

Molecular Cloning of an Alternative Human α E-Catenin cDNA

Marianne Linkels, Marion J. G. Bussemakers, Friedel Nollet,* Charles M. Ewing,†
Frans van Roy,* and Jack A. Schalken

Urology Research Laboratory, University Hospital Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands;

**Department of Molecular Biology, VIB–University of Gent, 9000 Ghent, Belgium; and †Brady*

Urological Institute, Johns Hopkins Hospital, Baltimore, Maryland 21205

Received June 26, 1997

The cytoplasmic protein alpha-catenin plays a crucial role in E-cadherin mediated cell-cell adhesion by binding E-cadherin to the cytoskeleton via beta- or gamma-catenin and actin. Functional loss of one of these interacting components leads to decreased cell-cell adhesion, and therefore to loss of epithelial integrity. Northern analysis revealed two distinct α E-catenin transcripts in different cell lines, whereas apparently only one protein is expressed. Because of the biological importance of this protein we sought to molecularly characterize the differences between the two observed transcripts. cDNA cloning and sequence analysis revealed the earlier described 3.4 kb α E-catenin transcript and an α E-catenin transcript of approximately 3.8 kb. This larger transcript contains a 321 bp extension in the 3'UTR sequence, which probably arises as a result of alternative polyadenylation. Considering the presence of AU-rich sequences in the extension, it may be involved in mRNA stability. © 1997

Academic Press

The classical cadherins are transmembrane glycoproteins responsible for calcium dependent cell-cell adhesion (1). For example E-cadherin, localized in the adherent junctions of epithelial cells, plays an important role in maintaining epithelial integrity (2-4). Its intercellular adhesive function depends on the binding to catenins (α , β , γ) (5). Beta- and gamma-catenin (plakoglobin), both members of the armadillo gene family (6), can bind directly in a mutually exclusive manner to the cytoplasmic part of E-cadherin (7). α E-catenin connects these E-cadherin-beta/gamma-catenin complexes to the cytoskeleton (8, 9) through direct association with the actin filament network of the cell (10). Alpha-catenin shows homology to the cytoplasmic protein vinculin, a major structural component of vertebrate and invertebrate cell-matrix and cell-cell junctions.

Vinculin is known to associate with different components of cell-contact sites like talin, paxillin and F-actin (11, 12). The homology between vinculin and α E-catenin is restricted to three distinct regions (13, 14), suggesting the involvement of these regions in comparable biological functions.

Until now, two α -catenin genes have been identified, i.e. α N-catenin (neural) and α E-catenin (epithelial). Both types have been cloned in mouse (13, 14), human (15, 16, 17), and *Drosophila* (18). The human α E-catenin gene (CTNNA1) is mapped to chromosome band 5q31 and consists of 16 coding exons of which the last exon contains the TAA stopcodon and a 723 bp 3'UTR sequence (19). Besides this coding gene also a human α E-catenin pseudogene which shows 90% similarity to the α E-catenin mRNA at the nucleotide level, has been identified (20). The pseudogene is mapped to chromosome band 5q22 (20).

Recently, it has become clear that loss of E-cadherin mediated adhesion is an important step in the progression of carcinomas. One of the mechanisms leading to defective cadherin function is impaired association with the cytoskeleton due to the absence or dysfunction of α E-catenin. In human colon cancer cells, transition from a noninvasive to an invasive phenotype is associated with down-regulation of α -catenin (21). That this can be biologically relevant is supported by the fact that E-cadherin mediated interactions can be restored by transfecting α -catenin cDNAs into α -catenin-deficient PC3 or PC9 cells. These cells have lost α E-catenin expression due to a homozygous deletion of a large portion of the coding region of the α E-catenin gene (22, 23). They normally show a fibroblastic-like growth pattern, but transfection of either α N-catenin cDNA or α E-catenin cDNA into, respectively, PC9 (24) and PC3 (25) cells resulted in colonies that have an altered, tightly packed, epithelioid morphology, supporting the hypothesis that α -catenin plays a crucial role in E-cadherin mediated cell adhesion.

Two distinct α E-catenin transcripts with different relative expression levels have been observed in some cell lines, although, western analysis revealed a single 110 kD band, suggesting that only one α -catenin protein is expressed. In order to resolve the molecular basis for the complexity in the α -catenin mRNA expression pattern, we screened a human prostate cDNA library and identified two distinct human α -catenin cDNA transcripts. Here, we describe the molecular characterization of the differences between these two transcripts. Furthermore, we performed computer-assisted analysis of α E-catenin related expressed sequence tags (ESTs). This confirmed that the different transcripts arise due to alternative polyadenylation.

MATERIALS AND METHODS

Isolation of the human α E-catenin cDNA clones. A cDNA library was constructed from RNA isolated from the human cell line 267B1 (prostatic epithelial cell line derived from fetal prostate tissue (26)) using the ZAP-cDNA synthesis kit and ZAP-cDNA Gigapack 2 Gold cloning kit (Stratagene). The library was screened, using a 1.5 kb human α -catenin cDNA fragment as a probe (27). pBluescript SK(–) phagemids were obtained using helper phage R408 according to the in vivo excision protocol (Stratagene). A second cDNA library used, was constructed from human fetal kidney (Clontech).

cDNA sequence analysis. Conventional cloning methods were used to clone cDNA fragments into the plasmids pUC18 and pGEM3Zf(–) (27). The dideoxy termination method (28) was used for nucleotide sequence analysis. The gel readings were recorded and edited using IntelliGenetics computer software. Computer comparison studies were performed with sequences obtained from the EMBL and Genbank nucleotide sequence databases (29), using CAMMSA computer software.

EST alignment. α E-catenin specific EST (expressed sequence tags) sequences were selected from original ABI-sequence profiles from the Washington University-Merck EST project (genome.wustl.edu). These sequences were aligned and analyzed using Staden GAP4 software (30).

Northern and Southern blot analysis. Total RNA was isolated from the cell lines A431 (human cervix carcinoma cell line), LNCAP and PC3 (both human prostate cancer cell lines). 10 μ g of total RNA was glyoxalated, separated on a 1% agarose gel, capillary transferred to a Hybond-N+ membrane (Amersham) and immobilized by uv-crosslinking.

Chromosomal DNA was isolated from the above mentioned cell lines. 10 μ g of DNA was digested using endonuclease *Eco*RI, separated on a 1% agarose gel and capillary transferred to a Hybond-N+ membrane (Amersham).

The blots were hybridized according to Church and Gilbert (31) using either a human α E-catenin coding cDNA fragment of 1.5 kb, or a 350 bp 3'UTR specific fragment from clone 921 as a probe. Ribosomal RNA and γ -catenin were used as controls for the amounts of respectively total RNA and chromosomal DNA loaded. Probes were radioactive labeled with [α - 32 P]dATP by random prime labeling (32).

PCR analysis. PCR was performed in a reaction mixture of 50 μ l containing 1 \times PCR buffer (SphaeroQ), 0.2 mM of each dNTP (dATP, dCTP, dGTP, and dTTP), 0.5 pmoles of each primer, 1 mM $MgCl_2$ (primer set P3-P4) or 2.5 mM $MgCl_2$ (primer set P1-P2) and 1 unit Supertaq DNA polymerase (SphaeroQ). 2 ng plasmid DNA (positive control), 10 ng cDNA or 500 ng chromosomal DNA was used as a template. Using a Perkin Elmer thermocycler, the following program was performed: 1 min at 94°C, 1 min at 60°C, 1 min at 72°C, for 35

cycles followed by a final extension of 10 min at 72°C. The following α E-catenin specific primers were used (see also fig. 4A):

sense primer P1: 5'- GGG CAT CTC AGA AGA AGC AC -3'
anti-sense primer P2: 5'- AAA TTC ACC ACC TGG ACT GG -3'
sense primer P3: 5'- GGG AAT TTG GCT CAA CTT CA -3'
anti-sense primer P4: 5'- GTT GTC AGG TTT TTC TGC TGC -3'

10 μ l of total PCR mixture was analyzed on a 2% agarose gel.

RESULTS

Isolation of Human α -Catenin cDNA Clones

Upon screening a human prostate cDNA library with the 1.5 kb human α E-catenin probe, we obtained 120 positive clones. Analysis of 20 clones suggested that 7 clones contain the full length human α E-catenin cDNA (data not shown). Sequence analysis of these 7 clones revealed that they contain the ATG start codon, followed by a 2718 bp open reading frame from the human α E-catenin cDNA as described by Furukawa et al. (19). However, differences in 3'UTR sequences were observed (fig 1): clone 921 showed a 321 bp extension of the 3'UTR when compared to clone 1021, which is identical to the earlier described α E-catenin cDNA. Clone 1021 contains an AATAAA polyadenylation signal that starts at position 3413, clone 921 contains this sequence at the same position, but has a second polyadenylation signal at position 3731. Three out of seven sequenced clones contain the extension. Upon screening a human fetal kidney cDNA library, 6 out of 7 α E-catenin specific clones showed the originally described poly(A) addition site (19). The remaining clone was likely to be artificially primed.

EST Alignment

Database searches revealed 59 α E-catenin related ESTs available from the Washington University-Merck EST project. 32 sequences out of 59 ESTs contained clear polyadenylation tails, from which 27 start at 2 dominant positions i.e. for 13 ESTs at position 3434-3437 and for 14 ESTs at position 3755-3757 (fig 1). The position of the first predominant polyadenylation site is equal or nearby (+1 or +3 bp) to that of the previously described α E-catenin transcript (19), whereas the position of the second one is equal or nearby (+2 bp) to the one identified in clone 921. The poly(A) tail start sites of the remaining poly(A) tail containing α E-catenin specific ESTs are indicated in fig 1.

Northern Blot Analysis

In order to test the hypothesis that the extended transcript we identified is in fact the explanation for the earlier observed two distinct α E-catenin transcripts in different cell lines (33), we performed a northern blot analysis. Using an α -catenin cDNA probe, two transcripts of respectively 3.4 kb and 3.8 kb were ob-

```

                                Poly-A 1xEST
                                Poly-A clone 1021, 4xEST
                                Poly-A 8xEST
                                Poly-A 1xEST
                                Poly-A 2xEST
                                Poly-A 1xEST
                                Poly-A 6xEST
                                Poly-A clone 921,
                                8xEST
3401  CTGCAGGACA TTAATAAAGT TGCTTTTITA GGCTacagtg tctcgatgcc ataatcagaa
3461  cacactttttt ttcctcttttc tcccagcttc aaatgcaaat tcatcattgg gctcacttct
3521  aataactgca gtgtttcccg ccttgggctt gcagcagaaa aacctgacaa catagtgttt
3581  gctaaggcag taatttagac ttaccttat ttgtgattac tgtagtgatt gattgattga
3641  ttactattaa ctacaaggta taatttacta tcaccttatt taaattttat gaattaattt
3701  gaatgttttt tacactaact aacttttccc aataaagtcc actatgaaac cacgac

```

FIG. 1. Nucleotide sequence analysis of the 3'UTR extension of clone 921. The sequence of the previously described α E-catenin cDNA (clone 1021) is in capitals (position 3401-3433). Clone 921 contains an extension of 321 bp that starts at position 3434. Poly-adenylation signals are underlined and the start-sites of the poly(A) tails are indicated by vertical lines.

served in the cell lines A431 and LNCAP (fig 2A). When we used the unique 350 bp 3'UTR fragment from clone 921 as a probe, no hybridization with the 3.4 kb fragment was found, whereas the larger 3.8 kb fragment was hybridized (fig 2B). Alternative use of polyadenylation signals is therefore a likely explanation for the two α -catenin transcripts observed in different cell lines.

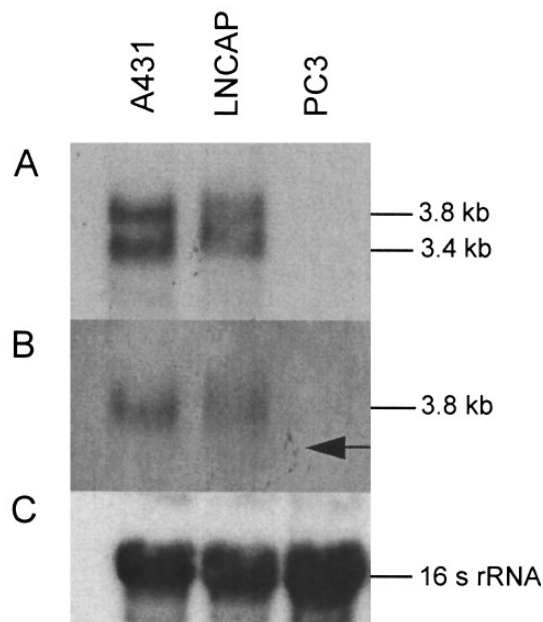


FIG. 2. Northern blot analysis of α E-catenin. Ten μ g of total RNA of different cell lines was loaded per lane. (A) α E-catenin coding cDNA probe; (B) clone 921 3'UTR extension specific probe, the arrow indicates the lost hybridization signal as compared to A; (C) rRNA was used as a control for the amounts of RNA loaded.

As expected, no transcripts were detected in cell line PC3 (fig 2A, B) due to the homozygous deletion of the functional α E-catenin gene (22).

Southern Blot Analysis

To determine whether both α E-catenin forms are derived from one α E-catenin gene, a Southern blot analysis was performed. Using the coding cDNA probe it was shown that both the α E-catenin gene and the α E-catenin-pseudogene are present in the cell lines A431 and LNCAP. Cell line PC3 contains the α E-catenin pseudogene only (fig 3A). Apparently the 90% similarity between the functional gene and the pseudogene is enough to detect the pseudogene in Southern analysis. From hybridization with the 3'UTR specific probe it became clear that the 3'UTR extension as identified in clone 921 is present in the α E-catenin gene, but absent in the pseudogene (no fragment detected in the lane containing PC3, fig 3B). The fact that the hybridizing fragment using the 3'UTR probe is of identical size (± 9.5 kb) as one of the fragments using the coding probe, suggests that there is indeed only one gene responsible for the α E-catenin transcription. The origin of the larger fragment observed in the lane containing LNCAP chromosomal DNA (fig. 3B) is not clear.

PCR Analysis

To determine if the 3'UTR extension of clone 921 is a part of the last exon (exon 16) of the α -catenin gene, we performed a PCR analysis. The forward primer is located 5' to the first polyadenylation signal of the α E-catenin transcript (P3, fig 4A) and the reverse primer is located within the extended 3'UTR sequence of clone

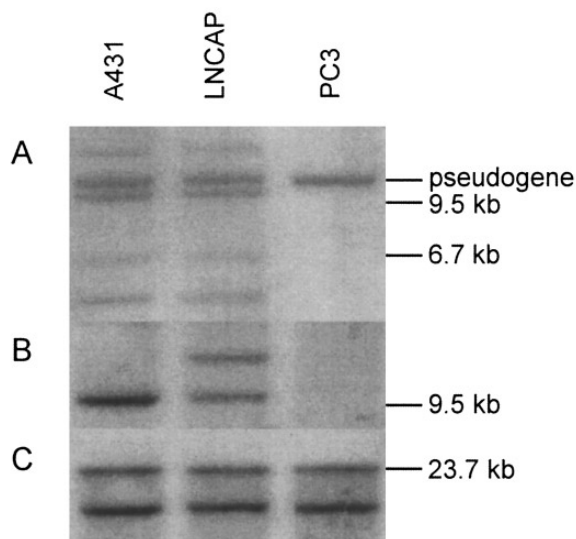


FIG. 3. Southern blot analysis of α E-catenin. Ten μ g of chromosomal DNA was loaded per lane after digestion with *Eco*RI. (A) α E-catenin coding cDNA probe, (B) clone 921 3'UTR extension specific probe, (C) γ -catenin was used as a control for the amounts of chromosomal DNA loaded.

921 (P4, fig 4A). Using these primers with both cDNA and chromosomal DNA (gDNA) from LNCAP cells as templates showed that fragments of identical sizes were amplified (fig 4B). This indicates that there is no intron present in the genomic sequence between these two primers. The 3'UTR extension in clone 921 is therefore probably transcribed from the same exon, exon 16, of which the polyadenylation signal of the shorter α -catenin transcript is described (19).

Using the same primer set P3-P4 on PC3 chromosomal DNA as template, we could not identify an amplified product. This agrees with the Southern blot analysis, where we couldn't detect the 3'UTR fragment in the PC3 cell line either. The primer set P1-P2 was used as a control for the DNA quality of the used cell lines, while clone 921 plasmid DNA and normal prostate cDNA were used as control templates for the PCR analysis.

DISCUSSION

Recent observations indicated that the α E-catenin gene is transcribed in at least two different mRNAs, even though only one protein seems to be expressed. This complex transcription of the α -catenin gene might be associated with functional regulation of the gene. In order to gain more insight into the molecular mechanism behind this phenomenon we isolated a number of human α E-catenin cDNA clones of different sizes from a prostate cDNA library. Sequence analysis showed that besides the human α E-catenin cDNA as previously described (19) another human α E-catenin transcript

was identified. This variant, represented in cDNA clone 921, showed an extension of 321 bp in the 3'UTR sequence and has a length of 3754 bp, when compared to the shorter, previously described transcript of 3433 bp. The newly identified transcript apparently results from using an alternative polyadenylation site. These data were confirmed upon aligning 59 sequences of α E-catenin EST clones. The total alignment revealed two predominant α E-catenin transcripts, with a length comparable to the transcripts we identified.

From the observation of two transcripts in the cell lines A431 and LNCAP (northern analysis) and the presence of only one α E-catenin gene (Southern analysis) it can be concluded that both transcripts are transcribed from the same gene. Furthermore, this provides the first evidence for alternative usage of polyadenylation signals in the expression of the α E-catenin gene. The finding that 6 out of 7 α E-catenin cDNA clones, derived from a fetal kidney library, showed the same poly(A) addition site, suggests some tissue specificity in the use of poly(A) addition signals. However, the 13 EST clones using the first poly(A) addition signal (position 3413) were derived from 8 different tissues, whereas the 14 EST clones using the second signal (position 3757) were derived from 11 tissues, of which 6 are shared with the first category. These preliminary data agrees against high tissue-specific preference for a particular poly(A) addition signal.

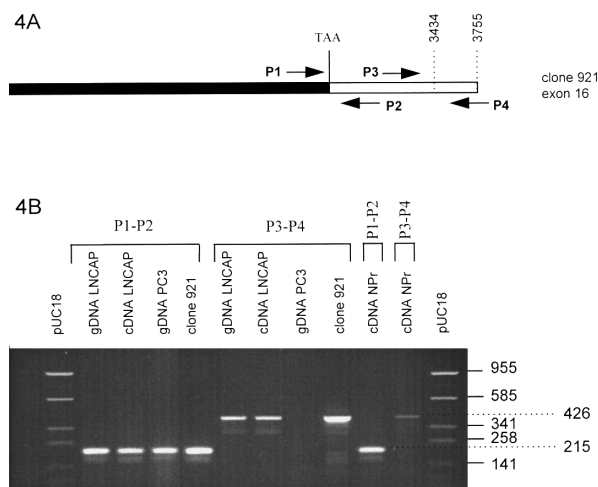


FIG. 4. (A) A schematic representation of clone 921 exon 16. The black box indicates the coding region of exon 16, the white box indicates the 3'UTR of exon 16. The position of the stop codon TAA is marked with a vertical line. The dotted vertical lines represent the ends of the transcripts of respectively α E-catenin (3434) and clone 921 (3755). The location of the primers are indicated with arrows. Coding primer set: P1-P2, fragment expected of 215 bp. 3'UTR primer set: P3-P4, fragment expected of 426 bp. (B) 1/5 of each PCR reaction mix was separated on a 2% agarose gel. gDNA: chromosomal DNA, cDNA: reverse transcriptase reaction of total RNA, NPr: normal prostate tissue. *Sau*III A digested pUC18 was used as a marker.

The α E-catenin isoforms identified before can be distinguished by a deletion (17) or an insertion (33) in the *coding* region of the α -catenin cDNA. The possible function of the larger 3'UTR which occurs in the *non-coding* region can therefore not be predicted from previous examples. Assuming that the relative expression levels of the two different transcripts are functionally relevant for the cells, it can be suggested that the 3'UTR extension has a regulatory function in mRNA stability. Studies by Zubiaga et al. (34) showed that different sequence motifs present in the 3' UTR of human genes can play a role in mRNA stability. For instance, the nonamer UUAUUUAUU is a AU-rich sequence motif that mediates mRNA degradation, whereas the length of the poly A tail might be involved in mRNA stability. Although the newly identified 3'UTR extension is an AU-rich sequence, no specific mRNA degradation accelerating or stabilizing elements can be distinguished.

In conclusion, we isolated an alternative human α E-catenin transcript which contains an additional 321 bp in the 3'UTR sequence. The extension probably arises due to alternative polyadenylation of the α E-catenin transcription unit. This additional sequence might play a role in the mRNA stability, however, further investigations will be necessary to confirm this hypothesis.

ACKNOWLEDGMENTS

We thank Dr. J. S. Rhim, National Cancer Institute, Bethesda, Maryland, for providing us with the cell line 267B1.

REFERENCES

1. Takeichi, M. (1990) *Annu. Rev. Biochem.* **59**, 237–252.
2. Nagafuchi, A., Shirayoshi, Y., Okazaki, K., Yasuda, K., and Takeichi, M. (1987) *Nature* **329**, 341–343.
3. Ringwald, M., Schuh, R., Vestweber, D., Eistetter, H., Lottspeich, F., Engel, J., Dölz, R., Jähnig, F., Epplen, J., Mayer, S., Muller, C., and Kemler, R. (1987) *EMBO J.* **6**, 3647–3653.
4. Takeichi, M. (1991) *Science* **251**, 1451–1455.
5. Ozawa, M., Baribault, H., and Kemler, R. (1989) *EMBO J.* **8**, 1711–1717.
6. McCrea, P. D., Turck, C. W., and Gumbiner, B. (1991) *Science* **254**, 1359–1361.
7. Hinck, L., Nathke, I. S., Papkoff, J., and Nelson, W. J. (1994) *J. Cell. Biol.* **125**, 1327–1340.
8. Sacco, P. A., McGranahan, T. M., Wheelock, M. J., and Johnson, K. R. (1995) *J. Biol. Chem.* **270**, 20201–20206.
9. Jou, T., Stewart, D. B., Stappert, J., Nelson, W. J., and Marrs, J. A. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5067–5071.
10. Rimm, D. L., Koslov, E. R., Kebriaei, P., Cianci, C. D., and Morrow, J. S. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 8813–8817.
11. Johnson, R. P., and Craig, S. W. (1995) *Nature* **373**, 261–264.
12. Menkel, A. R., Kroemker, M., Bubeck, P., Ronsiek, M., Nikolai, G., and Jockusch, B. M. (1994) *J. Cell. Biol.* **126**, 1231–1240.
13. Herrenknecht, K., Ozawa, M., Eckerskorn, C., Lottspeich, F., Lenter, M., and Kemler, R. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9156–9160.
14. Nagafuchi, A., Takeichi, M., and Tsukita, S. (1991) *Cell* **65**, 849–857.
15. Hirano, S., Kimoto, N., Shimoyama, Y., Hirohashi, S., and Takeichi, M. (1992) *Cell* **70**, 293–301.
16. Claverie, J. M., Hardelin, J. P., Legouis, R., Levilliers, J., Bougueleret, L., Mattei, M. G., and Petit, C. (1993) *Genomics* **15**, 13–20.
17. Oda, T., Kanai, Y., Shimoyama, Y., Nagafuchi, A., Tsukita, S., and Hirohashi, S. (1993) *Biochem. Biophys. Res. Commun.* **193**, 897–904.
18. Oda, H., Uemura, T., Shiomi, K., Nagafuchi, A., Tsukita, S., and Takeichi, M. (1993) *J. Cell. Biol.* **121**, 1133–1140.
19. Furukawa, Y., Nakatsuru, S., Nagafuchi, A., Tsukita, S., Muto, T., Nakamura, Y., and Horii, A. (1994) *Cytogen. Cell Genet.* **65**, 74–78.
20. Nollet, F., van Hengel, J., Berx, G., Molemans, F., and van Roy, F. (1995) *Genomics* **26**, 410–413.
21. Vermeulen, S. J., Bruyneel, E. A., Bracke, M. E., De Bruyne, G. K., Vennekens, K. M., Vleminckx, K. L., Berx, G. J., van Roy, F. M., and Mareel, M. M. (1995) *Cancer Res.* **55**, 4722–4728.
22. Morton, R. A., Ewing, C. M., Nagafuchi, A., Tsukita, S., and Isaacs, W. B. (1993) *Cancer Res.* **53**, 3585–3590.
23. Shimoyama, Y., Nagafuchi, A., Fujita, S., Gotoh, M., Takeichi, M., Tsukita, S., and Hirohashi, S. (1992) *Cancer Res.* **52**, 5770–5774.
24. Hirano, S., Kimoto, N., Shimoyama, Y., Hirohashi, S., and Takeichi, M. (1992) *Cell* **70**, 293–301.
25. Ewing, C. M., Ru, N., Morton, R. A., Robinson, J. C., Wheelock, M. J., Johnson, K. R., Barrett, J. C., and Isaacs, W. B. (1995) *Cancer Res.* **55**, 4813–4817.
26. Parda, D. S., Thraves, P. J., Kuettel, M. R., Lee, M., Arnstein, P., Kaighn, M. E., Rhim, J. S., and Dritschilo, A. (1993) *Prostate* **23**(2), 91–98.
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
28. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
29. Devereux, J., Haeblerli, P., and Smithies, O. (1984) *Nucl. Acids Res.* **12**, 387–395.
30. Bonfield, J. K., Smith, K. F., and Staden, R. (1995) *Nucleic Acids Res.* **23**(24), 4992–4999.
31. Church, G. M., and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1991–1995.
32. Feinberg, A. P., Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
33. Rimm, D. L., Kebriaei, P., and Morrow, J. S. (1994) *Biochem. Biophys. Res. Commun.* **203**, 1691–1699.
34. Zubiaga, A. M., Belasco, J. G., and Greenberg, M. E. (1995) *Mol. Cell. Biol.* **15**(4), 2219–2230.