 specifically inhibited ELISA activity, but preincubation with GST alone did not reduce reactivity measured by ELISA (not shown), suggesting that the Ab responses in patient sera to CML66-L or CML66-L-38 are specific. Taken together, these results further confirmed that CML66-L, predominantly expressed in tumor cells, is mainly responsible for epitope generation of this tumor Ag.

Discussion

The transcription of CT Ags in tumor cells was found to be associated with general demethylation in tumor cells (24). However, the precise relation between the expression of CT Ags and their immunogenicity in patients with tumors remains largely elusive, although the threshold levels of tumor Ag expression were recently considered to be important for tumor cell recognition by CTLs (25). The expression profile suggested, by the broadened definition of CT Ags that are mostly expressed in testis and a variety of tumor cells, that CML66-L may be a new member of CT Ag group (4, 10). Therefore, we chose CML66 as a model to further study the mechanism of immunogenicity of CT Ags and self-tumor Ags with similar expression pattern. In this report, we identified a new CML66 isoform, CML66-S. CML66-S is exclusively expressed in testis but expressed in tumor cells at much lower levels, whereas CML66-L is highly expressed in tumor cells in addition to testis. A previous report showed a correlation between mRNA levels and protein expression for CT Ags (25). Our data on the immunogenicity of these CML66 isoforms correlated well with the expression of CML66-L and CML66-S in tumor cells, suggesting that RNA levels of CML66 isoforms could be extrapolated to protein values. This conclusion was further supported by the recent results from Dr. J. Ritz's lab (Harvard) in the examination of CML66-L protein expression in tumor cells from patients with specific mAbs (26). More importantly, the overexpression of CML66-L in tumor cells is associated with the immunogenicity of this Ag in patients with tumors. We recently characterized another new CT Ag termed CML28 (9, 15). Similar to CML66-L, CML28 was broadly immunogenic in patients with tumors, and the immunogenicity of CML28 is likely to be related to the overexpression of its mRNA transcript and protein in tumor cells (15). This mechanism of immunogenicity has been proposed for other CT Ags, such as NY-ESO-1 (27–29) and the MAGE Ags (30).

To test the hypothesis that CML66-L, predominantly expressed in tumor cells, is mainly responsible for epitope generation of this tumor Ag, we have used two approaches. First, we showed that IgG Ab responses to CML66-L peptide in patients with tumors were significantly higher than those to the CML66-S peptide, suggesting that CML66-L is the dominant immunogenic form. It was our intention in choosing the CML66-S-specific peptide as an appropriate control to demonstrate that CML66-L is the dominant immunogenic form, rather than emphasizing that CML66-S, which is hardly expressed in tumor cells, was not immunogenic. Despite the difference in their N-terminal 9 aa, antigenic index analysis of CML66-L peptide and CML66-S peptide with the Jameson-Wolf algorithm (31) (Lasergene; DNASTAR, Madison, WI) indicated that both peptides were very similar in the antigenicity index (not shown). Furthermore, the N-terminal 10 aa from CML66-L or CML66-S are similar in hydrophilicity (not shown). Taken together, the difference in Ab titers to CML66-L peptide or to CML66-S peptide reflects the difference in Ag expression in patients with tumors but not their differences in the antigenicity index per se of the two peptides. Second, we used the GST-fusion protein approach to test our hypothesis that, if CML66-L is the dominant immunogenic form, specific Ab responses to full-length CML66-L should be correlated with the specific Abs to the CML66-L-specific N-terminal region. The short length of the CML66-S-specific N terminus (10 aa) prevented us from using the GST fusion protein approach to compare the CML66-S N terminus with the CML66-L N terminus. In some patients with prostate cancer, the Ab titers to CML66-L were much higher than those to CML66-L-38. These differences may be due to the fact that, first, the common region of CML66-L also contributed to the epitope generation significantly, and second, Ab responses to CML66-L in patients with prostate cancer were stronger than those in patients with CML, as described previously (10).

There are numerous oncogenes to be amplified in human tumors, including *ABL*, *NRAS*, and *MYCL* (32, 33). Ab response against *MYCL* was shown to correlate with overexpression of the gene (34). In addition, overexpression of self-Ag in patients with tumors and autoimmune diseases is among the major mechanisms to break immune self-tolerance (35–37). Possibly, the overexpression of self-Ags, such as CML66-L, overcomes the threshold of Ag concentration at which an immune response is initiated, as Zinkernagel and coworkers (38, 39) recently suggested. Overexpressed Ags must access the Ag presentation pathway and immune system. Some mechanisms have been proposed (37). First, overexpressed Ags may be released from damaged tumor cells due to spontaneous necrosis and then become available in the extracellular environment for attack by the immune cells, potentially through cross-presentation (40, 41). Alternatively, tumor-expressed Ags can translocate across the intracellular membranes (42) via binding to heat shock protein 70 and enter the membrane exosome for the MHC class II Ag presentation pathway (43).

It is noteworthy that the *CML66* gene is located in a complex structure of human genome (8q23) (10), where *CML66* promoter region overlaps with that of the gene encoding a small protein DC6 with 102 aa (GenBank accession no. NM_020189) on the antisense strand. However, similar sense/antisense gene structure has been reported for other genes (44). Lack of the significant homology (16% with gaps) in the amino acid sequences and the dissimilarity of the secondary structures and the antigenicity profiles between *CML66* and DC6 (not shown) suggest that the DC6 protein could not have any significant contribution to the epitope generation of *CML66* proteins, regardless of potential coregulation of the two genes.

Because *CML66-L* shares the C-terminal 544-aa common region with *CML66-S*, the comparison of these two forms in the epitope generation had to focus in their distinct N termini, although it was not our intention to emphasize that the contribution to the epitope generation of the Ag by the N-terminal epitope(s) was larger than that in the C-terminal common region. In addition to our findings that up-regulation of tumor Ag *CML66-L*, but not *CML66-S*, is the mechanism for the epitope generation, we further elucidate the control mechanism in the regulation of the generation of *CML66-L* vs *CML66-S*. Our results suggest that alternative promoter usage in combination with alternative splicing is the mechanism for the differential expression of *CML66-L* and *CML66-S* in tumor cells and, ultimately, the epitope generation of the Ag. Alternative use of multiple promoters, each of them associated with a specific 5'-UTR exon in a tissue-specific manner, was reported for other genes, such as fibroblast growth factor-1 (45, 46). A previous report identified breast cancer Ag ING1 with several transcripts (47). In addition, aberrant tumor Ag PDZ-73 isoform expression in tumors together with amplified expression of isoforms and unidentified mutation as three possibilities have been proposed to be the immunogenic event for PDZ-73 (48). Therefore, this regulation mechanism on the immunogenicity of unmutated self-Ags was not unique to the *CML66* gene. The regulation mechanism reported in this study could be recognized as one of the major mechanisms for unmutated self-Ags, because more studies are rapidly being accomplished in the postgenomics era with regard to alternative transcription and processing of the tissue-specific expression of isoforms of self-Ags. Some other factors contributing to the immunogenicity of autoantigens and self-tumor Ags have also been proposed (49). For example, antigenic determinants with low avidity in their interaction with the T cells are displayed in thymus, allowing self-reactive T cells to escape into the periphery (50). The low levels of self-reactive T cells in the peripheral repertoire may contribute to immune responses to the concentration changes of self-Ags (51). In this report, we have intentionally demonstrated that the alternative promoter in combination with alternative splicing is a novel mechanism for the regulation of the immunogenicity of self-tumor Ags, which could be of importance in the understanding of

potential roles of autoantigens in autoimmune diseases (52) as well as development of self-tumor Ag-based immunotherapy for tumors (53).

Acknowledgments

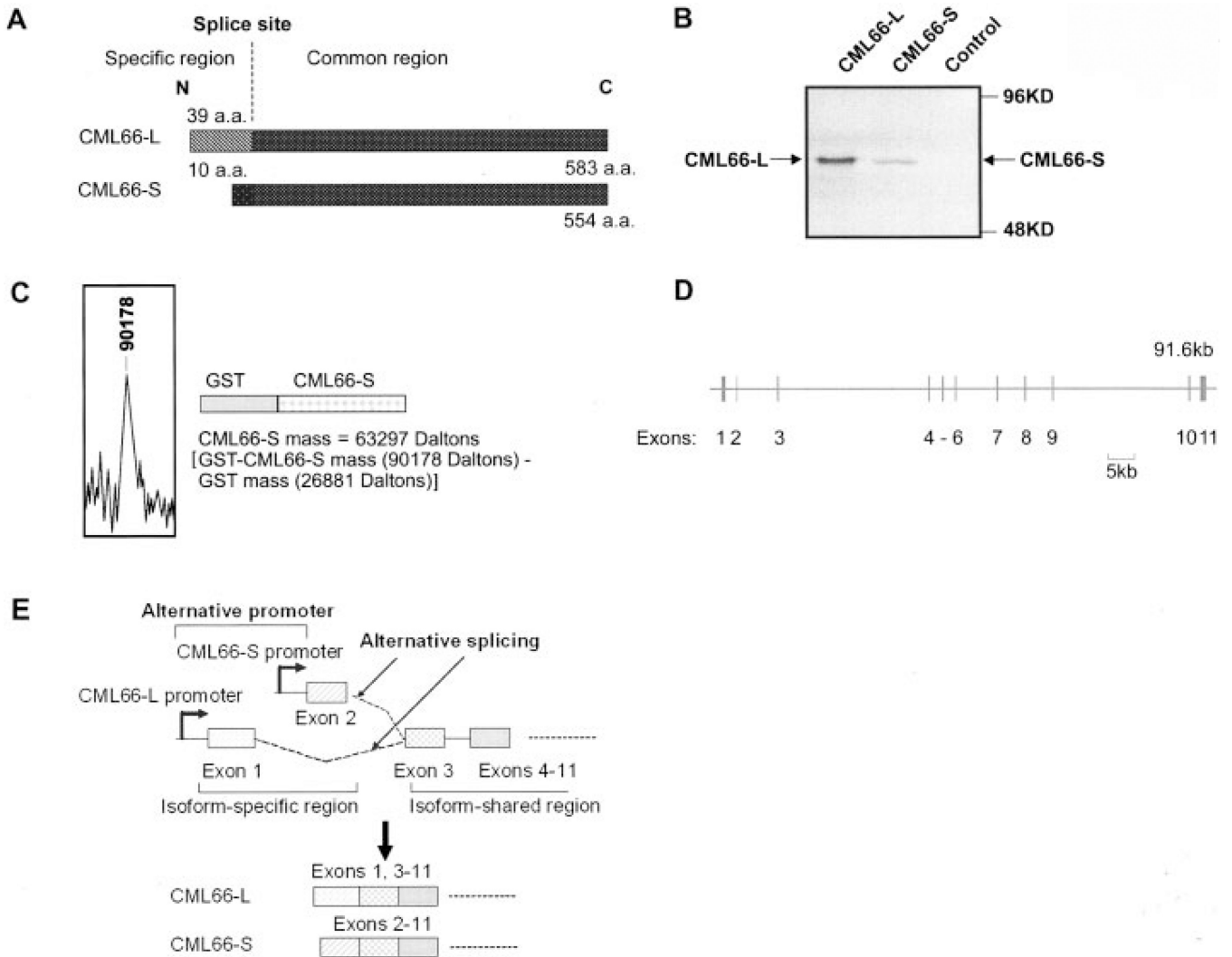
We are grateful to Drs. H. Cantor and J. Ritz at Harvard for mentoring (of X.-F. Yang); Drs. G. Dranoff and P. Kantoff at Harvard; R. B. Ren at Brandeis (Waltham, MA); J. C. D'Halluin in France; R. Wang, D. P. Huston, A. I. Schafer, T. Thompson, J. Prchal, M. Estes, D. Graham, M. Lieberman, M. Brenner, R. Cook, D. Lewis, and T. H. Tan at Baylor; Drs. Lachman at M. D. Anderson Cancer Center and T. D. Marriott at Rice University for their support and suggestions; and also Ms. K. Franks, T. Kocurek, A. Wirt, and K. Jolivette in our section, and Ms. A. Green of M. D. Anderson Cancer Center, for their assistance.

References

1. Boon T, van der Bruggen P. Human tumor antigens recognized by T lymphocytes. *J. Exp. Med.* 1996; 183:725. [PubMed: 8642276]
2. Chen YT, Scanlan MJ, Sahin U, Tureci O, Gure AO, Tsang S, Williamson B, Stockert E, Pfreundschuh M, Old LJ. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc. Natl. Acad. Sci. USA.* 1997; 94:1914. [PubMed: 9050879]
3. Old LJ, TChen Y. New paths in human cancer serology. *J. Exp. Med.* 1998; 187:1163. [PubMed: 9547328]
4. Chen YT, Old LJ. Cancer-testis antigens: targets for cancer immunotherapy. *Cancer J. Sci. Am.* 1999; 5:16. [PubMed: 10188055]
5. Streilein JW. Unraveling immune privilege. *Science.* 1995; 270:1158. [PubMed: 7502038]
6. Sahin U, Tureci O, Pfreundschuh M. Serological identification of human tumor antigens. *Curr. Opin. Immunol.* 1997; 9:709. [PubMed: 9368781]
7. Tureci O, Sahin U, Pfreundschuh M. Serological analysis of human tumor antigens: molecular definition and implications. *Mol. Med. Today.* 1997; 3:342. [PubMed: 9269687]
8. Chen YT, Gure AO, Tsang S, Stockert E, Jager E, Knuth A, Old LJ. Identification of multiple cancer/testis antigens by allogeneic antibody screening of a melanoma cell line library. *Proc. Natl. Acad. Sci. USA.* 1998; 95:6919. [PubMed: 9618514]
9. Wu CJ, Yang XF, McLaughlin S, Neuberg D, Canning C, Stein B, Alyea EP, Soiffer RJ, Dranoff G, Ritz J. Detection of a potent humoral response associated with immune-induced remission of chronic myelogenous leukemia. *J. Clin. Invest.* 2000; 106:705. [PubMed: 10974024]
10. Yang XF, Wu CJ, McLaughlin S, Chillemi A, Wang KS, Canning C, Alyea EP, Kantoff P, Soiffer RJ, Dranoff G, Ritz J. CML66, a broadly immunogenic tumor antigen, elicits a humoral immune response associated with remission of chronic myelogenous leukemia. *Proc. Natl. Acad. Sci. USA.* 2001; 98:7492. [PubMed: 11416219]
11. Gure AO, Stockert E, Arden KC, Boyer AD, Viars CS, Scanlan MJ, Old LJ, TChen Y. CT10: a new cancer-testis (CT) antigen homologous to CT7 and the MAGE family, identified by representational-difference analysis. *Int. J. Cancer.* 2000; 85:726. [PubMed: 10699956]
12. Yang XF, Weber GF, Cantor H. A novel Bcl-x isoform connected to the T cell receptor regulates apoptosis in T cells. *Immunity.* 1997; 7:629. [PubMed: 9390687]
13. Dawson-Saunders, B.; Trapp, RG. *Basic & Clinical Biostatistics.* Norwalk, CT: Appleton & Lange; 1994.
14. Kerr MK, Churchill GA. Statistical design and the analysis of gene expression microarray data. *Genet. Res.* 2001; 77:123. [PubMed: 11355567]
15. Yang XF, Wu CJ, Chen L, Alyea EP, Canning C, Kantoff P, Soiffer RJ, Dranoff G, Ritz J. CML28 is a broadly immunogenic antigen, which is overexpressed in tumor cells. *Cancer Res.* 2002; 62:5517. [PubMed: 12359762]
16. Kozak M. Interpreting cDNA sequences: some insights from studies on translation. *Mamm. Genome.* 1996; 7:563. [PubMed: 8679005]
17. Lewin, B. Nuclear splicing. In: Lewin, B., editor. *Genes VII.* New York: Oxford Univ. Press; 2000.

18. Ladd AN, Cooper TA. Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biol.* 3:reviews0008. 2002
19. Gallego ME, Nadal-Ginard B. Myosin light-chain 1/3 gene alternative splicing: cis regulation is based upon a hierarchical compatibility between splice sites. *Mol. Cell. Biol.* 1990; 10:2133. [PubMed: 2325649]
20. Johnson, W.; Jameson, JL. Transcriptional control of gene expression. In: Jameson, JL., editor. *Principles of Molecular Medicine*. Totowa, NJ: Humana; 1998. p. 25
21. Lewin, B. Initiation of transcription. In: Lewin, B., editor. *Genes VII*. New York: Oxford Univ. Press; 2000. p. 616
22. Grillot DA, Gonzalez-Garcia M, Ekhterae D, Duan L, Inohara N, Ohta S, Seldin MF, Nunez G. Genomic organization, promoter region analysis, and chromosome localization of the mouse *bcl-x* gene. *J. Immunol.* 1997; 158:4750. [PubMed: 9144489]
23. Tolnay M, Vereshchagina LA, Tsokos GC. NF- κ B regulates the expression of the human complement receptor 2 gene. *J. Immunol.* 2002; 169:6236. [PubMed: 12444129]
24. De Smet C, Lurquin C, Lethe B, Martelange V, Boon T. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol. Cell. Biol.* 1999; 19:7327. [PubMed: 10523621]
25. Riker AI, Kammula US, Panelli MC, Wang E, Ohnmacht GA, Steinberg SM, Rosenberg SA, Marincola FM. Threshold levels of gene expression of the melanoma antigen gp100 correlate with tumor cell recognition by cytotoxic T lymphocytes. *Int. J. Cancer.* 2000; 86:818. [PubMed: 10842196]
26. Wu CJ, Chen L, Kutok JL, Canning CM, Alyea EP, Yang XF, Soiffer RJ, Ritz J. CML28 and CML66 are GVL-associated antigens that are highly expressed in proliferating cells. *Blood.* 2002; 100:581a.
27. Jager E, Chen YT, Drijfhout JW, Karbach J, Ringhoffer M, Jager D, Arand M, Wada H, Noguchi Y, Stockert E, et al. Simultaneous humoral and cellular immune response against cancer-testis antigen NY-ESO-1: definition of human histocompatibility leukocyte antigen (HLA)-A2-binding peptide epitopes. *J. Exp. Med.* 1998; 187:265. [PubMed: 9432985]
28. Jager E, Gnjjatic S, Nagata Y, Stockert E, Jager D, Karbach J, Neumann A, Rieckenberg J, Chen YT, Ritter G, et al. Induction of primary NY-ESO-1 immunity: CD8⁺ T lymphocyte and antibody responses in peptide-vaccinated patients with NY-ESO-1⁺ cancers. *Proc. Natl. Acad. Sci. USA.* 2000; 97:12198. [PubMed: 11027314]
29. Jager E, Nagata Y, Gnjjatic S, Wada H, Stockert E, Karbach J, Dunbar PR, Lee SY, Jungbluth A, Jager D, et al. Monitoring CD8 T cell responses to NY-ESO-1: correlation of humoral and cellular immune responses. *Proc. Natl. Acad. Sci. USA.* 2000; 97:4760. [PubMed: 10781081]
30. Jungbluth AA, Busam KJ, Kolb D, Iversen K, Coplan K, Chen YT, Spagnoli GC, Old LJ. Expression of MAGE-antigens in normal tissues and cancer. *Int. J. Cancer.* 2000; 85:460. [PubMed: 10699915]
31. Beattie J, Fawcett HA, Flint DJ. The use of multiple-pin peptide synthesis in an analysis of the continuous epitopes recognised by various anti-(recombinant bovine growth hormone) sera: comparison with predicted regions of immunogenicity and location within the three-dimensional structure of the molecule. *Eur. J. Biochem.* 1992; 210:59. [PubMed: 1280219]
32. Schwab M. Oncogene amplification in solid tumors. *Semin. Cancer Biol.* 1999; 9:319. [PubMed: 10448118]
33. Meese E, Comtesse N. Cancer genetics and tumor antigens: time for a combined view? *Genes Chromosomes Cancer.* 2002; 33:107. [PubMed: 11793436]
34. Yamamoto A, Shimizu E, Sumitomo K, Shinohara A, Namikawa O, Uehara H, Sone S. L-Myc overexpression and detection of autoantibodies against L-Myc in both the serum and pleural effusion from a patient with non-small cell lung cancer. *Intern. Med.* 1997; 36:724. [PubMed: 9372336]
35. Chen YT. Cancer vaccine: identification of human tumor antigens by SEREX. *Cancer J. Sci. Am.* 2000; 6(Suppl. 3):S208.
36. Basten. Basis and mechanisms for self-tolerance. In: Rose, N.; Mackay, IR., editors. *The Autoimmune Diseases*. San Diego: Academic; 1998. p. 9

37. Bauer C, Diesinger I, Brass N, Steinhart H, Iro H, Meese EU. Translation initiation factor eIF-4G is immunogenic, overexpressed, and amplified in patients with squamous cell lung carcinoma. *Cancer*. 2001; 92:822. [PubMed: 11550153]
38. Ludewig B, Junt T, Hengartner H, Zinkernagel RM. Dendritic cells in autoimmune diseases. *Curr. Opin. Immunol.* 2001; 13:657. [PubMed: 11677086]
39. Zinkernagel RM, Hengartner H. Regulation of the immune response by antigen. *Science*. 2001; 293:251. [PubMed: 11452115]
40. Nagata Y, Ono S, Matsuo M, Gnjatic S, Valmori D, Ritter G, Garrett W, Old LJ, Mellman I. Differential presentation of a soluble exogenous tumor antigen, NY-ESO-1, by distinct human dendritic cell populations. *Proc. Natl. Acad. Sci. USA*. 2002; 99:10629. [PubMed: 12138174]
41. Steinman RM, Inaba K, Turley S, Pierre P, Mellman I. Antigen capture, processing, and presentation by dendritic cells: recent cell biological studies. *Hum. Immunol.* 1999; 60:562. [PubMed: 10426272]
42. Davidoff AM, Iglehart JD, Marks JR. Immune response to p53 is dependent upon p53/HSP70 complexes in breast cancers. *Proc. Natl. Acad. Sci. USA*. 1992; 89:3439. [PubMed: 1373500]
43. Thery C, Boussac M, Veron P, Ricciardi-Castagnoli P, Raposo G, Garin J, Amigorena S. Proteomic analysis of dendritic cell-derived exosomes: a secreted subcellular compartment distinct from apoptotic vesicles. *J. Immunol.* 2001; 166:7309. [PubMed: 11390481]
44. Coriton O, Lepourcelet M, Hampe A, Galibert F, Mosser J. Transcriptional analysis of the 69-kb sequence centromeric to HLA-J: a dense and complex structure of five genes. *Mamm. Genome*. 2000; 11:1127. [PubMed: 11130983]
45. Bouraima H, Hanoux V, Mittre H, Feral C, Benhaim A, Leymarie P. Expression of the rabbit cytochrome P450 aromatase encoding gene uses alternative tissue-specific promoters. *Eur. J. Biochem.* 2001; 268:4506. [PubMed: 11502211]
46. Chiu IM, Touhalisky K, Baran C. Multiple controlling mechanisms of *FGF1* gene expression through multiple tissue-specific promoters. *Prog. Nucleic Acid Res. Mol. Biol.* 2001; 70:155. [PubMed: 11642361]
47. Jager D, Stockert E, Scanlan MJ, Gure AO, Jager E, Knuth A, Old LJ, Chen YT. Cancer-testis antigens and ING1 tumor suppressor gene product are breast cancer antigens: characterization of tissue-specific ING1 transcripts and a homologue gene. *Cancer Res.* 1999; 59:6197. [PubMed: 10626813]
48. Scanlan MJ, Williamson B, Jungbluth A, Stockert E, Arden KC, Viars CS, Gure AO, Gordan JD, Chen YT, Old LJ. Isoforms of the human PDZ-73 protein exhibit differential tissue expression. *Biochim. Biophys. Acta.* 1999; 1445:39. [PubMed: 10209257]
49. Moudgil, K.; Sercarz, EE. Antigenic determinants involved in induction and propagation of autoimmunity. In: Rose, N.; Mackay, IR., editors. *The Autoimmune Diseases*. San Diego: Academic; 1998. p. 45
50. Liu GY, Fairchild PJ, Smith RM, Prowle JR, Kioussis D, Wraith DC. Low avidity recognition of self-antigen by T cells permits escape from central tolerance. *Immunity*. 1995; 3:407. [PubMed: 7584132]
51. Anderton SM, Wraith DC. Selection and fine-tuning of the autoimmune T-cell repertoire. *Nat. Rev. Immunol.* 2002; 2:487. [PubMed: 12094223]
52. Robinson WH, Steinman L, Utz PJ. Proteomics technologies for the study of autoimmune disease. *Arthritis Rheum.* 2002; 46:885. [PubMed: 11953963]
53. Rosenberg SA. Progress in human tumour immunology and immunotherapy. *Nature*. 2001; 411:380. [PubMed: 11357146]

**FIGURE 1.**

Structure of CML66-S in comparison with that of CML66-L. **A**, Schematic comparison of the CML66-S ORF with that of CML66-L. CML66-S has an ORF with 554 aa. The C-terminal 544-aa region and the 3'-UTR are the same in both CML66-L and CML66-S. The novel sequence in the 5' end of CML66-S results in a replacement of 39 aa of CML66-L by a novel N terminus of 10 aa of CML66-S. The dashed line shows the splicing site. **B**, [³⁵S]Methionine-labeled in vitro TNT products of CML66-L and CML66-S cDNAs separated on SDS-PAGE and revealed by autoradiography. The cloning vector without cDNA insert was used as a control in TNT. CML66-S cDNA can be transcribed and translated into a protein with an apparent molecular mass of 63.5 kDa, slightly smaller than that of CML66-L (66 kDa). The arrows show bands of CML66-L and CML66-S, respectively. **C**, The mass spectra of GST-CML66-S fusion protein. On the *right* side, a schematic representation of GST-CML66-S fusion protein is shown together with the CML66-S mass (63,297 Da) calculated by subtracting the the GST mass (26,881 Da) from the GST-CML66-S mass (90,178 Da). **D**, Schematic representation of overall structure of *CML66* gene locus. By comparison of CML66-L and CML66-S cDNA sequences with the sequence of the human genomic clone RP11-122A21, an overall structure of the *CML66* gene locus was generated. The *CML66* gene has a total of 11 exons and spans a 91.6-kb-

long region. The CML66 common region is encoded by exons 3–11. Therefore, CML66-L is encoded by all of the exons except exon 2 (exons 1 and 3–11). In contrast, CML66-S is encoded by all of the exons except exon 1 (exons 2–11). *E*, Schematic representation of alternative promoter and alternative splicing mechanisms in the generation of isoforms CML66-L and CML66-S. In addition to alternative promoter functioning in the transcription of the CML66 L-specific 5' transcripts encoded by exon 1 driven by CML66-L promoter, or the CML66-S-specific 5' transcripts encoded by exon 2 driven by CML66-S promoter, the alternative splicing mechanism also participates in splicing either exon 1 transcripts into the isoform-shared exons 3–11 to make CML66-L transcripts, or exon 2 transcripts into the exon 3–11 to make CML66-S transcripts. The schematic organizations of isoforms CML66-L transcripts and CML66-S transcripts are also presented.

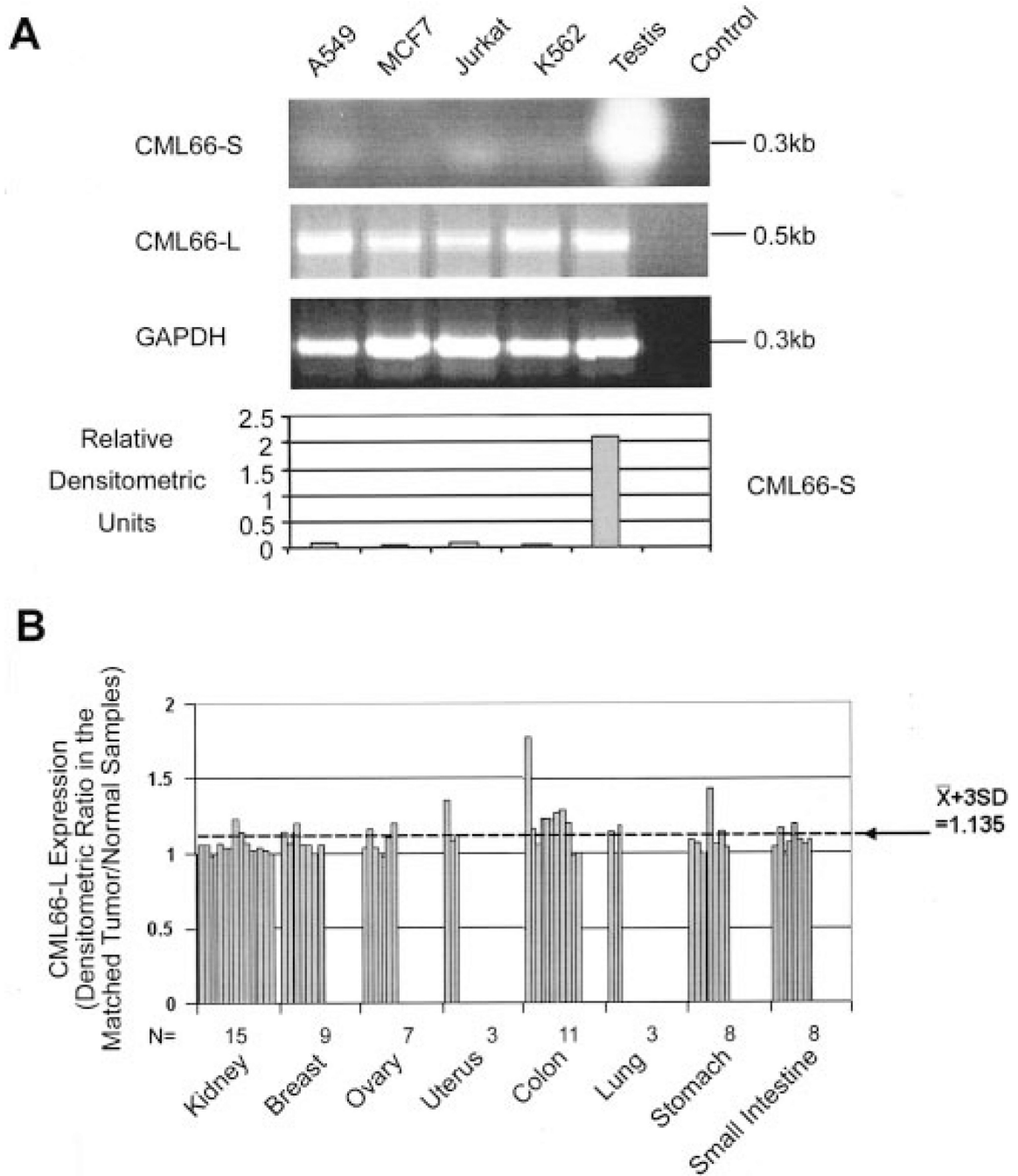
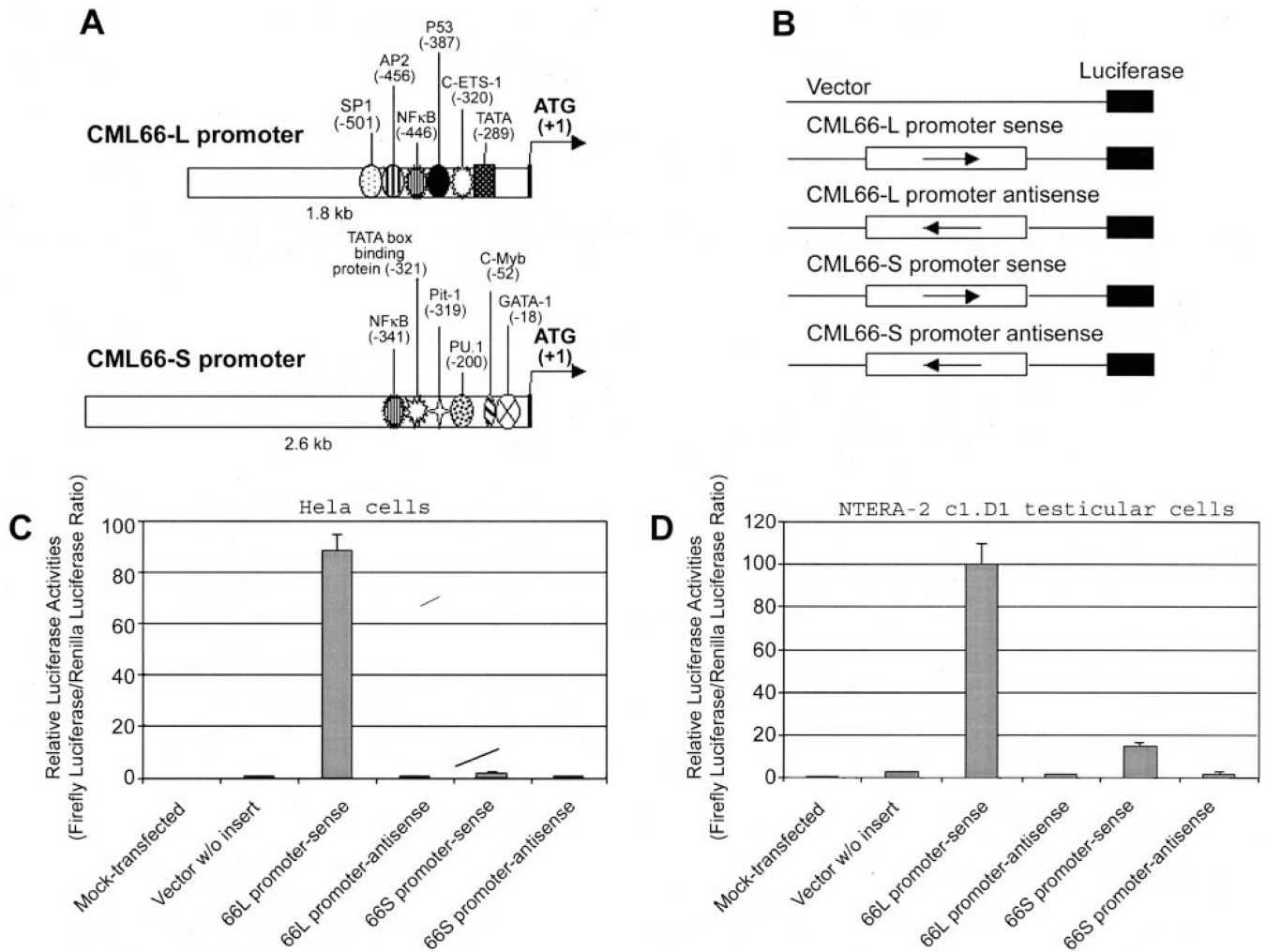


FIGURE 2.

Predominant expression in tumor cells of CML66-L but not CML66-S. A, The tissue expression of CML66-S in tumor cell lines vs testis was examined by semiquantitative RT-PCR revealed by 1% agarose gel followed by Southern blot hybridization (not shown). In addition to RNA samples from cells, no RNA control was used as a negative control in RT-PCR. CML66-S-specific transcripts were amplified in RT-PCR with 35 cycles to detect CML66-S transcripts at very low levels in tumor cell lines. As controls, CML66-L-specific transcripts and GAPDH transcripts were also examined in RT-PCR with 30 cycles. The lower panels show the ratios of CML66-S signal over GAPDH signal (internal housekeeping gene control) as the normalized quantitation (relative densitometric units) of CML66-S

expression in each sample, respectively. The expression of CML66-S was mostly found in testis but was expressed in very low levels in tumor cell lines even after 35 cycles of PCR amplification. In contrast, the CML66-L transcript was detected in four tumor cell lines and human testis with 30 cycles of PCR amplification. *B*, The expression of CML66-L in tumor samples from patients. The hybridization analyses of the matched tumor/normal expression array (BD Clontech) with ³²P-labeled CML66-L-specific probe, CML66-S-specific probe, and ubiquitin probe were performed, respectively. The densitometric ratios of CML66-L expression in the matched tumor samples vs that in matched normal tissues were calculated. The statistical mean (\bar{X}) plus 3 SD (1.135) of CML66-L expression signals among normal samples is shown as the variation range of CML66 signal among the normal samples. CML66-L expression signals in tumors higher than this upper limit of normal range are considered positive.

**FIGURE 3.**

Predominant expression of CML66-L in tumor cells resulted from tumor-specific transcription of CML66-L promoter. *A*, The schematic representation of CML66-L-specific promoter (1.8 kb) and CML66-S promoter (2.6 kb). Several transcription factor binding sites and their positions are marked on the CML66-L promoter and CML66-S promoter, respectively. *B*, Constructs of CML66-L promoter and CML66-S promoter in both sense and antisense orientations on the pGL-3 luciferase enhancer reporter vector. CML66-L promoter antisense and CML66-S promoter antisense were also constructed as controls, because it is well accepted that promoter activities are orientation dependent. *C*, CML66-L and CML66-S promoter activities in HeLa cells. The mean ratios of firefly luciferase activities conferred by reporter constructs vs *Renilla* luciferase activities (internal control) and 2 SD are shown as relative luciferase activities for every experimental group. The error bars indicating the upper limit in all the groups, except CML66-L promoter sense orientation, are hardly visible due to the scale. The experiments were repeated four times, and the representative results are shown. Notably, similar results were obtained in HCCT, a hepatic carcinoma cell line, after transfection with the same sets of reporter luciferase vectors (not shown). *D*, CML66-L and CML66-S promoter activities in NTERA-2 cl.D1 human testicular embryonal carcinoma cells. The mean ratios of firefly luciferase activities conferred by reporter constructs vs *Renilla* luciferase activities (internal control) and 2 SD are shown as relative luciferase activities for every experimental group. The error bars indicating the upper limit in all the

groups, except CML66-L promoter sense orientation construct and CML66-S promoter sense orientation construct, are hardly visible due to the scale. The experiments were repeated four times, and the representative results are shown.

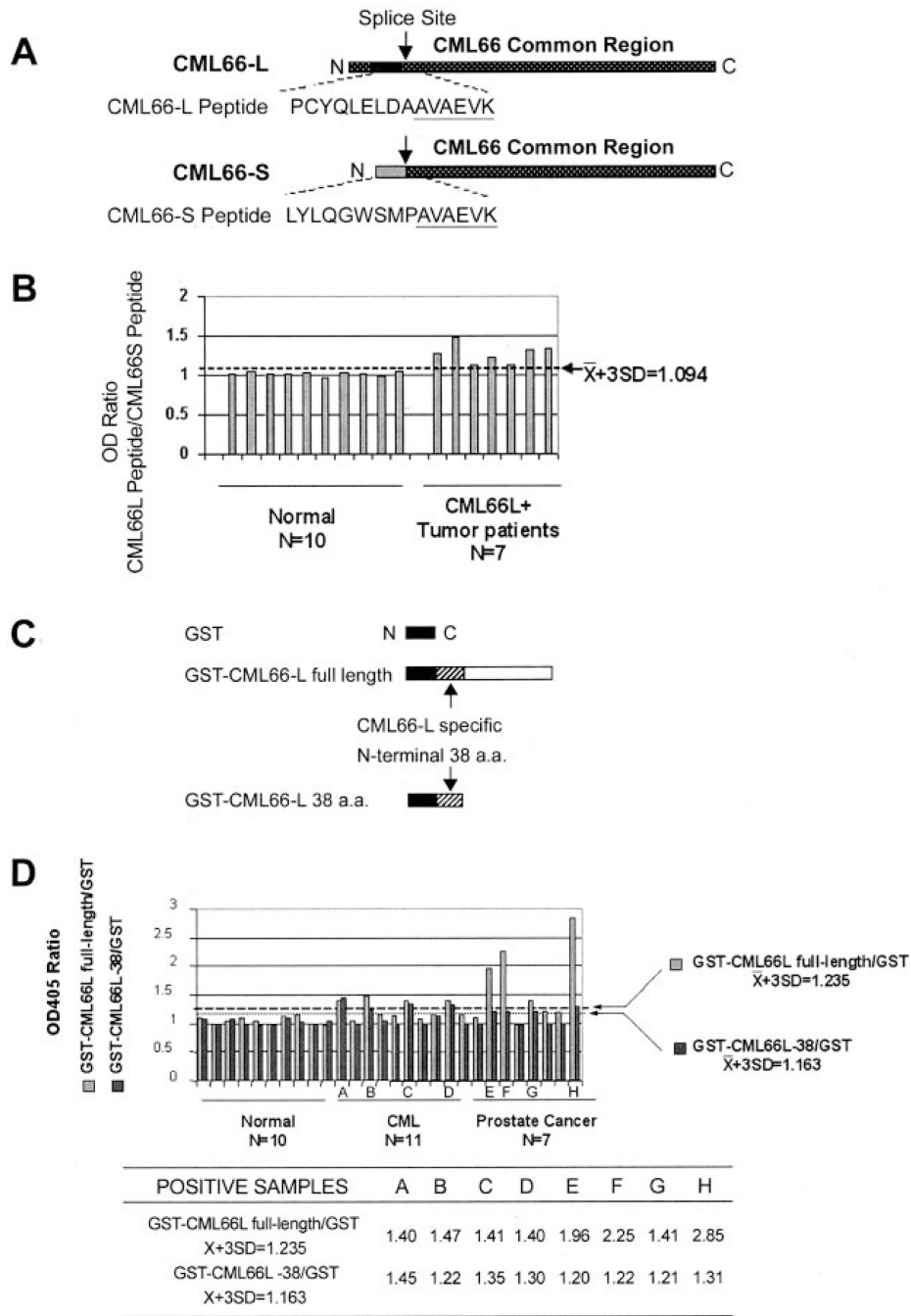


FIGURE 4. Immunogenicity of CML66 contributed mostly by CML66-L, but not CML66-S. **A**, Design of CML66-L peptide and CML66-S peptide. The arrow indicates the splice site in CML66-L and CML66-S, respectively. The 15-mer CML66-L-specific peptide covering the splicing site was synthesized with the N-terminal 9 aa unique to the N terminus of CML66-L and the C-terminal 6 aa (underlined) in the CML66 common region. In comparison, the 15-mer CML66-S peptide, consisting of the N-terminal 9 aa from the CML66-S and the C-terminal 6 aa from CML66 common region (underlined), was also synthesized. **B**, Positive Ab responses to the CML66-L-specific peptide but not to the CML66-S peptide in patients with

tumors. The mean OD_{405} ratio plus 3 SD of Abs to CML66-L peptide/Abs to CML66-S peptide in normal controls was 1.094. In contrast, the OD ratios of Abs to CML66-L peptide/Abs to CML66-S peptide in the sera from CML66-L⁺ patients with tumors (four with CML and three with prostate cancer) were significantly higher than the mean plus 3 SD in normal donors. The CML66-L⁺ patients with tumors were chosen based on the previous results showing positive Abs responses to CML66-L full-length recombinant protein (not shown) (10). *C*, Schematic representation of GST fusion proteins of GST-CML66-L full length and GST-CML66-L-38. *D*, Positive correlation of Ab responses to CML66-L-specific N terminus with that to full-length CML66-L. The OD_{405} ratio of GST-CML66-L over GST was considered as the normalized Ab response to CML66-L. Similarly, the OD_{405} ratio of GST-CML66-L-38 over GST was considered as the normalized Ab response to CML66-L-38. The mean plus 3 SD of the OD_{405} ratios of GST-CML66-L full length over GST from 10 normal donors are calculated as the upper limit (1.235) of the normal range of Ab responses to CML66-L full length. Similarly, the mean plus 3 SD (1.163) of the OD_{405} ratios of GST-CML66-L-38 over GST from 10 normal donors is also calculated. Ab responses to CML66-L or CML66-L-38 in sera from patients with tumor, including 11 patients with CML and 7 patients with prostate cancer, higher than the normal ranges are considered as positive. The experiments were repeated three times, and the representative results are shown. For eight patients (A–H) with tumors whose sera had both positive IgG Ab responses to CML66-L full-length and positive IgG Ab responses to CML66-L-38, the OD_{405} ratios of GST-CML66-L full length/GST and the OD_{405} ratios of GST-CML66-L-38/GST are listed in a table in the lower panel of *D*. The results show that Ab responses to CML66-L-38 in patients with tumors are correlated with those to CML66-L full length.