

# Attenuated *APC* alleles produce functional protein from internal translation initiation

Kathleen Heppner Goss\*, Chris Trzepacz\*†, Thérèse M. F. Tuohy‡, and Joanna Groden§

Howard Hughes Medical Institute, Department of Molecular Genetics, Biochemistry, and Microbiology, University of Cincinnati College of Medicine, Cincinnati, OH 45267

Edited by Albert de la Chapelle, Ohio State University, Columbus, OH, and approved April 10, 2002 (received for review February 6, 2002)

**Some truncating mutations of the *APC* tumor suppressor gene are associated with an attenuated phenotype of familial adenomatous polyposis coli (AAPC). This work demonstrates that *APC* alleles with 5' mutations produce *APC* protein that down-regulates  $\beta$ -catenin, inhibits  $\beta$ -catenin/T cell factor-mediated transactivation, and induces cell-cycle arrest. Transfection studies demonstrate that cap-independent translation is initiated internally at an AUG at codon 184 of *APC*. Furthermore, *APC* coding sequence between *AAPC* mutations and AUG 184 permits internal ribosome entry in a bicistronic vector. These data suggest that *AAPC* alleles *in vivo* may produce functional *APC* by internal initiation and establish a functional correlation between 5' *APC* mutations and their associated clinical phenotype.**

Germline mutation of the *APC* gene is associated with familial adenomatous polyposis coli (APC), an inherited disease in which the colonic mucosa is carpeted with adenomatous polyps early in adulthood. A subset of persons with *APC* mutations displays a less severe form of the disease, attenuated adenomatous polyposis coli (AAPC); they carry mutations in exons 3 and 4 at the 5' end of the gene (1–6), the extreme 3' end of the gene (4, 7–9), or at splice junctions (10, 11). Genetic and clinical analyses of AAPC and APC families reveal a phenotypic boundary in the *APC* gene between the most distal “attenuated” mutation at codon 157 and most proximal “classical” mutation at codon 168 (3, 12).

The *APC* tumor suppressor gene follows the paradigm of retinoblastoma in that the rate-limiting step in tumor formation is somatic mutation of the normal allele in adenomas from persons with germline mutation. Adenomas from individuals with AAPC are characterized by somatic mutation of the normal *APC* allele, but they also can carry somatic mutations in the *AAPC* allele (13, 14), suggesting that the *AAPC* allele may be functional and that its inactivation is rate-limiting for tumorigenesis. Furthermore, these data imply that three hits may be required for tumor development in persons with AAPC, which may in turn be reflected in the later age of onset and reduced tumor burden of AAPC. The mechanism by which *AAPC* alleles maintain function has not yet been addressed.

We tested the hypothesis that 5' *AAPC* alleles produce functional protein by initiation of translation downstream of truncating mutations in *APC* associated with attenuated polyposis. Our data demonstrate that an *AAPC* allele with a chain-terminating mutation at codon 157 of *APC* produces protein with normal subcellular localization that maintains its ability to down-regulate both  $\beta$ -catenin and Wnt signaling and induce a cell-cycle arrest. To address the mechanism by which *AAPC* mutation produces functional protein, short 5' *APC* fragments containing *AAPC* mutations were placed upstream of green fluorescent protein (GFP)-encoding sequences and transfected into COS-1 cells. Western analyses demonstrate that translation is initiated internally at an AUG at codon 184 of *APC*. Analysis of the region between codons 157 and 184 of *APC* suggests that *APC* sequence and RNA secondary structure facilitate internal ribosome entry to initiate translation. These data suggest that 5' *AAPC* mutations not only may be attenuated because a nearly

full-length protein is translated, but also that the amino terminus of *APC*, including a homo-dimerization domain and nuclear export signal, may not be critical to *APC* tumor-suppressor function.

## Methods

**Plasmid Constructs.** The *APC-GFP* expression vector was generated by digesting full-length *APC* cDNA in *pBS(SK)APC* (15) with *Bam*HI and subcloning it into the *Bam*HI site of *pEGFP-N3* (CLONTECH). Digestion of *APC-GFP* with *Kpn*I liberates a fragment encoding the first 621 nucleotides of *APC*; subcloning this fragment into *pEGFP-N3* generated *APCKpn-GFP*. Mutations within the *APC* coding sequence were generated by site-directed mutagenesis of *APCKpn-GFP* with the Quick-Change mutagenesis kit (Stratagene). *AAPCKpn-GFP* has a TGG  $\rightarrow$  TAG change at *APC* codon 157. Constructs *AAPCKpn-GFP/AUG184* and *AAPCKpn-GFP/AUG200* have ATG  $\rightarrow$  GCG and ATG  $\rightarrow$  GCA mutations, respectively, resulting in Met to Ala substitutions at codons 184 and 200; *AAPCKpn-GFP/AUG184,200* includes both mutations. *AAPCKpn-GFP/AUGGFP* contains a mutation of the initiating ATG in GFP to GCG, creating a Met to Ala substitution. *168UAG-GFP* has an AGA  $\rightarrow$  TAG mutation at codon 168.  $\Delta$ 170-*GFP* contains a one base deletion (A) in codon 170 to generate a frameshift and downstream stop codon at position 174. The *AAPCKpn-GFP* construct was subcloned by *Kpn*I digestion into the *APC-GFP* expression vector to create *AAPC-GFP*. *E1A-GFP*, *E1A-APCKpn-GFP*, and *E1A-AAPCKpn-GFP* were generated by subcloning the *Eco*RI/*Xba*I fragments of *pEGFP*, *APCKpn-GFP*, and *AAPCKpn-GFP*, respectively, into *pCDEF3* downstream of the *E1A* promoter. The bicistronic construct, *pCAT-IRES-GFP*, was created by cloning chloramphenicol acetyl transferase (CAT) into the *Pst*I-*Bam*HI restriction sites in *pIRES2GFP* (CLONTECH) by PCR. The encephalomyocarditis virus (ECMV) internal ribosome entry site (IRES) was removed from this vector with *Bam*HI and *Bst*XI to generate *pCAT-GFP*. The *APC* segments (nucleotides 474–565 and nucleotides 516–565) were ligated as oligonucleotide linkers into the *Bam*HI/*Bst*XI sites to create *pCAT-APCIRES-GFP* and *pCAT-3'APCIRES-GFP*, respectively. All constructs were verified by automated nucleotide sequencing at the University of Cincinnati DNA Core Laboratory.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: APC, adenomatous polyposis coli; AAPC, attenuated adenomatous polyposis coli; GFP, green fluorescent protein; CAT, chloramphenicol acetyl transferase; ECMV, encephalomyocarditis virus; IRES, internal ribosome entry site; TCF, T cell factor; CMV, cytomegalovirus.

\*K.H.G. and C.T. contributed equally to this work.

†Present address: Program in Molecular Medicine, University of Massachusetts Medical Center, Worcester, MA 01605.

‡Present address: Department of Cell Biology, University of Cincinnati College of Medicine, 231 Albert Sabin Way, Cincinnati, OH 45267-0521.

§To whom reprint requests should be addressed at: Howard Hughes Medical Institute, University of Cincinnati College of Medicine, 231 Albert Sabin Way, Cincinnati, OH 45267-0524. E-mail: joanna.groden@uc.edu.

**Cell Culture and Transfection.** COS-1 (African green monkey) and SW480 (human colorectal carcinoma) cells (American Type Culture Collection) were maintained in DMEM with 10% (vol/vol) FBS (HyClone). Cells ( $10^5$ ) were plated on coverslips in 35-mm plates and transfected with 1.0  $\mu\text{g}$  of each plasmid, unless noted, by using Fugene-6 transfection reagent (Roche Molecular Biochemicals). For rapamycin experiments, cells were transfected in serum-free conditions with Lipofectamine (GIBCO/BRL) and washed after 45 min. They were placed in media with serum and 100 ng/ml rapamycin (Sigma) or vehicle (ethanol) for 24 h before harvesting.

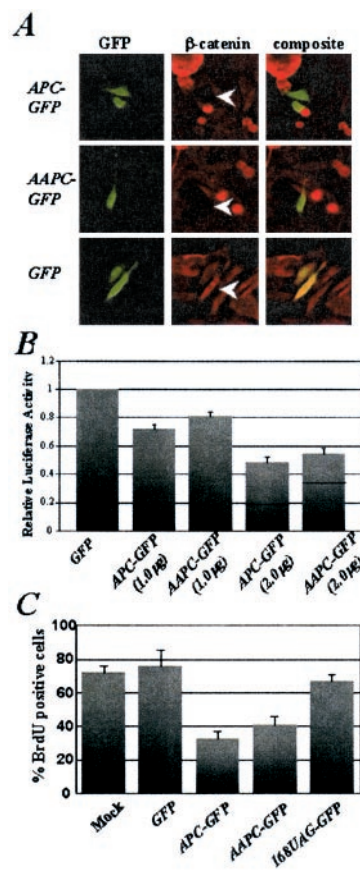
**T Cell Factor (TCF) Reporter Assays.** Cotransfections were performed in SW480 cells with *pEGFP* (0.5  $\mu\text{g}$ ), *APC-GFP* (1.0 or 2.0  $\mu\text{g}$ ), or *AAPC-GFP* (1.0 or 2.0  $\mu\text{g}$ ) and 0.5  $\mu\text{g}$  *TOPFLASH* (Upstate Biotechnology, Lake Placid, NY) and 0.5  $\mu\text{g}$  *pCMV $\beta$ Gal* (CLONTECH). pBluescript II KS(+) (Stratagene) was used to normalize the amount of DNA in each transfection. After transfection (24 h), cells were harvested and analyzed with the Luciferase Assay System (Promega). Luciferase activity was normalized to transfection efficiency by using the  $\beta$ -galactosidase Enzyme Assay System (Promega). Transfections were performed in triplicate, and the results of three independent experiments were averaged.

**Western Analyses.** Cells were lysed in Laemmli sample buffer 24–48 h after transfection. Samples were resolved by denaturing PAGE on 12% gels and transferred to Immobilon-P nylon membrane (Millipore). The samples were analyzed by Western blotting with a 1:2,000 dilution of an anti-GFP monoclonal antibody (Roche Molecular Biochemicals) or a 1:200 dilution of an anti-CAT-digoxigenin antibody (Roche Molecular Biochemicals). Blots were probed with a 1:5,000 dilution of a peroxidase-conjugated goat anti-mouse secondary antibody (Kirkegaard & Perry Laboratories) or a 1:2,500 dilution of peroxidase-labeled anti-digoxigenin secondary antibody (Roche Molecular Biochemicals) and visualized with chemiluminescence.

**Immunohistochemistry and BrdUrd Incorporation.** Transfected SW480 cells were labeled with 10  $\mu\text{M}$  BrdUrd (Amersham Pharmacia) in culture medium for 16 h, beginning 24 h after transfection. Cells were fixed in 3.7% (vol/vol) formaldehyde and permeabilized with 0.3% Triton X-100. Coverslips were probed with a 1:500 dilution of rat monoclonal anti-BrdUrd antibody (Accurate Scientific, Westbury, NY) or a 1:200 dilution of mouse monoclonal anti- $\beta$ -catenin antibody (Transduction Laboratories, Lexington, KY) with 20 units/ml RNase-free DNase I (Roche Molecular Biochemicals) in blocking buffer (5 mg/ml BSA, 0.5% Nonidet P-40 in PBS). Slides were probed with a 1:200 dilution of rhodamine-conjugated donkey anti-rat antibody (Jackson ImmunoResearch) or a 1:200 dilution of rhodamine-conjugated rat anti-mouse antibody (Kirkegaard & Perry Laboratories). Slides were mounted with Fluoromount G (Electron Microscopy Sciences, Fort Washington, PA) containing 10  $\mu\text{g}/\text{ml}$  Hoechst stain and evaluated with an Axioplan2 microscope (Zeiss); images were captured with a digital camera (Hamamatsu, Middlesex, NJ) and imaging software (QED Imaging) or with a 35-mm camera. More than 150 GFP<sup>+</sup> cells were scored blindly in each experiment; the results of at least three independent experiments were averaged.

## Results

**AAPC Alleles Produce Functional APC Protein.** The identification of two somatic *APC* mutations in adenomas from persons with AAPC indicates that *AAPC* alleles may need to be inactivated by mutation *in vivo* and first suggested that these alleles might encode protein with residual tumor suppressor activity (13, 14). To assess the ability of an *AAPC* allele to produce APC protein,



**Fig. 1.** AAPC protein down-regulates  $\beta$ -catenin, inhibits  $\beta$ -catenin/TCF-mediated transactivation, and induces a G<sub>1</sub>/S phase cell-cycle arrest as efficiently as APC. (A) *APC-GFP*, *AAPC-GFP*, and *GFP* were transfected into SW480 colon cancer cells. Cells were stained with an anti- $\beta$ -catenin antibody. The white arrows show GFP<sup>+</sup> cells in the center panels. (B) SW480 cells were cotransfected with *APC-GFP*, *AAPC-GFP*, or *GFP*, a TCF luciferase reporter vector (*TOPFLASH*), and a  $\beta$ -galactosidase expression vector (*pCMV $\beta$ -gal*). The relative luciferase activity is normalized to transfection efficiency ( $\beta$ -galactosidase activity) and shown relative to SW480 cells transfected with *GFP*. The transfections were performed in triplicate in three independent experiments. (C) BrdUrd incorporation was measured in SW480 cells after mock transfection or transfection with *GFP*, *APC-GFP*, *AAPC-GFP*, or *168UAG-GFP*. The percentage of GFP<sup>+</sup> cells that stained positively with an anti-BrdUrd antibody is shown. More than 150 GFP<sup>+</sup> cells were counted in each experiment. The experiments were performed in triplicate.

a characterized *AAPC* mutation (16) was introduced into a full-length *APC* cDNA cloned in-frame with GFP (*APC-GFP*). Site-directed mutagenesis replaced the normal codon 157 (TGG) with an in-frame stop codon (TAG) to generate *AAPC-GFP* under the control of the cytomegalovirus (CMV) promoter. Transfection of SW480 colorectal cancer cells resulted in detectable expression of *AAPC-GFP* in transfected cells, although transfected cells displayed less fluorescence than cells transfected with *APC-GFP* (Fig. 1A). The localization of *AAPC-GFP* was identical to *APC-GFP*: cytoplasmic, filamentous, and concentrated at the membrane extensions of migrating cells. This localization is identical to immunostaining of endogenous APC (17, 18), ectopic expression of other GFP-tagged APC fusion proteins (19), and costaining of transfected cells with an anti-APC antibody (not shown).

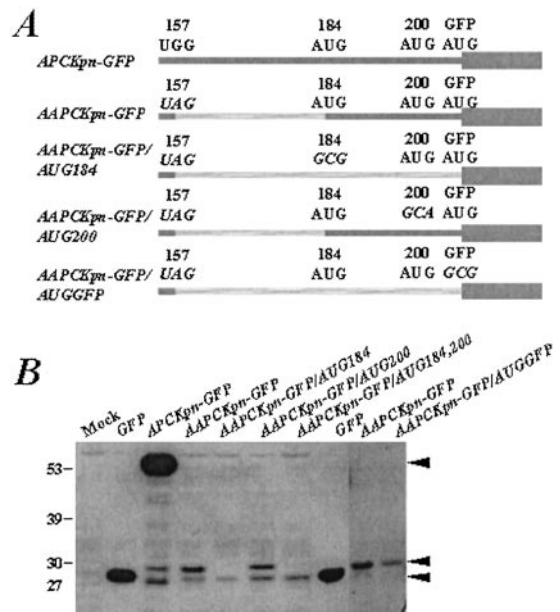
One important function of the APC tumor suppressor is to modulate Wnt signal transduction by regulating the cytoplasmic levels of  $\beta$ -catenin (20). Therefore, the ability of *AAPC-GFP* to down-regulate  $\beta$ -catenin and modulate  $\beta$ -catenin/TCF-

mediated transactivation was tested. SW480 cells carry an *APC* mutation and consequently accumulate both cytoplasmic and nuclear  $\beta$ -catenin that is readily detected by immunofluorescence with an anti- $\beta$ -catenin antibody. Transient transfection of SW480 cells with *APC-GFP* or *AAPC-GFP* efficiently reduced the levels of  $\beta$ -catenin (Fig. 1A). Furthermore, expression of *APC-GFP* or *AAPC-GFP* in SW480 cells, which have a high level of  $\beta$ -catenin/TCF-mediated transcriptional activity, resulted in a dose-dependent reduction of activity from a TCF reporter construct (Fig. 1B). These results demonstrate that *AAPC* facilitates  $\beta$ -catenin degradation and inhibits transactivation by  $\beta$ -catenin/TCF complexes.

Ectopic expression of *APC* in fibroblasts or colon cancer cells reproducibly inhibits entry into S phase of the cell cycle, an effect that is both  $\beta$ -catenin-dependent and -independent (ref. 21; C. D. Heinen, K.H.G., J. R. Cornelius, G. F. Babcock, E. S. Knudsen, T. Kowalik, and J.G., unpublished work). Transient transfection of *AAPC-GFP* into SW480 cells significantly decreased the percentage of cells incorporating BrdUrd after 16 h compared with cells transfected with *GFP* or mock-transfected cells (Fig. 1C). Although the suppression of S phase entry by *AAPC-GFP* was slightly less than that of *APC-GFP*, these results demonstrate that *AAPC* can function as a cell cycle regulator. Interestingly, when a stop codon replaced codon 168 in *APC* (*168UAG-GFP*) to recreate a classical *APC* mutation, the resulting protein product is not capable of inducing a  $G_1/S$  arrest (Fig. 1C), supporting a correlation between this mutation boundary and *APC* function.

**Internal Initiation of *AAPC* in Vivo.** One mechanism by which *AAPC* alleles maintain activity may be through translation initiation at AUG codons downstream of *AAPC* nonsense mutations. To test this hypothesis, a series of short *APC-GFP* and *AAPC-GFP* fusions were generated. A 5' *KpnI* fragment of *APC*, including nucleotides 1–664, was cloned in-frame and 5' of the *GFP* coding sequence to generate *APCKpn-GFP* (Fig. 2A). After transfection into COS-1 cells, Western analysis with an anti-GFP antibody demonstrated that *APCKpn-GFP* produced a 52-kDa protein (Fig. 2B). A stop-codon (UAG) was introduced by site-directed mutagenesis at codon 157 of *APCKpn-GFP* to recreate an *AAPC* mutation and generate *AAPCKpn-GFP* (Fig. 2A). *AAPCKpn-GFP* produced two protein products detectable by Western analysis: one that comigrated with GFP (28 kDa) and one that was smaller than that produced by *APCKpn-GFP* but slightly larger than GFP (approximately 31 kDa). Longer exposures of Western blots of lysates from cells transfected with *APCKpn-GFP* also showed both of these bands (Fig. 2B).

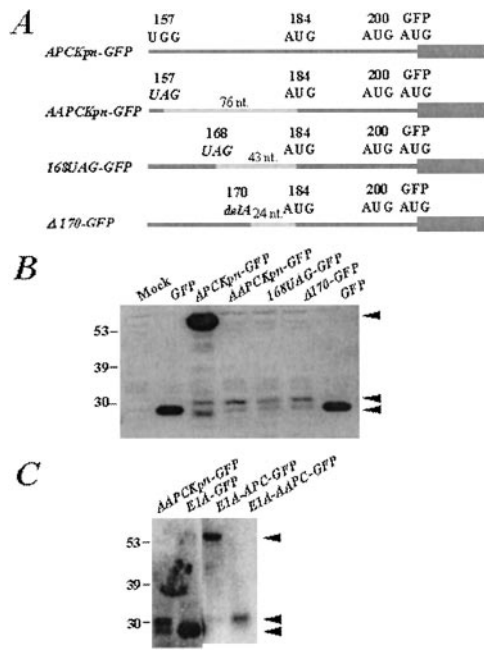
*AAPCKpn-GFP* contains three in-frame AUG initiation codons (at codons 184 and 200 in *APC* and one at the beginning of *GFP*) downstream of the *AAPC* stop codon at 157. To test whether translation initiation was occurring at any of these AUG codons to generate proteins detected with the anti-GFP antibody, derivatives of *AAPCKpn-GFP* were created in which each of these AUGs (*APC* codon 184, *APC* codon 200 and the *GFP* AUG) were mutated independently or in combination (Fig. 2A). Western analysis was performed with lysates from COS-1 cells transfected with each construct. The *AAPCKpn-GFP* band that comigrates with GFP is absent from cells transfected with *AAPCKpn-GFP/AUGGFP* in which the initiation codon for *GFP* was mutated, indicating that this protein was derived from internal translation of *GFP* (Fig. 2B). Mutation of the first in-frame downstream AUG in *AAPCKpn-GFP* at codon 184 resulted in the loss of the 31-kDa protein (Fig. 2B). In contrast, mutation of the AUG at codon 200 did not alter the proteins detectable by Western analysis with the GFP antibody. Ablation of all three AUGs downstream of the *AAPC* mutation abrogated expression of both proteins (Fig. 2B). These data suggest that the AUG at codon 184 in *AAPC*, 27 codons downstream of the



**Fig. 2.** Initiation of translation occurs at the first available AUG after the *AAPC* nonsense mutation. (A) Schematic of the *APCKpn-GFP* and *AAPCKpn-GFP* constructs used for transfection experiments. Dark gray bars represent predicted translated products, and hatched gray bars represent untranslated regions of *APC*. Wide gray bar represents translation of *GFP* coding sequence. Mutated codons in each construct are italicized. (B) Western analysis of lysates from COS-1 cells transfected with one of the constructs is shown in A. The blot was probed with a mouse monoclonal anti-GFP antibody. Molecular weight protein standards are represented on the far left (kDa). Mock and *GFP*-transfected cells are shown as controls. Black arrows on the right show proteins generated from translation initiation at codons 1 and 184 in *APC* and at the beginning of *GFP*.

engineered stop codon at 157, facilitates translation initiation, whereas codon 200 cannot. The relative intensity of the proteins produced by *APCKpn-GFP* and *AAPCKpn-GFP* suggests that internal initiation at codon 184 is relatively inefficient compared with initiation at the conventional *APC* AUG (codon 1; Fig. 2B). This suggestion is also supported by the relatively low intensity of fluorescence observed in cells transfected with the full-length *AAPC-GFP* construct when compared with cells transfected with *APC-GFP*.

Although initiation of translation usually occurs at an AUG triplet proximal to the 5' end of an mRNA, it is possible for leaky scanning to bypass this first AUG and initiate translation downstream. This possibility is an unlikely mechanism of initiation at codon 184 in *APC*, because the AUG at codon 1 of *APC* meets the Kozak sequence requirements for efficient initiation (22). Alternatively, it is possible for ribosomes to reinitiate translation of an ORF after termination. Reinitiation conventionally requires a minimum of 50 bps between the stop codon and AUG (22, 23). To address whether initiation of *AAPC* alleles could be considered *bona fide* reinitiation, the interval between the stop codon and initiation codon in *APC* was altered. A derivative of *APCKpn-GFP*, *168UAG-GFP*, was generated by site-directed mutagenesis in which a stop codon was introduced at codon 168 to decrease the interval between the stop codon and the in-frame AUG 184 (76 nucleotides to 43 nucleotides; Fig. 3A). Proteins of approximately 28 and 31 kDa were detected by Western analysis with an anti-GFP antibody, although these bands are less intense than those produced by *AAPCKpn-GFP* (Fig. 3B). Another derivative of *APCKpn-GFP*,  $\Delta$ 170-*GFP*, was generated in which a deletion of one nucleotide (an A at position 509) produced a downstream stop codon at 174 to decrease the

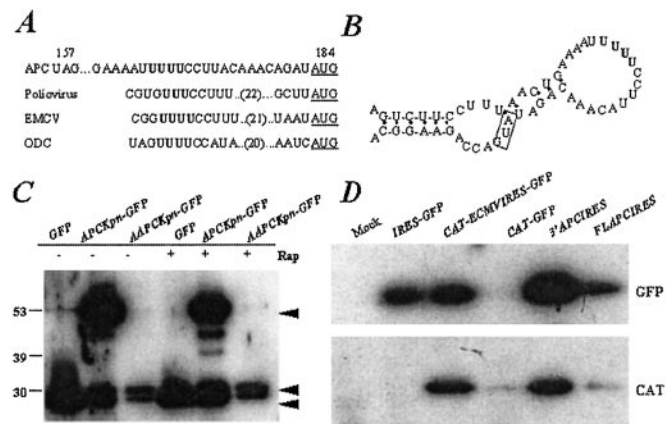


**Fig. 3.** Initiation of translation at codon 184 in *APC* represents internal initiation rather than reinitiation. (A) Schematic of the *AAPCKpn-GFP* constructs used for transfection experiments in which the spacing between the stop codon and the AUG at codon 184 is reduced from 76 nucleotides (in *AAPCKpn-GFP*) to 43 nucleotides by a stop codon at 168 (in *168UAG-GFP*) and to 24 nucleotides by a one-base deletion at codon 170 (in *Δ170-GFP*). Dark gray bars represent predicted translated products; hatched gray bars represent untranslated regions of *APC*. Wide gray bar represents translation of *GFP* coding region. (B) Western analysis of lysates from COS-1 cells transfected with the constructs shown in A. The blot was probed with a mouse monoclonal anti-GFP antibody; molecular weight protein standards are represented on the far left (kDa). Mock and *GFP*-transfected cells are shown as controls. Black arrows on the right show proteins generated from translation initiation at codons 1 and 184 in *APC* and at the beginning of *GFP*. (C) Western analysis of cells transfected with *E1A-GFP*, *E1A-APCKpn-GFP*, and *E1A-AAPCKpn-GFP* probed with the anti-GFP antibody. *APCKpn-GFP* and *AAPCKpn-GFP*-transfected cells are shown as controls. Black arrows show proteins generated from translation initiation at codons 1 and 184 in *APC* and at the beginning of *GFP*. Molecular weight standards are represented on the far left (kDa).

interval between this stop and AUG 184 (24 nucleotides; Fig. 3A). This construct produced similar proteins as *AAPCKpn-GFP* (Fig. 3B), suggesting that the mechanism of initiation at AUG 184 of *APC* is not reinitiation. In addition, read-through of the *APC* stop codon at 157 was ruled out by inserting an additional in-frame stop codon into *AAPCKpn-GFP* at codon 159 without any effect on translation initiation at AUG 184 (not shown).

Finally, to rule out the possibility that internal initiation of *AAPCKpn-GFP* was an artifact of the strong *CMV* promoter, it was replaced by the *E1A* promoter to create *E1A-AAPCKpn-GFP*. Western analysis showed that the 31-kDa band is evident in lysates from cells transfected with *E1A-AAPCKpn-GFP*, as in *AAPCKpn-GFP*-transfected cell lysates (Fig. 3C). These results suggest that initiation at codon 184 of *APC* is independent of the promoter driving expression of these constructs in mammalian cells.

**Internal Initiation Is Mediated by an IRES in the *APC* Coding Sequence.** For some mRNAs with long, highly structured 5' untranslated regions, other mechanisms may mediate translation initiation. Internal initiation is a cap-independent translation process that uses an IRES within the mRNA to recruit the small ribosomal subunit, thereby bypassing the requirement for ribosome scanning. IRESs were first described in the 5' untranslated region of



**Fig. 4.** *APC* contains an IRES within its coding sequence. (A) Sequence alignment of *APC* RNA sequence between the *AAPCKpn-GFP* stop codon at 157 and the initiation codon at 184 with characterized IRES elements. The AUG initiation codon is underlined, and a pyrimidine-rich tract that is required for IRES activity in characterized IRESs is shown in bold. (B) Predicted RNA secondary structure of the putative *APC* IRES as modeled by M-FOLD RNA-folding software. A characteristic stem-loop structure is juxtaposed to the initiating AUG at codon 184 (box). Black circles show base pairing; the pyrimidine-rich tract is shown in bold. (C) Western analysis of cells transfected with *GFP*, *APCKpn-GFP*, or *AAPCKpn-GFP* and treated with 100 ng/ml of rapamycin or vehicle for 24 h. The blot was probed with a mouse monoclonal GFP antibody; molecular weight standards are represented on the far left (kDa). Black arrowheads show proteins generated from translation initiation at codons 1 and 184 in *APC* and at the beginning of *GFP*. Three independent experiments were performed; one representative experiment is shown. (D) Western analyses of cells transfected with a bicentric vector in which *CAT* is downstream of the *CMV* promoter and *GFP* is downstream of the *ECMV* IRES (in *CAT-IRES-GFP*), the putative *APC* IRES (*FLAPCIREs* and 3'*APCIREs*), or nothing (*CAT-GFP*). *IRES-GFP* represents cells transfected with the parental vector in which *GFP* is downstream of the *ECMV* IRES but does not contain *CAT*. The top blot is probed with anti-GFP antibody, and the bottom blot is probed with anti-CAT antibody.

RNAs from picornaviruses, plus-strand RNA viruses including *ECMV*, hepatitis A virus, and poliovirus, all of which lack a 5' cap (24–26). Several cellular mRNAs have been identified that contain an IRES in the 5' untranslated region, such as *ODC*, *FGF-2*, *VEGF*, and the *c-myc* proto-oncogene (27). Such cellular RNAs suggest that IRESs can mediate expression of eukaryotic mRNAs during apoptosis, mitosis, and other stress conditions in which cap-dependent translation is not used. Furthermore, the identification of an IRES in the coding region of the cyclin-dependent kinase *PITSLRE* suggests that these sequences can provide a tightly regulated mode of protein isoform expression (28).

A conserved core IRES structural element has been identified in several cellular and viral IRES elements (29). Sequence alignments of characterized IRESs with the *APC* mRNA sequence between codons 157 and 184 shows significant similarity of a polypyrimidine tract 15 nucleotides upstream of AUG 184 (Fig. 4A). A similar pyrimidine tract in *ODC*, for example, is required for IRES activity, as its function is abrogated when mutated to a polyA tract (30). In addition, this tract and the AUG at codon 184 are perfectly conserved in the *APC* gene from human, mouse, rat, and *Xenopus* (not shown).

In general, IRES-containing mRNAs have complex 5' untranslated regions that form extensive RNA secondary structures, including stem-loops, that may be repressive for cap-dependent ribosome scanning (31). The Zuker RNA folding software M-FOLD was used to model the *APC* mRNA structure and predict that the pyrimidine-rich tract of the putative IRES in *APC* could form a stem-loop structure juxtaposed to the AUG

at codon 184 (Fig. 4B). This predicted structure is consistent with the hypothesis that this region of *APC* recruits the ribosome for internal initiation via a characteristic secondary structure.

Internal ribosome entry mediates cap-independent translation. To verify that translation from AUG 184 is cap-independent, COS-1 cells transfected with *GFP*, *APCKpn-GFP*, or *AAPCKpn-GFP* were treated with rapamycin, a specific inhibitor of cap-dependent translation initiation (32), for 24 h after transfection and analyzed by Western analysis. Treatment of cells with rapamycin inhibited expression of proteins at 52 kDa (*APCKpn-GFP*) and 28 kDa (*GFP*) at least two-fold, but had no effect on the expression of the 31-kDa protein (*AAPCKpn-GFP*) in transfected cells (Fig. 4C). In addition, translation from the *GFP* AUG to produce the 28-kDa protein in cells transfected with *APCKpn-GFP* and *AAPCKpn-GFP* was similarly unchanged by rapamycin treatment, suggesting that translation from both AUGs in these constructs is cap-independent.

To test directly whether the *APC* sequence between codons 157 and 184 (nucleotides 441 and 551) is sufficient to mediate cap-independent translation, this fragment of *APC* was cloned between two reporter genes in a bicistronic vector. In this construct, *CAT* expression is driven by the CMV promoter in a cap-dependent fashion, whereas *GFP* expression measures cap-independent translation mediated by the putative *APC* IRES. The construct was transfected into COS-1 cells and the expression of each reporter gene expression was evaluated by Western analysis. The *APC* segment, as well as the 3' portion containing the polypyrimidine stretch, was sufficient to express *GFP* in transfected cells (Fig. 4D). As a positive control for IRES activity, the IRES from the picornavirus ECMV was used (*CAT-ECMVIRES-GFP*); a negative control had no intervening sequence between *CAT* and *GFP* (*CAT-GFP*). These results indicate that the region within the *APC* coding sequence, between the *AAPC* mutation boundary and the first downstream AUG, is capable of mediating internal ribosome entry and translation initiation and provide a molecular mechanism to explain the genotype-phenotype correlation in persons affected with attenuated polyposis.

## Discussion

These experiments test the hypothesis that *AAPC* patients with 5' proximal mutations are protected from classical polyposis by internal translation initiation of *APC*. It was demonstrated that a nonsense *AAPC* mutation at codon 157 introduced into full-length *APC* cDNA does not alter its ability to degrade  $\beta$ -catenin, inhibit  $\beta$ -catenin/TCF-mediated transactivation, or induce a G<sub>1</sub>-phase arrest. Western analyses of short 5' *APC* fragments containing *AAPC* mutations indicated that translation initiates at an internal AUG at codon 184 of *APC*. Analysis of the *APC* sequence between codons 157 and 184 suggested that RNA secondary structure facilitates internal ribosome entry to initiate translation at this codon via a *bona fide* IRES. Our data not only suggest that 5' *AAPC* mutations may be less severe because a nearly full-length protein is translated, but they also indicate that residues amino-terminal to codon 184 are not critical to *APC* function. These results provide a molecular explanation for the phenotype-genotype correlation in persons with *AAPC*.

Most classical polyposis and somatic *APC* mutations are localized to the central third of the coding sequence (33), a region designated the "mutation cluster region" which encodes the domain of the protein necessary and sufficient for  $\beta$ -catenin down-regulation (34). Localization of *AAPC* mutations to the 5' and 3' ends of the gene, as well as to splice junctions, has been difficult to reconcile with the known protein functions of *APC* and the phenotype of patients with classical *APC* mutations. Analyses of *APC* alleles with splice site mutations (10) and the presence of somatic mutations in both the normal allele of *APC*

and *AAPC* alleles in adenomas (13, 14) suggest that *AAPC* alleles produce functional tumor suppressor protein. Internal initiation of translation downstream of the 5' *AAPC* nonsense mutation provides one explanation for this finding.

In our model, *AAPC* alleles produce nearly full-length protein that is functional by the assays we have used. It is likely that internal initiation is relatively inefficient, so that *AAPC* with 5' mutations of *APC* could be considered a haploinsufficient phenotype. This hypothesis is supported by the fluorescence and Western analyses presented here. *AAPC* alleles with 3' mutations may be more easily reconciled phenotypically, in that they may produce proteins retaining at least the  $\beta$ -catenin down-regulation domain but lacking the carboxyl-terminal functional domain, such as that for discs large binding, required for full *APC* tumor suppressor activity. Classical *APC* mutations and *AAPC* alleles with 3' mutations produce proteins fully capable of dimerization with full-length *APC* (35), an activity predominantly conferred by the heptad repeat sequences encoded by exon 1 of *APC* (35, 36). Although the identification of *APC* patients with cytogenetic deletions of *APC* argues strongly against a dominant-negative activity of mutant *APC* allele (37–39), recent *in vivo* and biochemical experiments suggest a dominant-negative function of some mutant *APC* alleles (40). Nevertheless, it is unlikely that 5' *AAPC* alleles would produce products that function as dominant-negative proteins, given that the internally initiated protein lacks the first heptad repeat.

If *AAPC* protein lacks the first 183 aa and retains tumor suppressor activity, is this region of the protein dispensable? Exon 1 encodes most of the *APC* homo-dimerization domain (35, 36); however, no mutations have been detected in this exon, and some alternatively spliced isoforms lack this exon entirely (41, 42), suggesting that it is not necessary for *APC* activity. A nuclear export sequence (NES) at the amino terminus has been identified immediately upstream of the *AAPC* stop codon 157 and is sufficient to confer nuclear export of a reporter protein (43, 44). However, no evidence indicates that mutation of the NES affects localization or function of full-length *APC* (K.H.G., C. D. Heinen, T.M.F.T., and J.G., unpublished work). Further experiments are required to understand the role, if any, of the amino terminus of *APC* in tumor suppression.

Our data demonstrate that internal initiation of translation downstream of *AAPC* nonsense mutations is mediated by an IRES in the *APC* sequence encoded by a portion of exons 4 and 5. This sequence is highly conserved in *APC*, consistent with functional significance. Although the function of IRESs in cellular RNAs is not entirely clear, they may be important in regulating protein expression at times when cap-dependent translation is compromised, such as during apoptosis, mitosis, or in response to hypoxia, serum deprivation and DNA damage (27). The only other gene known to contain an IRES within its coding sequence is the *PITSLRE* protein kinase, a cell cycle-regulated kinase related to the cyclin-dependent kinase CDK1 (28). The p58<sup>PITSLRE</sup> isoform is produced by internal initiation of translation mediated by its IRES during the G<sub>2</sub>/M phase of the cell cycle and is critical to controlling proliferation of some cell types (28). Although it is not known whether *APC* IRES usage is cell cycle-dependent, *APC* is a substrate of CDK1, localizes to the kinetochore during the early stages of mitosis, and is implicated in maintaining chromosomal stability and segregation (45–47). These data suggest that the *APC* IRES may be important during mitosis.

There are several examples of apoptotic or survival factors whose translation is regulated by an IRES, including c-myc, DAP5, XIAP, and Apaf-1 (27). It is intriguing to speculate that cap-independent translation of *APC* may be differentially regulated during programmed cell death or in response to death signals. In colon cancer cells, introduction of *APC* increases apoptosis (48), and, *in vitro*, exogenous *APC* accelerates apo-

ptosis in a caspase-8-dependent manner (K. Steigerwald, G. K. Behbehani, K. A. Combs, M. C. Barton, and J.G., unpublished work). In addition, a caspase-3 cleavage product of APC has been identified in cells undergoing apoptosis (49). It is possible that IRES-mediated translation of APC is important for an apoptotic cascade, and that dysregulation of this function of APC may swing the balance in cancer cells from death to growth.

These data illustrate the need to understand more fully the regulation of APC expression and how this regulation may be critical to its function as a tumor suppressor. Furthermore, these data highlight the utility of dissecting the genotype-phenotype

correlations in hereditary diseases, because such an understanding may contribute to the design of treatment strategies for polyposis and colorectal cancer.

We thank J. Kordich, L. Slovek, A. Combs, and K. Marshall for technical assistance and K. Lillard-Wetherell for insightful discussions and advice. This work was supported by the Ohio Cancer Research Foundation, the Center for Environmental Genetics at the University of Cincinnati (ES06096), and National Institutes of Health awards CA63507 (to J.G.), CA71155 (to T.M.F.T.), and CA88460 (to K.H.G.). J.G. is an Assistant Investigator with the Howard Hughes Medical Institute.

- Leppert, M., Burt, R., Hughes, J. P., Samowitz, W., Nakamura, Y., Woodward, S., Gardner, E., Lalouel, J. M. & White, R. (1990) *N. Engl. J. Med.* **322**, 904–908.
- Spirio, L., Otterud, B., Stauffer, D., Lynch, H., Lynch, P., Watson, P., Lanspa, S., Smyrk, T., Cavalieri, J., Howard, L., et al. (1992) *Am. J. Hum. Genet.* **51**, 92–100.
- Spirio, L., Olschwang, S., Groden, J., Robertson, M., Samowitz, W., Joslyn, G., Gelbert, L., Thliveris, A., Carlson, M. & Otterud, B. (1993) *Cell* **75**, 951–957.
- Walton, C., Kartheuser, A., Michils, G., Smaers, M., Lannoy, N., Ngounou, P., Mertens, G. & Verellen-Dumoulin, C. (1997) *Hum. Genet.* **100**, 601–605.
- Soravia, C., Berk, T., Madlensky, L., Mitri, A., Cheng, H., Gallinger, S., Cohen, Z. & Bapat, B. (1998) *Am. J. Hum. Genet.* **62**, 1290–1301.
- Fidalgo, P. O., Maia, L. C., Cravo, M. L., Albuquerque, C. M., Suspiro, A., Ramalho, E. & Nobre-Leitao, C. (1999) *Cancer Genet. Cytogenet.* **111**, 152–156.
- Friedl, W., Meuschel, S., Caspari, R., Lamberti, C., Krieger, S., Sengteller, M. & Propping, P. (1996) *Hum. Genet.* **97**, 579–584.
- van der Luijt, R. B., Meera Khan, P., Vasen, H. F., Breukel, C., Tops, C. M., Scott, R. J. & Fodde, R. (1996) *Hum. Genet.* **98**, 727–734.
- Kartheuser, A., Walton, C., West, S., Breukel, C., Detry, R., Gribomont, A. C., Hamzehloei, T., Hoang, P., Maiter, D., Pringot, J., et al. (1999) *J. Med. Genet.* **36**, 65–67.
- Varesco, L., Gismondi, V., Presciuttini, S., Groden, J., Spirio, L., Sala, P., Rossetti, C., De Benedetti, L., Bafico, A., Heouaine, A., et al. (1994) *Hum. Genet.* **93**, 281–286.
- van der Luijt, R. B., Vasen, H. F., Tops, C. M., Breukel, C., Fodde, R. & Meera Khan, P. (1995) *Hum. Genet.* **96**, 705–710.
- Scott, R. J., van der Luijt, R., Spycher, M., Mary, J. L., Muller, A., Hoppeler, T., Haner, M., Muller, H., Martinoli, S. & Brazzola, P. L. (1995) *Gut* **36**, 731–736.
- Spirio, L. N., Samowitz, W., Robertson, J., Robertson, M., Burt, R. W., Leppert, M. & White, R. (1989) *Nat. Genet.* **20**, 385–388.
- Su, L. K., Barnes, C. J., Yao, W., Qi, Y., Lynch, P. M. & Steinbach, J. (2000) *Am. J. Hum. Genet.* **67**, 582–590.
- Groden, J., Joslyn, G., Samowitz, W., Jones, D., Bhattacharyya, N., Spirio, L., Thliveris, A., Robertson, M., Egan, S., Meuth, M. & White, R. (1995) *Cancer Res.* **55**, 1531–1539.
- Olschwang, S., Laurent-Puig, P., Groden, J., White, R. & Thomas, G. (1993) *Am. J. Hum. Genet.* **52**, 273–279.
- Nathke, I. S., Adams, C. L., Polakis, P., Sellin, J. H. & Nelson, W. J. (1996) *J. Cell Biol.* **134**, 165–179.
- Neufeld, K. L. & White, R. L. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 3034–3039.
- Mimori-Kiyosue, Y., Shiina, N. & Tsukita, S. (2000) *J. Cell Biol.* **148**, 505–518.
- Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B. & Polakis, P. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3046–3050.
- Baeg, G. H., Matsumine, A., Kuroda, T., Bhattacharjee, R. N., Miyashiro, I., Toyoshima, K. & Akiyama, T. (1995) *EMBO J.* **14**, 5618–5625.
- Kozak, M. (1989) *Mol. Cell Biol.* **9**, 5134–5142.
- Kozak, M. (1987) *Mol. Cell Biol.* **7**, 3438–3445.
- Jang, S. K., Davies, M. V., Kaufman, R. J. & Wimmer, E. (1989) *J. Virol.* **63**, 1651–1660.
- Glass, M. J., Jia, X. Y. & Summers, D. F. (1993) *Virology* **193**, 842–852.
- Pelletier, J. & Sonenberg, N. (1989) *J. Virol.* **63**, 441–444.
- Hellen, C. U. T. & Sarnow, P. (2001) *Genes Dev.* **15**, 1593–1612.
- Cornelis, S., Bruynooghe, Y., Denecker, G., Van Huffel, S., Tinton, S. & Beyaert, R. (2000) *Mol. Cell* **5**, 597–605.
- Le, S. Y. & Maizel, J. V., Jr. (1997) *Nucleic Acids Res.* **25**, 362–369.
- Pyronnet, S., Pradayrol, L. & Sonenberg, N. (2000) *Mol. Cell* **5**, 607–616.
- Niepel, M. & Gallie, D. R. (1999) *J. Virol.* **73**, 9080–9088.
- Beretta, L., Gingras, A. C., Svitkin, Y. V., Hall, M. N. & Sonenberg, N. (1996) *EMBO J.* **15**, 658–664.
- Goss, K. H. & Groden, J. (2000) *J. Clin. Oncol.* **18**, 1967–1979.
- Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S. & Polakis, P. (1995) *J. Biol. Chem.* **270**, 5549–5555.
- Su, L. K., Johnson, K. A., Smith, K. J., Hill, D. E., Vogelstein, B. & Kinzler, K. W. (1993) *Cancer Res.* **53**, 2728–2731.
- Joslyn, G., Richardson, D. S., White, R. & Alber, T. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11109–11113.
- Herrera, L., Kakati, S., Gibas, L., Pietrzak, E. & Sandberg, A. A. (1986) *Am. J. Med. Genet.* **25**, 473–476.
- Solomon, E., Voss, R., Hall, V., Bodmer, W. F., Jass, J. R., Jeffreys, A. J., Lucibello, F. C., Patel, I. & Rider, S. H. (1987) *Nature (London)* **328**, 616–619.
- Lindgren, V., Bryke, C. R., Ozcelik, T., Yang-Feng, T. L. & Francke, U. (1992) *Am. J. Hum. Genet.* **50**, 988–997.
- Mahmoud, N. N., Boolbol, K. K., Bilinski, R. T., Martucci, C., Chadburn, A. & Bertagnolli, M. M. (1997) *Cancer Res.* **57**, 5045–5050.
- Santoro, I. M. & Groden, J. (1997) *Cancer Res.* **57**, 488–494.
- Pyles, R. B., Santoro, I. M., Groden, J. & Parysek, L. M. (1998) *Oncogene* **16**, 77–82.
- Henderson, B. R. (2000) *Nat. Cell Biol.* **2**, 653–660.
- Neufeld, K. L., Nix, D. A., Bogerd, H., Kang, Y., Beckerle, M. C., Cullen, B. R. & White, R. L. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 12085–12090.
- Trzepacz, C., Lowy, A. M., Kordich, J. J. & Groden, J. (1997) *J. Biol. Chem.* **272**, 21681–21684.
- Fodde, R., Kuipers, J., Rosenberg, C., Smits, R., Kielman, M., Gaspar, C., van Es, J. H., Breukel, C., Wiegant, J., Giles, R. H. & Clevers, H. (2001) *Nat. Cell Biol.* **3**, 433–438.
- Kaplan, K. B., Burds, A. A., Swedlow, J. R., Bekir, S. S., Sorger, P. K. & Nathke, I. S. (2001) *Nat. Cell Biol.* **3**, 429–432.
- Morin, P. J., Vogelstein, B. & Kinzler, K. W. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 7950–7954.
- Webb, S. J., Nicholson, D., Bubb, V. J. & Wylie, A. H. (1999) *FASEB J.* **13**, 339–346.