THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 38, Issue of September 22, pp. 29660–29671, 2000
Printed in U.S.A.

Novel Alternative Splicing and Nuclear Localization of Human *RGS12* Gene Products*

Received for publication, January 14, 2000, and in revised form, May 30, 2000
Published, JBC Papers in Press, June 26, 2000, DOI 10.1074/jbc.M000330200

Tapan K. Chatterjee and Rory A. Fisher‡

From the Department of Pharmacology, University of Iowa College of Medicine, Iowa City, Iowa 52242

RGS proteins are GTPase-activating proteins for certain Gα subunits, accelerating the shutoff mechanism of G protein signaling, and also may interact with receptors and effectors to modulate G protein signaling. Here, we report identification of 12 distinct transcripts of human RGS12 that arise by unusually complex splicing of the *RGS12* gene, which spans 70 kilobase pairs of genomic DNA and contains 16 exons. These transcripts arise by both *cis*- and *trans*-splicing mechanisms, are expressed in a tissue-specific manner, and encode proteins ranging in size from 356 to 1447 amino acids. Both 5′- and 3′-splicing of two primary RGS12 transcripts occur to generate RGS12 mRNAs encoding proteins with four distinct N-terminal domains, three distinct C-terminal domains, and a common internal region where the semiconserved RGS domain is located. Confocal microscopy and subcellular fractionation of COS-7 cells expressing RGS12 proteins with three different N termini (brain (B), peripheral (P), and *trans*-spliced (TS)) and a shared short (S) C-terminal domain demonstrated exclusive nuclear localization of these proteins and an influence of the N-terminal region on the pattern of intranuclear distribution. Both native RGS12TS-S in HEK-293T cells and ectopically expressed RGS12TS-S localized to discrete nuclear foci (dots), a characteristic of various tumor suppressor proteins. Subnuclear localization of RGS12TS-S into nuclear dots was cell cycle-dependent. Native RGS12TS-S associated with the metaphase chromosome during mitosis, and ectopically expressed RGS12TS-S induced formation of abnormally shaped and multiple nuclei in COS-7 cells. Expression of RGS12 proteins with long and intermediate C-terminal domains was not observed in COS-7 cells, suggesting that 3′-splicing of RGS12 transcripts may influence the expression or stability of the encoded proteins. These results document extraordinary structural complexity in the RGS12 family and the role of alternative splicing and cell cycle-dependent mechanisms in expression and subnuclear targeting of RGS12 proteins.

Heterotrimeric G proteins are components of many major signaling systems used by cells to transduce a variety of signals (neurotransmitters, hormones, light, and olfactory and taste signals) from specific cell-surface receptors to effector proteins (see Refs. 1 and 2 for review). These receptors activate G proteins by stimulating exchange of GTP for GDP on the α subunit (Gα) of the inactive G protein heterotrimer (Gαβγ) to promote dissociation into Gα-GTP and Gβγ subunits. Both of these G protein subunits function as signal-transducing molecules by regulating activities of various effector proteins including enzymes and ion channels. The intensity and duration of signaling by G proteins are highly regulated. The key element that controls the lifetime of active Gα and Gβγ is the intrinsic GTPase activity of the Gα subunits, producing Gα-GDP and its reassembly with Gβγ.

RGS (regulators of G protein signaling) proteins constitute a family of proteins originally defined by the presence of a semiconserved region of ~120 amino acids called the RGS domain (3). These proteins appear to function as negative regulators of G protein signaling in organisms ranging from yeast to man (4, 5). The discovery that RGS proteins, or their isolated RGS domains, function as GTPase-activating proteins for certain Gα subunits *in vitro* (6, 7) provided the first insight into how these proteins may exert regulatory influences on G protein signaling. The crystal structure of RGS4 bound to $G\alpha_{i1}$ demonstrated interaction of the RGS domain with the G protein switch regions and suggested that the mechanism of GTPase activation by RGS proteins may be due to a reduction in the free energy of the transition state (8). Coleman and Sprang (9) recently suggested that RGS proteins may both stabilize the transition state and release Gα subunits from an autoinhibited ground state to enhance their GTPase activity. Additional studies have raised the possibility that some RGS proteins may interact with effectors or receptors to attenuate G protein signaling. These studies showed that recombinant RGS proteins can block phosphoinositide-dependent phospholipase C activation by active G_q *in vitro* and produce receptor-selective attenuation of G_q signaling when added to permeabilized cells (10, 11). Although Koelle and Horvitz (3) first predicted the existence of at least 15 mammalian RGS family members, this number has grown to exceed that of the number of identified Gα subunits; yet the GTPase-activating protein activity of RGS proteins appears to be limited to proteins in the G_i and G_q family (4), raising an interesting dilemma regarding the physiological significance of such a large family of RGS proteins.

We undertook studies to clone members of the human RGS protein family to further our understanding of the structural diversity within this family. This knowledge is crucial to understanding the structural determinants that might be involved in regulating the specificity and function(s) of RGS proteins. Here we report the identification of 12 distinct transcripts of human RGS12 and define their molecular basis of origin. These transcripts arise by extraordinarily complex splicing of the *RGS12* gene, providing the first documentation of alternative splicing of an RGS protein gene. Transcripts

* This work was supported by National Institutes of Health Grants HL-41071 (to R. A. F.) and DK-25295 (to the University of Iowa Diabetes and Endocrinology Research Center). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF030109, AF030110, AF030111, and AF030112.

‡ To whom correspondence should be addressed. Tel.: 319-335-8330; Fax: 319-335-8930; E-mail: rory-fisher@uiowa.edu.

encoding the different N-terminal forms of RGS12 were expressed in a tissue-specific fashion, and the expression and intracellular pattern of distribution of RGS12 proteins were affected by splicing at the 5′- and 3′-ends of their corresponding mRNAs. Our results demonstrate the exclusive nuclear localization of three distinct N-terminal forms of RGS12 in COS-7 cells, with one form (RGS12TS-S) localizing to discrete foci (nuclear dots), a characteristic of various tumor suppressor proteins. We developed an antibody to RGS12TS proteins and identified the endogenous expression of RGS12TS-S in HEK-293T cells. Localization of native RGS12TS-S in nuclear dots was cell cycle-dependent and independent of changes in the level of RGS12TS-S protein. Native RGS12TS-S associated with the metaphase chromosome in mitotic cells and ectopically expressed RGS12TS-S induced formation of abnormally shaped and multiple nuclei in COS-7 cells.

EXPERIMENTAL PROCEDURES

Materials—5′-RACE[1]-Ready cDNA, Marathon-Ready cDNA, a Quick-Screen cDNA library panel, and the pEGFP vector were purchased from CLONTECH. pCR2.1 and pCR3.1 were from Invitrogen. Elongase was from Life Technologies, Inc. Antibodies to lamin A/C, c-Myc (9E10), and 14-3-3β were purchased from Santa Cruz Biotechnology, and an antibody to Na^+/K^+-ATPase was from Upstate Biotechnology, Inc. Cell culture medium and serum were provided by the University of Iowa Diabetes and Endocrinology Research Center. Oligonucleotide primers and other molecular biological reagents were obtained from the University of Iowa DNA Core Facility. Polyclonal anti-RGS12TS antibodies were generated with a synthetic peptide immunogen corresponding to residues 1–15 of RGS12TS by Biosynthesis Inc. (Lewisville, TX).

PCR Amplification of RGS12 cDNAs—Full-length cDNAs encoding various forms of RGS12 were amplified using a PCR-based strategy we described previously (12). We utilized a 489-bp expressed sequence tag identified as RGS12 (GenBank™/EBI Data Bank accession number T57943) to design two "RGS12" forward primers (5′-CTTCATGAGATTGAAGATCTGCAGCTGCTG and 5′-GCTAGCTGGGCCTGGCTGTCGATGTTGACCG) and two RGS12 reverse primers (5′-ACATGTTCAAGGAGCAGCAGCTGCAGATC and 5′-GATAGCTACACTCGCTTTCTGAAGTCCCC) for use in 5′- and 3′-RACE to amplify overlapping segments of RGS12 cDNAs essentially as we described previously (13). Semi-nested 5′-RACE was performed using 5′-RACE-Ready human brain cDNA (anchor sequence at 5′-cDNA end) as template with a forward primer to the 5′-anchor sequence and nested RGS12 reverse primers. The same template was used for semi-nested 3′-RACE using nested RGS12 forward primers and an oligo(dT) reverse primer. Marathon-Ready human lung and placenta cDNAs (adapter sequences on both cDNA ends) were also used as templates for 5′- and 3′-RACE using adapter-specific forward or reverse primers in combination with appropriate nested RGS12 forward or reverse primers. Resulting PCR products were cloned into pCR2.1, and sequence analysis of multiple clones revealed successful amplification of overlapping 5′- and 3′-cDNA fragments of RGS12 from brain, lung, and placenta cDNAs. Sequence analysis of these clones revealed the existence of 12 different splice variant forms of RGS12. Full-length cDNAs encoding these splice variant forms of RGS12 were amplified using forward and reverse primers encompassing the translational start and stop sites, respectively, specific for each splice variant form. Resulting full-length cDNAs were cloned into pCR3.1, and double-stranded sequencing was performed by automated fluorescent dideoxynucleotide sequencing by the University of Iowa DNA Core Facility.

PCR Analysis of Tissue Distribution of Transcripts Encoding RGS12 Splice Variants—Human tissue cDNAs from the Quick- Screen human cDNA library panel were used as templates in PCR for amplification of the three 5′-splice variant forms of RGS12 that we identified by 5′-RACE and cloning. PCR was performed using a forward primer specific for each 5′-splice variant form of RGS12 (p1, p2, or p3) and a reverse primer (p4) to regions common to all RGS12 splice variants. The resulting products were diluted 1:100 and subjected to PCR using the same forward primer and a nested reverse primer (p5), again to a region shared by all RGS12 splice forms. Amplified products were separated by agarose gel electrophoresis. The identity of the amplified products was confirmed by dideoxynucleotide sequencing (University of Iowa DNA Core Facility) following their cloning into pCR2.1.

Preparation of EGFP Constructs of RGS12—Various RGS12 protein cDNAs were PCR-amplified using gene-specific primers incorporating restriction sites to facilitate their cloning into the EGFP vector. First, amplified RGS12 protein cDNAs were cloned in the T/A cloning vector pCR2.1 (Invitrogen). Then, restriction enzyme digestion and agarose gel purification of the cloned RGS12 protein cDNAs were performed. RGS12 protein cDNAs were ligated to the EGFP vector in frame with its C-terminal EGFP sequence. Double-stranded sequencing of all cloned RGS12 protein cDNAs was performed by automated fluorescent dideoxynucleotide sequencing by the University of Iowa DNA Core Facility.

cDNA encoding the 310-amino acid C-terminal domain of long splice forms of RGS12 was generated by PCR. A Kozak consensus sequence and an ATG start codon were included for proper translation of the truncated protein. Double-stranded sequencing of the cDNA construct was performed by automated fluorescent dideoxynucleotide sequencing by the University of Iowa DNA Core Facility.

Cell Culture and Transfection—COS-7, HEK-293, and HEK-293T cells were grown in DMEM supplemented with 10% fetal bovine serum and gentamycin (50 μg/ml) (complete DMEM) in a 5% CO_2 humidified atmosphere at 37 °C.

COS-7 cells were transiently transfected with vectors containing various RGS12 protein cDNAs by electroporation using a Bio-Rad Gene-Pulser. Typically, COS-7 cells (10^7/ml) were transfected with 40 μg of plasmid DNA at settings of 0.22 kV and 950 microfarads. Cells were diluted in complete DMEM and plated on two-chambered slides (Nunc) at a density of ~10^6 cells/well. For use in immunoblotting, cells were plated at an equivalent density on 10-cm culture plates. Transfected cells were used in experiments 40 h following transfection.

HEK-293 cells stably transfected with the pVgRXR vector were purchased from Invitrogen. These cells were transfected with C-terminal GFP-tagged RGS12TS-S (RGS12TS-S-GFP) in the pIND vector by electroporation, and stable transfectants were isolated and propagated in complete DMEM supplemented with G418 (800 μg/ml) and Zeocin (200 μg/ml; Invitrogen). These cells, designated EcR293–12TS-S-GFP, express RGS12TS-S-GFP within 12 h following ponasterone A (5 μM) induction.

Immunofluorescence—Cells were rinsed three times with DPBS before fixation for immunofluorescence. For visualization of GFP-tagged RGS12 proteins in COS-7 cells, cells were fixed by treatment with 4% paraformaldehyde for 20 min at room temperature, followed by permeabilization with DPBS containing 0.1% Triton X-100 and 0.1% Nonidet P-40 for 10 min at room temperature. After permeabilization, cells were treated with DPBS containing 100 μg/ml RNase A (Roche Molecular Biochemicals) for 20 min at room temperature prior to staining with propidium iodide. Cells were stained with propidium iodide in DPBS for 20 min at room temperature, followed by three washes with DPBS. Cells were air-dried and then mounted using VectaShield mounting solution. For visualization of endogenous RGS12TS in HEK-293T cells, cells were fixed with 70% ethanol at −20 °C for 1 h, followed by treatment with 4 N HCl for 30 min at room temperature. Cells were then incubated with anti-RGS12TS antibody in phosphate-buffered saline containing 2% bovine serum albumin and 0.2% Tween 20, followed by washing and incubation with fluorescein isothiocyanate-conjugated goat anti-rabbit secondary antibody (Sigma). RNase A treatment and propidium iodide staining were performed as described for COS-7 cells.

Confocal microscopy was performed with a Bio-Rad MRC 1024 confocal microscope equipped with a krypton/argon laser at the University of Iowa Central Microscopy Research Facility. EGFP fluorescence was examined under a fluorescein isothiocyanate filter, and propidium iodide fluorescence was examined under a Texas Red filter using 60× oil lenses. Images were captured after Kalman averaging. Images shown are representative of a minimum of 1600 cells derived from four or more separate transfections.

Subcellular Fractionation and Immunoblotting—Subcellular fractionation of COS-7 cells expressing Myc-tagged RGS12 proteins and HEK-293T cells expressing native RGS12TS-S was performed as we described previously (14). Polyclonal anti-RGS12TS antibodies were used at 1:10,000 dilution. Immunoblotting for specific organelle markers using lamin A/C as a nuclear marker, Na^+/K^+-ATPase as a plasma membrane marker, and 14-3-3β as a cytosolic marker in COS-7 cells confirmed the identity of these fractions. Immunoblotting was performed essentially as we described previously (13, 14).

[1] The abbreviations used are: RACE, rapid amplification of cDNA ends; PCR, polymerase chain reaction; bp, base pair(s); GFP, green fluorescent protein; EGFP, enhanced GFP; DMEM, Dulbecco's modified Eagle's medium; DPBS, Dulbecco's phosphate-buffered saline.

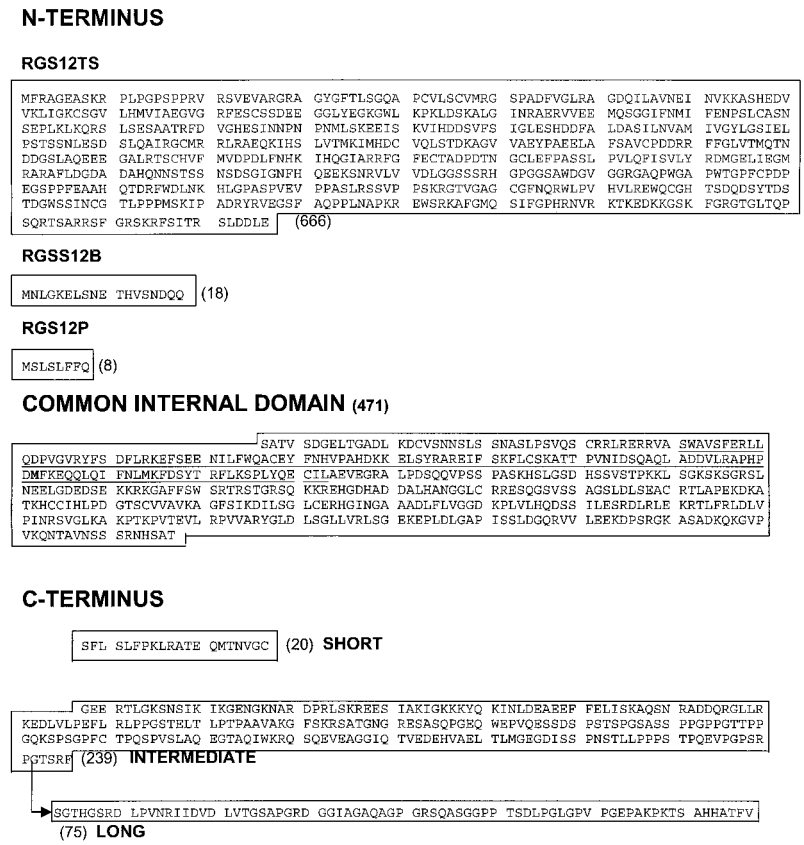

FIG. 1. **Organization of the amino acid sequences encoded by 12 RGS12 transcripts into variable N-terminal, common internal, and variable C-terminal cassettes.** The unique N-terminal sequences of RGS12TS, RGS12B, and RGS12P are shown. The common internal region is encoded by all splice variants of RGS12, although N-terminally truncated for RGS12L (the methionine start codon of RGS12L is shown in *boldface*). The unique C-terminal sequences of short, long, and intermediate forms of RGS12 are shown. The intermediate and long C-terminal forms of RGS12 share the first 235 amino acids with the C-terminal TSRF sequence of intermediate forms replaced with a unique 75-amino acid sequence in long forms. The 12 forms of RGS12 encoded by RGS12 mRNAs are produced by assembling different combinations of these domains.

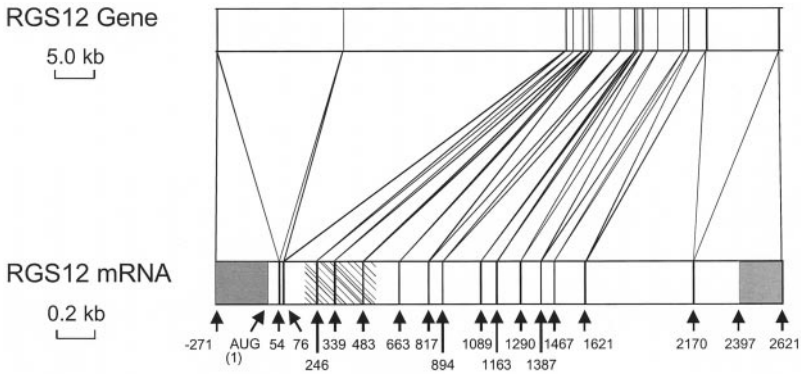

FIG. 2. **Intron-exon organization of the human *RGS12* gene in relation to the coding sequence of human RGS12B-L mRNA.** The position of exons (*filled*) and introns (*empty*) of the human *RGS12* gene are shown at the top. The structure of the human RGS12B-L mRNA is shown at the bottom with the locations of introns (*vertical lines*) indicated by the mRNA nucleotide number, with base 1 corresponding to the A of the AUG start codon. Sequences encoding the RGS domain of human RGS12B-L are indicated with *hatching*. *kb*, kilobases.

RESULTS

Identification of Splice Variant Human RGS12 cDNAs—We employed a PCR-based strategy to amplify and identify 12 splice variant forms of human RGS12 using sequence information from a 489-bp expressed sequence tag (GenBank™/EBI Data Bank accession number T57943). Koelle and Horvitz (3) first identified this expressed sequence tag as a potential RGS protein and named it RGS12. Based upon this sequence, we designed two forward and two reverse primers for use in 5′- and 3′-RACE to generate overlapping cDNA fragments of RGS12. 5′-RACE-Ready human brain cDNA and Marathon-Ready human lung and placenta cDNAs were used as templates in these reactions. Sequence information obtained from the resulting RACE products was used to design additional primers for amplification of full-length RGS12 cDNAs. Sequence analysis identified two forms of RGS12 amplified from human brain that differed at their 5′-ends. Similarly, the 5′-ends of RGS12 cDNAs amplified from lung and placenta were unique from each other and distinct from the brain-specific RGS12 cDNAs. In addition to tissue-specific diversity at the 5′-ends of mRNAs encoding RGS12, sequence analysis revealed that the 3′-ends of all of the different RGS12 cDNA forms existed in three different forms. The coding sequences of the 12 different RGS12 cDNAs ranged from 1068 to 4341 nucleotides, encoding proteins ranging from 356 to 1447 amino acids.

The four different types of 5′-ends of RGS12 cDNAs found in brain, placenta, and lung encode four different N-terminal sequences for the RGS12 proteins. We designated the proteins encoded by the two different RGS12 cDNAs amplified from brain as RGS12B (GenBank™/EBI Data Bank accession number AF030109) and RGS12TS (accession number AF030111) for *trans*-spliced (see below). The proteins encoded by RGS12 cDNAs amplified from placenta and lung were accordingly designated RGS12P (accession number AF030110) and RGS12L (accession number AF030112), respectively. The three different types of 3′-ends of the RGS12 cDNAs encoding these four types of RGS12 proteins encode three different C-terminal sequences that differed in both sequence and length. These three types of sequences were designated L for long, S for short, and I for intermediate to distinguish also the predicted RGS12

TABLE I
Intron-exon organization of the RGS12 gene

5′-Donor site	3′-Acceptor site	Intron no.	Intron size
			bp
TCT AAT GAC CAG CAGgtaag.....ttcagTCT GCA ACT GTG TCT Ser Asn Asp Gln Gln Ser Ala Thr Val Ser		1	15746
ACT GTG TCT GAT GGC Ggtaag..... cgcagAG TTG ACG GGC GCC Thr Val Ser Asp Gly G lu Leu Thr Gly Ala		2	27641
GTC CGC TAC TTC TCTgtgag .....cttagGAT TTT CTA AGG AAA Val Arg Tyr Phe Ser Asp Phe Leu Arg Lys		3	510
CAT GAC AAA AAG GAGgtaag .....gtcagCTT TCC TAC AGG GCC His Asp Lys Lys Glu Leu Ser Tyr Arg Ala		4	1054
GAG CAG CAG CTG CAGgtaac..... cccagATC TTC AAT CTC ATG Glu Gln Gln Leu Gln Ile Phe Asn Leu Met		5	791
TCC ACG CCA AAA AAGgtgac..... gacagTTA AGT GGA AAA TCA Ser Thr Pro Lys Lys Leu Ser Gly Lys Tyr		6	295
GGG GAC CAC GCA GAC Ggtttg.....tcaagAC GCC CTG CAT GCC Gly Asp His Ala Asp A sp Ala Leu His Ala		7	3099
GGG AGC CTG GAC CTGgtgag.....cccagTCG GAG GCC TGC AGG Gly Ser Leu Asp Leu Ser Asp Ala Cys Arg		8	1657
GTG GGC GGG GAC AAGgtact.....ctcagCCT CTG GTG CTG CAC Val Gly Gly Asp Lys Leu Leu Val Leu His		9	334
AAG CGC ACC TTG TTT CGgtaag.....accagG CTG GAT CTT GTT Lys Arg Thr Leu Phe Ar g Leu Asp Leu Val		10	536
CTG CTG GTG AGG CTGgtgag.....cctagAGT GGA GAG AAG GAG Leu Leu Val Arg Leu Ser Gly Glu Lys Asp		11	1822
CCT TCC AGA GGA AAG Ggtgag.....tttagCA TCC GCA GAT AAA Pro Ser Arg Gly Lys A la Ser Ala Asp Lys		12	3124
AAC CAC TCG GCT ACGgtaat.....tgcagGGA GAG GAA AGA ACA Asn His Ser Ala Thr Gly Glu Glu Arg Thr		13	388
TTG GAC GAA GCA GAG Ggtatg.....aatagAG TTT TTT GAG CTT Leu Asp Glu Ala Glu G lu Phe Phe Glu Leu		14	1691
GGA CCT TCC AGA CCA Ggtacc.....gacagGA AGT GGG ACC CAT Gly Pro Ser Arg Pro G ly Ser Gly Thr His		15	8498

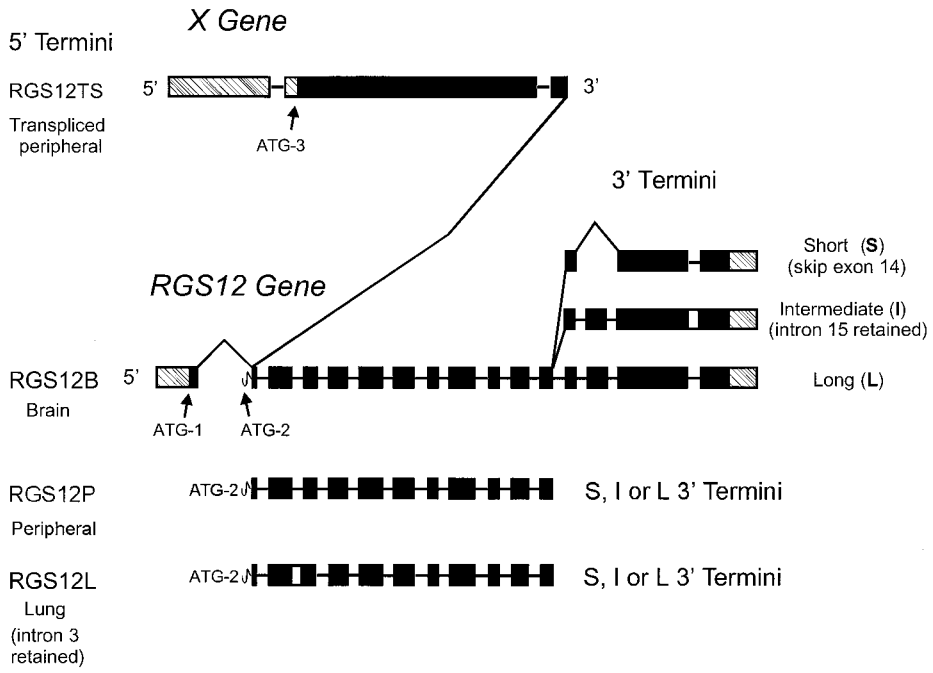

FIG. 3. **Representation of the splicing events occurring in the human *RGS12* gene to generate 12 splice variant RGS12 mRNAs.** Exons are shown as *filled boxes*, introns as *empty boxes*, and noncoding sequences as *hatched boxes*. Two primary transcripts encode the four 5′-splice forms of RGS12; one is processed to RGS12B and RGS12TS, and one is processed to RGS12P and RGS12L. The primary transcript encoding RGS12B and RGS12TS mRNAs includes exon 1 of the human *RGS12* gene, and that encoding RGS12P and RGS12L mRNAs uses an alternate transcriptional start site within intron 1 of the human *RGS12* gene. The *middle* of the figure shows how 3′-splicing generates the long, short, and intermediate transcripts of RGS12B as well as the 3′-splice forms of RGS12TS, RGS12P, and RGS12L mRNAs. *trans*-Splicing of a transcript encoded by three exons of an unknown gene (X gene) generates RGS12TS. RGS12L arises by retention of intron 3 of the human *RGS12* gene.

proteins on the basis of their C-terminal sequences. Thus, the long C-terminal tail form of RGS12B was designated RGS12B-L.

With the exception of the RGS12L isoforms, all RGS12 proteins share a common internal region of 471 amino acids flanked by unique N- and C-terminal sequences (Fig. 1). RGS12L proteins share only 336 amino acids of this common internal domain. This is a result of N-terminal truncation of the RGS12L proteins due to the presence of a stop codon in the 5′-end of their cDNAs that is in frame with the RGS domain. Fig. 1 shows that the RGS domain (*underlined*) is located in the common internal region. The RGS domain is N-terminally truncated in RGS12L proteins and comprises the N terminus of these RGS12 isoforms. RGS12TS, RGS12B, and RGS12P have unique N-terminal domains of 666, 18, and 8 amino acids, respectively. The region C-terminal to the common internal domain of RGS12 proteins exists as a short (20 amino acids), intermediate (239 amino acids), or long (310 amino acids) form. The intermediate and long C-terminal forms share a 235-amino acid sequence.

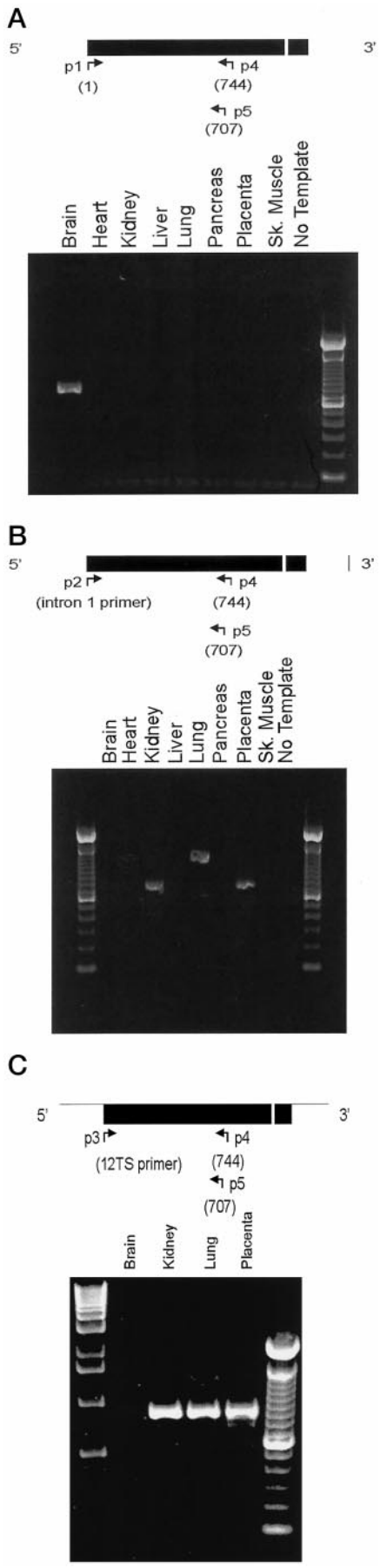

FIG. 4. **PCR analysis of tissue-specific expression of RGS12 transcripts with unique 5′-ends.** Semi-nested PCR was performed utilizing primers (p1, p2, and p3) specific for sequences encoding the unique 5′-ends of RGS12 transcripts and the Quick-Screen human cDNA library panel as templates. Two nested reverse primers (p4 and p5) utilized in these PCRs were from a region common to all RGS12 transcripts derived from exon 7 of the *RGS12* gene. Each panel shows a diagram of the RGS12 cDNA with locations of the forward and semi-nested reverse primers used in PCR and the agarose electrophoresis analysis of the amplified products. *A*, PCR amplification with a forward primer spanning the translational start site of RGS12B; *B*, PCR amplification with a forward primer designed to the sequence derived from intron 1 of the *RGS12* gene that represents the 5′-end of both RGS12P and RGS12L; *C*, PCR amplification with a forward primer designed to a sequence present in the N terminus of RGS12TS that is derived from the second exon of the X gene. Molecular sizes of amplified products were determined from 1-kilobase pair and 100-bp DNA ladders (Life Technologies, Inc.) run in parallel. *Sk.*, skeletal.

Structure of the Human RGS12 Gene—The observation that the 12 RGS12 isoforms could be constructed from different combinations of N- and C-terminal cassettes with a common internal domain suggested that these forms might arise by alternative splicing of a single gene. A BLAST nucleotide sequence similarity search of the human gene data banks with the cloned human RGS12 cDNAs revealed that the *RGS12* gene is located within the Huntington disease region of chromosome 4p16.3. RGS12B-L showed complete homology to interrupted segments of human genomic DNA sequences from cosmid HS361H4B, HS361H4C, HSL60G9A, and HSL60G9B. This enabled us to determine the complete structure of the *RGS12* gene. The coding region of the *RGS12* gene spans 69333 bp of DNA and is interrupted by 15 introns. The intron-exon organization of the *RGS12* gene in relation to the RGS12B-L mRNA is shown in Fig. 2. The *RGS12* gene contains 16 exons within its coding region that range in size from 22 to 549 bp. Table I shows the sizes of introns and the intron-exon splice junction sequences in the coding region of the *RGS12* gene. As shown, the introns vary in size from 295 to 27.6 kilobases, and all of the splice acceptor and donor sequences agree with the GT/AG consensus sequence (15). The *RGS12* gene intron phasing is type 0 (the intron occurs between codons) for introns 1, 3–6, 8, 9, 11, and 13; type 1 (the intron interrupts the first and second bases of the codon) for introns 2, 7, 12, 14, and 15; and type 2 (the intron interrupts the second and third bases of the codon) for intron 10.

Splicing of Human RGS12 Transcripts—We examined the relationship between exon and intron locations of the *RGS12* gene to the structure of the 12 RGS12 variants we isolated to gain insight into how the unique N- and C-terminal forms of RGS12 proteins arise. RGS12B-L results from splicing of exons 1–16 (Fig. 3). Exon 1 encodes both the 5′-untranslated region and the unique N-terminal domain of RGS12B, whereas the shared common internal domain found in all RGS12 proteins is encoded by exons 2–13. The short and intermediate C-terminal forms of RGS12B and all other RGS12 proteins arise by skipping of exon 14 and retention of intron 15, respectively (Fig. 3). Although the transcript encoding the intermediate C-terminal forms of RGS12 proteins is longer than that encoding the long forms, the presence of a stop codon produces a shorter protein. Thus, the *RGS12* gene can be alternatively spliced near its 3′-end to generate the three types of C-terminal sequences present on the four N-terminal forms of RGS12L.

Unlike RGS12B proteins, the unique N-terminal domains of RGS12P, RGS12L, and RGS12TS proteins are not encoded by sequences from exon 1 of the *RGS12* gene. Comparison of the sequence encoding the N-terminal domain of RGS12P with that of the *RGS12* gene revealed that the unique N-terminal domain of RGS12P proteins is encoded by sequences from intron 1 of the *RGS12* gene. The 5′-end of transcripts encoding RGS12P proteins retains 1384 bp of the 3′-end of intron 1, presumably

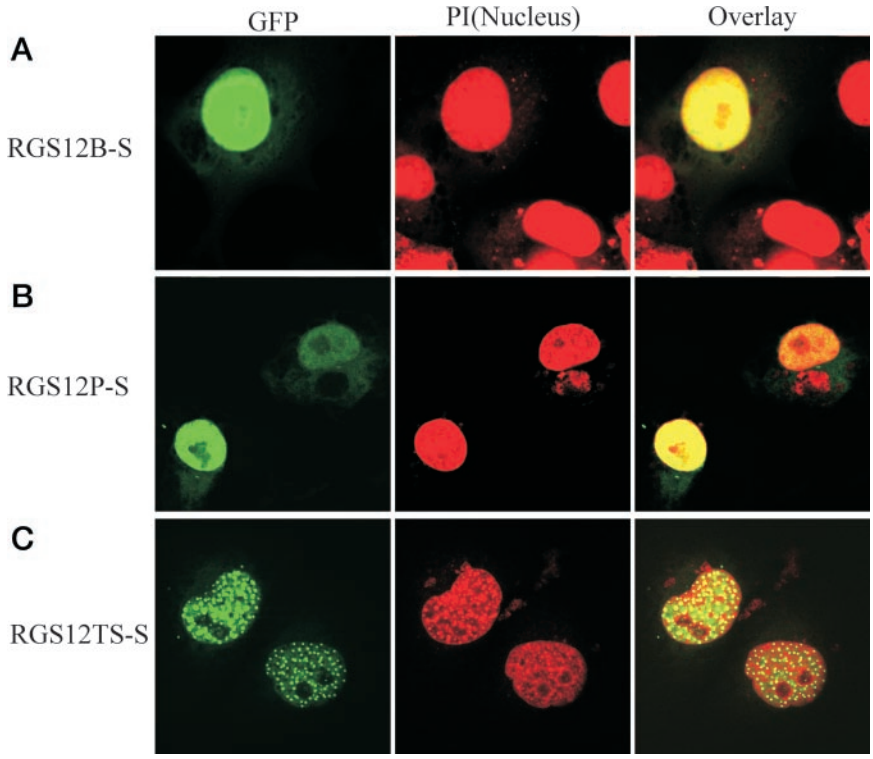

FIG. 5. **Confocal microscope images of COS-7 cells expressing short forms of RGS12B, RGS12P, and RGS12TS tagged at their C termini with GFP.** *Green* represents GFP fluorescence from expressed RGS12 proteins; *red* represents propidium iodide (*PI*)-stained cell nuclei; and *yellow* (in the overlay image) represents overlapping *green* and *red* fluorescence. Transfection of COS-7 cells and immunofluorescence measurements were performed as described under "Experimental Procedures."

a result of an alternate transcriptional start site at this location. An alternate translational start site in frame with the RGS domain sequence is located 1360 bp from the 5′-end of the RGS12P transcript. Thus, this translational start site is encoded by sequences located 24 bp from the 3′-end of intron 1 (Fig. 3). A similar analysis showed that the 5′-end of transcripts encoding RGS12L retains both the 1384-bp intron 1 sequence as well as all of intron 3 (Fig. 3). The N-terminal truncation of RGS12L within the common internal region of RGS12 proteins results from a stop codon present in the retained intron 3 sequences. The only start codon that could produce a translational product in frame with the RGS domain in RGS12TS transcripts is present in exon 5 and corresponds to Met^{136} of the common internal region (Fig. 1). As a result, the C-terminal half of the RGS domain encodes the N-terminal domain of RGS12L proteins. Thus, the 5′-ends of transcripts of both RGS12P and RGS12L are located within sequences spliced out of transcripts encoding RGS12B.

The sequence encoding the unique N-terminal region of RGS12TS proteins is not present within or upstream of the *RGS12* gene. A BLAST search of the human gene data banks with the sequence from the 5′-end of RGS12TS transcripts showed complete homology to interrupted segments from cosmids HSL185E6A and HSL21F12B. These cosmids map to a site on chromosome 4p16.3 located 170 kilobases from the 3′-end of the *RGS12* gene from the DNA (−)-strand. The 5′-ends of RGS12TS transcripts are derived from nucleotides 15808–15924 of the (−)-strand of HSL21F12B and nucleotides 12599–14580 and 15701–16502 of the (−)-strand of HSL185E6A. Thus, the unique N-terminal domain and the 5′-untranslated region of RGS12TS appear to be encoded by a transcript arising from three exons of an unknown gene ("X" gene) that splices onto exon 2 of the *RGS12* gene to form the full-length RGS12TS transcript. This phenomenon of splicing of two independent transcripts, known as "*trans*-splicing" (16), is much less common than the *cis*-splicing used to generate alternatively spliced variants from within a single gene. Accordingly, the proteins encoded by RGS12 transcripts with this unique 5′-end were designated RGS12TS.

Tissue Distribution of Human RGS12 Transcripts—The divergence in sequence of cDNAs encoding different N-terminal forms of RGS12 proteins enabled us to design primers to examine expression of their transcripts in tissues by PCR. For these experiments, we used the Quik-Screen library panel of cDNAs and primers specific for sequences encoding the unique 5′-ends of RGS12 transcripts. Semi-nested PCR was performed with a single specific forward primer and two nested reverse primers to a region common to all RGS12 transcripts derived from exon 7 of the *RGS12* gene. Sequencing of resulting PCR products confirmed their identities as the expected amplification products. Fig. 4*A* shows that PCR with a forward primer spanning the translational start site of RGS12B (exon 1 of the *RGS12* gene) amplified the expected product from brain cDNA, but not from other tissue cDNAs. Fig. 4*B* shows the results of PCR amplification using a forward primer designed to the sequence derived from intron 1 of the *RGS12* gene that serves as the 5′-end of both RGS12P and RGS12L. As shown, products of the expected size were amplified from kidney, placenta, and lung, *i.e.* the lung product is larger due to the retention of sequences derived from intron 3. Sequencing of the smaller faint band amplified from lung showed that it was identical to the products amplified from placenta and kidney that correspond to the 5′-end of RGS12P cDNAs. Thus, the major amplification product in lung corresponds to RGS12L, although a product corresponding to RGS12P can also be detected. Fig. 4*C* shows the results of PCR amplification using a forward primer designed to a sequence present in the N terminus of RGS12TS that is derived from the second exon of the X gene. The expected PCR products were amplified from kidney, lung, and placenta, but not brain. These results show that transcripts encoding different N-terminal forms of RGS12 proteins have both common and unique tissue expression patterns.

Subcellular Localization of Human RGS12 Proteins—The different forms of RGS12 were tagged at their C termini with either GFP or c-Myc to examine their expression by confocal microscopy and immunoblotting, respectively. We focused our experimental attention on the RGS12 proteins possessing a complete common internal region so that any differences in

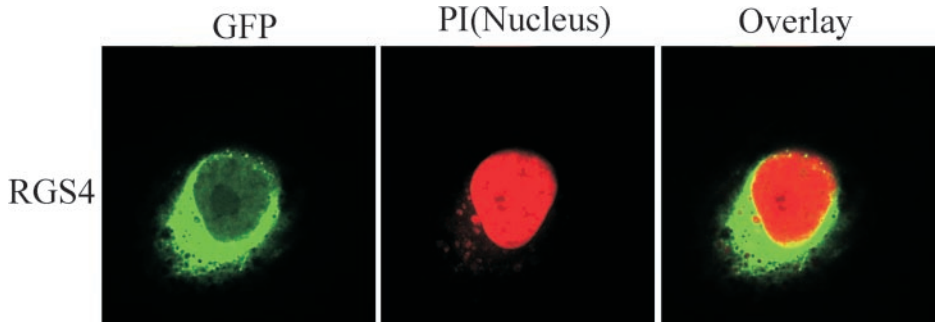

FIG. 6. **Confocal microscope images of COS-7 cells expressing human RGS4 protein tagged with GFP at its C terminus.** *Green* represents GFP fluorescence from expressed RGS4; *red* represents propidium iodide (*PI*)-stained cell nuclei; and *yellow* (in the overlay image) represents overlapping *green* and *red* fluorescence. Transfection of COS-7 cells and immunofluorescence measurements were performed as described under "Experimental Procedures."

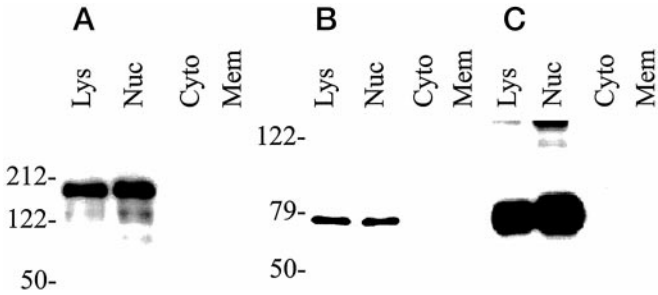

FIG. 7. **Immunoblots of subcellular fractions of COS-7 cells expressing C-terminal c-Myc-tagged forms of RGS12TS-S (*A*), RGS12B-S (*B*), and RGS12P-S (*C*).** COS-7 cell transfectants were lysed (*Lys*) and subfractionated into nuclear (*Nuc*), cytosolic (*Cyto*), and membrane (*Mem*) fractions as described under "Experimental Procedures." c-Myc immunoblotting was performed on the resulting cell lysates and an equivalent amount of each fraction as described under "Experimental Procedures." Locations of molecular mass markers (in kilodaltons) are shown to the left of the autoradiograms.

expression or intracellular distribution of these proteins could be correlated to splicing events leading to different N- or C-terminal forms of the proteins. Fig. 5 shows confocal microscope images of COS-7 cells expressing the GFP-tagged short forms of RGS12B, RGS12P, and RGS12TS. The *green color* represents the GFP fluorescence from expressed RGS12 proteins, and the *red color* represents fluorescence from propidium iodide staining of nuclei in these cells. To confirm any apparent nuclear localization of expressed RGS12 proteins by visualization of GFP fluorescence alone, an overlay image of the GFP and propidium iodide fluorescence is also shown. As shown, all three RGS12 proteins were localized exclusively in the nucleus of COS-7 cell transfectants. Although RGS12B-S and RGS12P-S exhibited a homogeneous pattern of distribution throughout the nucleus, RGS12TS-S exhibited a unique pattern of dotted distribution within the nucleus. No evidence was obtained for accumulation of any of these proteins in the cytoplasm or plasma membrane. The nuclear localization of RGS12 proteins is quite distinct from the predominantly cytoplasmic localization of RGS4 that is observed in COS-7 cells (Fig. 6). As an alternate approach to examine the intracellular distribution patterns of RGS12 proteins, we performed immunoblotting of subcellular fractions of COS-7 cells expressing Myc-tagged forms of these proteins. Fig. 7 shows that these three RGS12 proteins were localized nearly exclusively in the nuclear fraction of COS-7 cells, in agreement with the confocal microscopic evidence for their nuclear patterns of distribution in COS-7 cells.

In contrast to the reproducible nuclear expression of the short splice forms of RGS12 proteins in COS-7 cells, we were unable to detect expression of the long splice forms of these proteins. No GFP fluorescence was detected in COS-7 cells transfected with the GFP-tagged long splice form of RGS12B, RGS12P, or RGS12TS. Immunoblotting of lysates of cells transfected with c-Myc-tagged forms of these proteins similarly showed no detectable expression of the proteins. Treatment of transfectants with inhibitors of proteasome (10 μM MG-132)- or calpain (25 μM *N*-acetyl-leucyl-leucyl-norleucinal)-dependent proteolytic pathways did not promote expression of the long splice forms of RGS12, suggesting that these pathways likely do not contribute to the lack of expression of the long splice forms of RGS12. We prepared a GFP-tagged construct of the intermediate splice form of RGS12B and similarly were unable to detect its expression. These results suggested that lack of expression of the long and intermediate splice forms of RGS12 was due to sequences in the mRNA or protein that encode or comprise, respectively, their unique C-terminal domains. Therefore, we examined whether the unique 310-amino acid C-terminal domain of the long splice forms of RGS12 could be expressed as a GFP fusion protein. Fig. 8 shows that this domain was expressed in both the nucleus and cytoplasm of COS-7 cells. Thus, if this domain contributes to the inability to express the long splice forms of RGS12, it must do so within the context of a holoprotein or corresponding mRNA. Alternatively, we reasoned that the lack of expression of long splice forms of RGS12 could result from an inability of the cells expressing these proteins to survive. To consider this possibility, we cotransfected cells with untagged RGS12TS-L (in pCR3.1) and the GFP vector alone to determine whether expression of GFP would be diminished in cells cotransfected with RGS12TS-L compared with those cotransfected with the GFP and pCR3.1 vectors. Both transfectants exhibited comparable and robust expression of GFP (data not shown).

Subcellular Localization of Native RGS12TS-S—We raised an antibody to one of the N-terminal forms of RGS12 for use in determining whether native RGS12 proteins show a pattern of nuclear localization as observed in COS-7 cells expressing recombinant forms of these proteins. This antibody was raised to a synthetic peptide corresponding to residues 1–15 of RGS12TS proteins. This polyclonal anti-RGS12TS antibody recognized ectopically expressed RGS12TS-S following immunoblotting or immunocytochemistry (see below). We used the anti-RGS12TS antibody to screen various human cell lines for endogenous RGS12TS immunoreactivity and found that HEK-293T cells express appreciable levels of RGS12TS. Fig. 9 shows that RGS12TS immunoreactivity in HEK-293T cells migrated with the same molecular size as RGS12TS-S-Myc transiently expressed in COS-7 cells, suggesting that HEK-293T cells endogenously express the RGS12TS protein variant we designated RGS12TS-S. Interestingly, although SV40-transformed HEK-

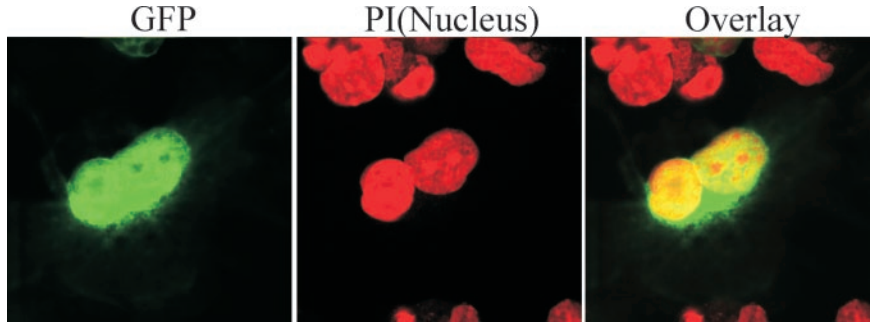

FIG. 8. **Confocal microscope images of COS-7 cells expressing the 310-amino acid C-terminal domain of long splice forms of RGS12 proteins tagged with GFP at their C termini.** *Green* represents GFP fluorescence from expressed RGS4; *red* represents propidium iodide (*PI*)-stained cell nuclei; and *yellow* (in the overlay image) represents overlapping *green* and *red* fluorescence. Transfection of COS-7 cells and immunofluorescence measurements were performed as described under "Experimental Procedures."

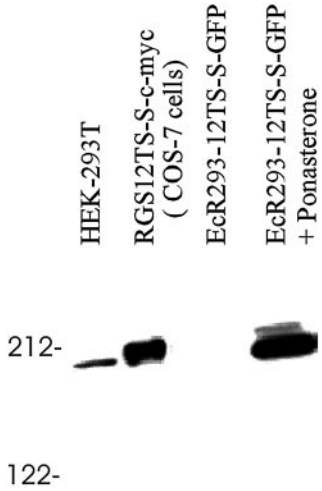

FIG. 9. **RGS12TS-S is expressed endogenously in HEK-293T cells.** RGS12TS-S immunoblotting was performed on lysates derived from HEK-293T cells, COS-7 cells transiently transfected with c-Myc-tagged RGS12TS-S, and EcR293–12TS-S-GFP cells before and after 5 μM ponasterone induction for 12 h. Immunoblotting was performed as described under "Experimental Procedures." Locations of molecular mass markers (in kilodaltons) are shown to the left.

293T cells express RGS12TS-S, HEK-293 cells do not. We developed stable transfectants of HEK-293 cells expressing C-terminal GFP-tagged RGS12TS-S under the control of an ecdysone-inducible promoter and named this stable cell line EcR293–12TS-S-GFP. As shown in Fig. 9, RGS12TS immunoreactivity was present in lysates of these cells following ponasterone (5 μM) treatment to induce expression of RGS12TS-S-GFP, but not in lysates from vehicle-treated cells. The RGS12TS immunoreactivity migrated as a doublet with the higher molecular mass band corresponding to the size expected for RGS12TS-S-GFP. Both RGS12TS-immunoreactive bands in ponasterone-induced EcR293–12TS-S-GFP lysates were detected also by immunoblotting with anti-GFP antibodies (data not shown). The nature of heterogeneity in size of induced RGS12TS-S-GFP is unclear, but may reflect post-translational changes or proteolysis of the protein. Together, these results illustrate that the anti-RGS12TS antibody detects both native and ectopically expressed RGS12TS-S.

We next examined the subcellular localization of native RGS12TS-S in HEK-293T cells by subcellular fractionation and confocal microscopy. Fig. 10*A* shows that RGS12TS-S is found exclusively in the nuclear fraction of HEK-293T cells. In agreement with these findings, confocal microscopy of HEK-293T cells showed that RGS12TS immunoreactivity in HEK-293T cells exhibits a dotted pattern of nuclear distribution (Fig. 10*B*). Thus, the pattern of expression of endogenously expressed RGS12TS-S in HEK-293T cells is very similar, if not the same, as that observed in COS-7 cells overexpressing RGS12TS-S-GFP (Fig. 5).

Cell Cycle Regulation of RGS12TS-S Nuclear Dots—The dotted intranuclear distribution of RGS12TS-S is strikingly similar to that of promyelocytic leukemia protein PML, leukemia-associated protein ALL-1, and breast cancer-associated proteins BRCA1 and BARD1 (17–20). At present, the precise functional significance of the distribution of these proteins in nuclear dots is not understood. However, PML protein promotes transcriptional silencing/enhancing and exhibits tumor suppressor activities, and its dotted nuclear distribution delocalizes during acute promyelocytic leukemia and certain viral infections (21–26). BRCA1 functions as a tumor suppressor protein and localizes to nuclear dots during the S phase of the cell cycle (19). Hydroxyurea-mediated DNA synthesis arrest induces loss of BRCA1 nuclear foci, a response accompanied by hyperphosphorylation of BRCA1. Thus, BRCA1 S phase nuclear dots are dynamic elements, responsive to DNA damage, and involved in replication checkpoint responses (19).

Therefore, we examined whether RGS12TS-S nuclear dots are cell cycle-regulated by determining the nuclear immunolocalization pattern of RGS12TS-S in HEK-293T cells blocked at different stages of the cell cycle (Fig. 11). A pattern of evenly distributed RGS12TS-S nuclear dots was present in asynchronously growing HEK-293T cells. RGS12TS nuclear dots were unevenly distributed and somewhat clumped together in cells blocked at the G_1/S cell cycle boundary by treatment with hydroxyurea, aphidicolin, or thymidine. However, nocodazole-mediated blockade of HEK-293T cells at the G_2/M cell cycle boundary produced loss of RGS12TS-S nuclear dots and a diffuse distribution of RGS12TS in the nucleoplasm. A similar loss of RGS12TS-S nuclear dots occurred in quiescent cells (G_0) that exited the cell cycle by serum deprivation. The loss of RGS12TS-S nuclear foci in HEK-293T cells at G_0 and the G_2/M cell cycle boundaries prompted us to examine the localization of RGS12TS-S in cells undergoing mitosis. Fig. 11*B* shows that RGS12TS-S overlapped with the metaphase chromosome and evenly distributed to dividing chromosomes. The amount of RGS12TS-S protein in HEK-293T cells arrested at various phases of the cell cycle did not differ significantly (Fig. 11*C*). These findings suggest that RGS12TS-S localizes to nuclear dots during the G_1/S phase of the cell cycle. The localization of RGS12TS-S to discrete nuclear foci in HEK-293T cells and the loss of this localization during the G_2/M and G_0 phases of the cell cycle appear to be independent of changes in the steady-state level of RGS12TS-S protein.

Aberrant Nuclear Morphology and Multiple Nuclei in COS-7 Cells Overexpressing RGS12TS-S—Approximately 10% of COS-7 cells transiently expressing RGS12TS-S-GFP exhibited aberrant nuclear morphologies. Fig. 12*A* illustrates an example of a dumbbell-shaped nucleus that also has nuclear buds or blebs. In addition, a significant number of RGS12TS-S-GFP-expressing COS-7 cells exhibited multiple nuclei (Fig. 12*B*). Both of these nuclear anomalies were not observed with other GFP-tagged RGS proteins including RGS2, RGS4, RGS10, RGS16, RGS12B-S, and RGS12P-S or in cells transfected with

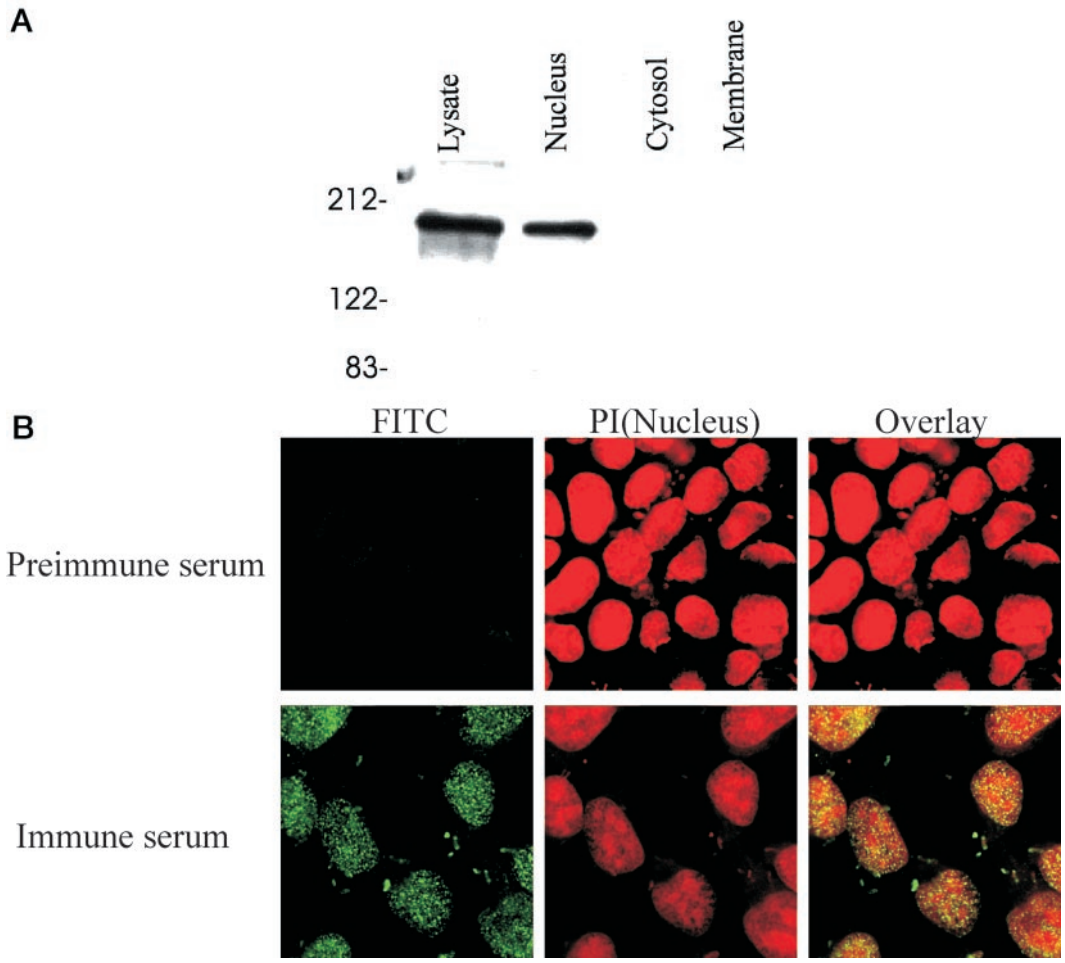

FIG. 10. **Subcellular localization of native RGS12TS-S.** *A*, immunoblot of subcellular fractions of HEK-293T cells endogenously expressing RGS12TS-S. HEK-293T cells were lysed and subfractionated into nuclear, cytosolic, and membrane fractions as described under "Experimental Procedures." Immunoblotting with RGS12TS-specific antibody was performed on the resulting lysate and an equivalent amount of each fraction as described under "Experimental Procedures." Locations of molecular mass markers (in kilodaltons) are shown to the left. *B*, confocal microscope images of HEK-293T cells endogenously expressing RGS12TS-S. *Green* represents fluorescein isothiocyanate (*FITC*) fluorescence from RGS12TS immunoreactivity; *red* represents propidium iodide (*PI*)-stained cell nuclei; and *yellow* (in the overlay image) represents overlapping *green* and *red* fluorescence. Immunofluorescence measurements of endogenously expressed RGS12TS were performed as described under "Experimental Procedures."

GFP vector alone (data not shown). These nuclear abnormalities also were not observed in EcR293-TS-S-GFP cells following induced expression of RGS12TS-S-GFP. These results suggest that ectopic expression of RGS12TS-S dysregulates nuclear division and/or cytokinesis in a subpopulation of COS-7 cells and that this response is cell type-dependent.

DISCUSSION

This study has elucidated a remarkable complexity in processing of the *RGS12* gene to produce 12 distinct forms of RGS12 differing in both N- and C-terminal sequences. These results are the first to document alternative splicing of an RGS protein gene. Thus, both genes for RGS proteins and their presumed regulatory targets, Gα subunits, may be spliced differentially to produce proteins with distinct structural features. Our results show that transcripts encoding the different N-terminal forms of RGS12 are expressed in a tissue-specific fashion and that the expression and intracellular pattern of distribution of RGS12 proteins are affected by splicing at the 5′- and 3′-ends of their corresponding mRNAs. Our demonstration that the short forms of RGS12 proteins are nuclear proteins suggests that the functional role of these proteins might be very different from that proposed for many members of the RGS protein family. Indeed, we demonstrate here that native RGS12TS-S exhibits a pattern of subnuclear targeting to nuclear dots (characteristic of various tumor suppressor proteins) that is cell cycle-dependent and associates with the metaphase chromosome during mitosis. Ectopic expression of RGS12TS-S in COS-7 cells induced nuclear abnormalities and multinucleated cells that were not observed in cells expressing other RGS proteins.

The complexity and mechanism of splicing of the *RGS12* gene are extraordinary. First, there appears to be two distinct primary transcripts arising from the *RGS12* gene, one that includes exon 1 and one that does not. The latter apparently arises from an alternate transcriptional start site resulting in retention of sequences from intron 1 and an in-frame translational start site located 24 bp 5′ to exon 2. Second, the primary transcript lacking exon 1 can be spliced to remove or retain intron 3, producing the RGS12P and RGS12L transcripts, respectively. Third, the primary transcript retaining exon 1 undergoes *cis*- or *trans*-splicing to generate RGS12B and RGS12TS, respectively. Finally, the 3′-ends of the RGS12 transcripts are spliced to retain exons 14–16, to skip exon 14, or to retain intron 15 to produce the long, short, and intermediate 3′-splice forms of RGS12. This splicing appears to occur in each of the four different 5′-splice forms of RGS12. Reverse transcription-PCR analysis showed that the RGS12 transcripts encoding the four N-terminal forms of RGS12 are expressed in a tissue-specific fashion, *i.e.* with transcripts encoding RGS12B

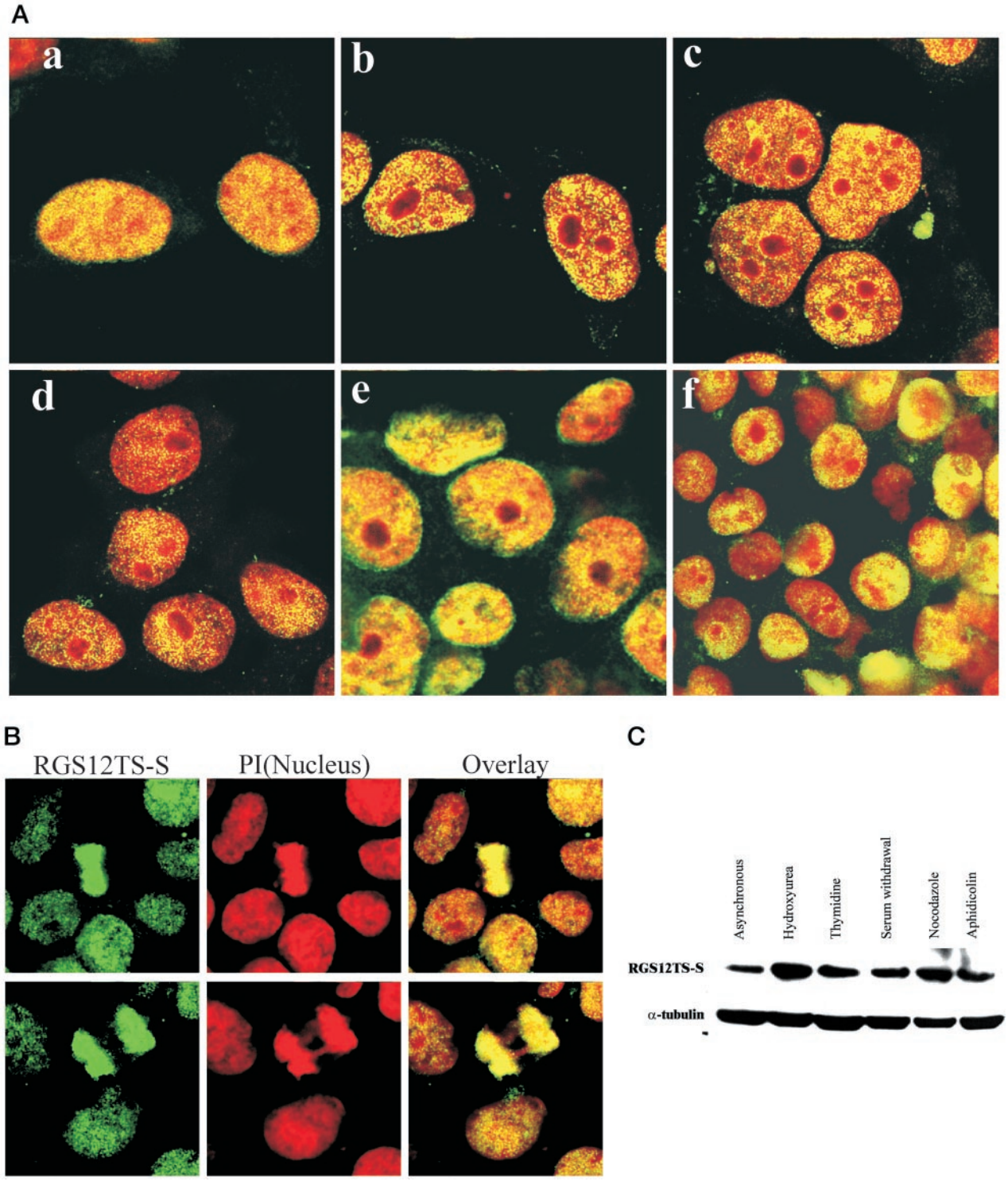

FIG. 11. **Cell cycle regulation of RGS12TS-S nuclear dots.** *A*, confocal microscope overlay images of endogenously expressed RGS12TS-S in HEK-293T cells growing asynchronously (*panel a*) or blocked at various phases of the cell cycle by treatment for 24 h with hydroxyurea (*panel b*), thymidine (*panel c*), aphidicolin (*panel d*), and nocodazole (*panel e*). HEK-293T cells were also deprived of serum (*panel f*) by culturing for 32 h in serum-free DMEM. Endogenously expressed RGS12TS-S (*green*) is seen against the propidium iodide-stained cell nucleus (*red*). *B*, confocal microscope images of RGS12TS-S localization within mitotic HEK-293T cells. *Green* fluorescence represents RGS12TS-S immunoreactivity, and *red* fluorescence represents propidium iodide (*PI*)-stained cell nuclei. *C*, immunoblot of RGS12TS-S in lysates of HEK-293T cells blocked at various phases of the cell cycle. Immunoblotting of α-tubulin is shown as a control for protein loading.

present only in brain, those encoding RGS12L present only in lung, and those encoding RGS12P and RGS12TS present in several tissues. Thus, these unique transcripts appear to originate in an orderly and tissue-specific fashion.

RGS12TS is unique in being the only RGS12 protein encoded by splicing of transcripts from discontinuous gene segments, a particularly novel *trans*-splicing mechanism that has been described previously in only a limited number of genes. Such intermolecular splicing has been implicated in the generation of mRNA species in protozoa, plants, and mammals (16). In mammals, rat androgen-binding protein/sex hormone-binding globulin gene transcripts, located on chromosome 10, are fused

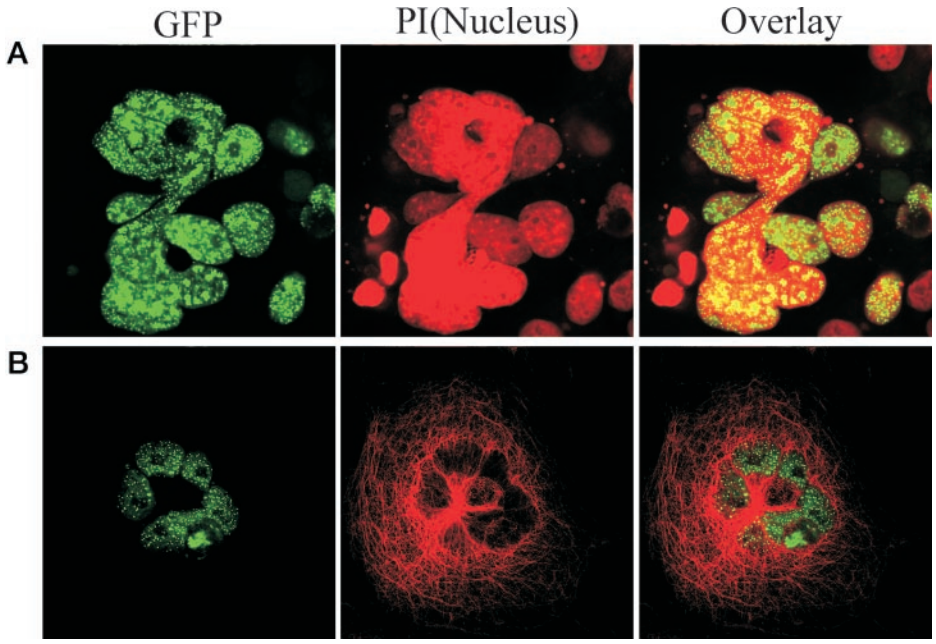

FIG. 12. **RGS12TS-S induces formation of abnormally shaped and multiple nuclei in COS-7 cells.** *A*, confocal microscope images depicting nuclear anomalies in COS-7 cells ectopically expressing RGS12TS-S-GFP. *Green* represents GFP fluorescence from expressed RGS12TS-S, and *red* fluorescence represents propidium iodide (*PI*)-stained cell nucleus. *B*, confocal microscope images showing formation of multiple nuclei in transfected COS-7 cells ectopically expressing RGS12TS-S-GFP. Immunostaining of α-tubulin was performed for visualization of the cell cytoplasm. *Green* represents GFP fluorescence from expressed RGS12TS-S, and *red* fluorescence represents endogenous α-tubulin immunoreactivity.

in mature mRNA transcripts to histidine decarboxylase gene transcripts, originating from chromosome 3 (27). Similar *trans*-splicing of coding sequences from different chromosomes generates c-Myb mRNA in human thymic cells (28). *trans*-Splicing also generates multi-isotype immunoglobulin transcripts from a single B lymphocyte and multiple types of chimeric germ-line immunoglobulin heavy chain transcripts in human B cells (29, 30). We considered the possibility that RGS12TS could arise from a recombinant gene, arising in a somatic cell from chromosomal breakage and gene rearrangement. However, no evidence indicates the occurrence of such widespread gene rearrangements in a subpopulation of cells in healthy individuals (*i.e.* CLONTECH cDNAs), and the existing gene data bases show that the sequence encoding the 5′-end of RGS12TS is located 170 kilobases downstream from the *RGS12* gene on the opposite DNA strand.

It seems likely that the complex pattern of splicing of RGS12 mRNAs is reflective of cellular mechanisms for orchestrating a diversity of functions for the encoded RGS12 proteins. Our results suggest that splicing at the 3′-end of RGS12 transcripts affects expression/stability of the protein, whereas RGS12 transcripts encoding three different N-terminal forms of RGS12 showed equivalent levels of expression in the nucleus. However, splicing at the 5′-end of the primary RGS12 transcript can produce RGS12 proteins that are targeted to discrete intranuclear sites as observed for RGS12TS-S. On the other hand, retention of RGS12 mRNA sequences encoding the 491-amino acid C-terminal region of all short forms of RGS12 may reflect the importance of this region for RGS12 activity. Although this region includes the conserved RGS domain, it shares no homology with any other protein except RGS14, whose function is not yet known.

Both ectopically expressed short forms of RGS12 proteins and native RGS12TS-S in HEK-293T cells are localized to the nucleus, suggesting that these proteins are constitutively nuclear proteins; yet we cannot discount the possibility that RGS12 proteins might translocate out of the nucleus in response to some type of cellular stimulus. The presence of RGS proteins in the nucleus suggests a dramatic separation in space from their presumed regulatory targets, *i.e.* Gα proteins. We are unaware of evidence for nuclear localization of Gα subunits, although Park *et al.* (31) recently showed that G$\beta\gamma_5$ complexes localize to the nucleus in NIH 3T3-L1 cells and attenuate the transcriptional repression activity of adipocyte enhancer-binding protein (AEBP1). If Gα subunits do not localize to and exert regulatory effects in the nucleus, it seems possible that RGS proteins could have other functions in the nucleus. This seems particularly likely for RGS12 proteins in view of their apparent exclusive localization within the nucleus and the considerable contribution of non-RGS domain sequences to the primary structure of these proteins. However, RGS12 proteins are not unique among RGS proteins in their nuclear localization. We recently showed that RGS4 and RGS16, although predominantly cytoplasmic proteins, shuttle in and out of the nucleus via a leptomycin B-sensitive pathway, whereas RGS2 and RGS10 are predominantly nuclear proteins (14).

However, RGS12TS-S is unique among RGS proteins in its subnuclear organization into nuclear dots, a feature shared by various tumor suppressor proteins. Our results suggest that RGS12TS-S localizes to nuclear dots during the G_1/S phase of the cell cycle and that this occurs independent of changes in the steady-state concentration of the protein. BRCA1 and BARD1 also localize to nuclear dots during progression to S phase and disperse thereafter (19, 20). Although BRCA1 localization into these structures is accompanied by increases in both its steady-state level and phosphorylation state, localization of BARD1 in these nuclear dots occurs independent of changes in its content (20). Whether the mechanism underlying formation of RGS12TS-S nuclear dots involves its phosphorylation or other modification and/or its interaction with other proteins as part of a multiprotein complex is not known at present. However, RGS12TS-S is rich in consensus sites for phosphorylation by various protein kinases and possesses various protein-binding modules including a PDZ domain (amino acids 30–99), a phosphotyrosine-binding domain (amino acids 225–374), a phosphotyrosine-interacting domain (amino acids 928–939), two Ras-like Raf-binding domains (amino acids 962–1032 and 1034–1104), and a guanylyl nucleotide exchange factor domain (amino acids 1187–1209) in addition to its RGS domain (amino acids 715–832). Although the precise function of BRCA1 is unknown, the sensitivity of S phase BRCA1 nuclear dots to DNA damage (19), the inability of BRCA1-deficient stem cells to carry out transcription-coupled DNA repair (32), and the presence of germ-line mutations in BRCA1 in approximately half the heritable forms of breast and ovarian cancers (33) suggest an important role in genome surveillance.

Thus, it is interesting to note that RGS12TS-S is unique among various RGS proteins we examined in its ability to produce abnormal nuclear morphology and multinucleated cells when ectopically expressed in COS-7 cells. The mechanism by which RGS12TS-S influences nuclear morphology is not clear; however, it appears that ectopic expression of RGS12TS-S dysregulates cytokinesis and/or nuclear division in COS-7 cells. The precise biochemical events involved in regulation of nuclear and cell division processes are not fully understood. Among various proteins implicated in cytokinesis and DNA replication checkpoint control are aurora and Ipl1-like midbody-associated protein (AIM-1) and ataxia telangiectasia mutated- and Rad3-related protein (ATR) (34, 35). Ectopic expression of AIM-1 or its kinase-inactive mutant induces abnormally shaped and multiple nuclei in cells (34), as we observed here for RGS12TS-S. Obviously, further studies are needed to define the precise biological role of RGS12TS-S and the mechanisms underlying its localization/delocalization in nuclear dots and its ability to induce nuclear anomalies upon its overexpression. The observed cell cycle-dependent localization of RGS12TS-S in nuclear dots and its ability to induce nuclear aberrations may indicate the involvement of RGS12TS-S in nuclear processes that are cell cycle-regulated.

It is important to consider the present results in relation to previous studies of RGS12. Koelle and Horvitz (3) first identified and named a partial expressed sequence tag as RGS12 based upon its homology to the RGS domains of Egl-10, RGS1, and RGS2. Although the identity of this transcript is unknown because it includes only 51 amino acids of the RGS domain, we utilized this sequence to amplify the RGS12 transcripts described here. Snow *et al.* (36) first isolated, by hybridization screening, a rat homolog of the cDNA named RGS12TS-I in the present study. We then reported sequences encoding the long forms of the four N-terminal forms of RGS12 in the GenBank™/EBI Data Bank. Snow *et al.* (37) used a probe to the unique 3′-end of one of the sequences we reported to show that RGS12TS transcripts also exist with this unique 3′-end (*i.e.* corresponding to the long forms described here). In this same report, these workers identified a cDNA encoding a rat homolog of RGS12B and showed that a recombinant fusion protein of the RGS domain of rat RGS12 acted as a GTPase-activating protein *in vitro* for recombinant $G\alpha_i$ and $G\alpha_o$, but not $G\alpha_s$ or $G\alpha_q$. However, we failed to detect $G\alpha_i$ immunoreactivity in RGS12TS-S nuclear dots or changes in plasma membrane-associated $G\alpha_i$ immunoreactivity in EcR293–12TS-GFP cells following inducible expression of RGS12TS-S (data not shown). Mao *et al.* (38) recently showed that transfection of NIH 3T3 cells with rat RGS12TS-I attenuated serum response factor activation mediated by GTPase-deficient forms of G_{12} and G_{13} as well as by receptors acting through both $G_{q/11}$-dependent and -independent pathways. These results raise interesting questions concerning how this effect of RGS12TS-I is mediated, in part due to the lack of GTPase-activating protein activity of the isolated RGS domain of this protein toward G_q (37). Moreover, the ability of rat RGS12 to inhibit responses mediated by GTPase-deficient forms of $G_{12/13}$ suggests this effect is not likely a result of GTPase activation. Although we cannot discount the sequence differences between the rat RGS12TS used in the study by Mao *et al.* (38) or their use of NIH 3T3 cells, our results show that this form of RGS12TS is not measurably expressed in COS-7 cells. In view of our evidence for nuclear expression of RGS12, its interesting to consider a possible direct nuclear role of RGS12 in inhibition of serum response factor-mediated gene transcription, particularly that mediated by GTPase-deficient forms of $G_{12/13}$.

It is still unclear why such a large family of genes encoding RGS proteins exists in man and other species. Here, we have complicated this issue by demonstrating that a single gene in this family can be processed to 12 distinct RGS protein transcripts. Perhaps the observed differential expression of these forms in a heterologous system and their nuclear patterns of expression reflect attributes fundamentally important to the cellular activities of these proteins. Despite the presence of an RGS domain, the apparent constitutive nuclear localization of RGS12 proteins implicates them in activities distinct from regulation of cell-surface G protein-coupled receptor signaling. Hopefully, these results will facilitate studies to define the biological functions of RGS proteins in the nucleus.

REFERENCES

1. Hamm, H. E. (1998) *J. Biol. Chem.* **273,** 669–672
2. Gilman, A. G. (1987) *Annu. Rev. Biochem.* **56,** 615–649
3. Koelle, M. R., and Horvitz, H. R. (1996) *Cell* **84,** 115–125
4. Berman, D. M., and Gilman, A. G. (1998) *J. Biol. Chem.* **273,** 1269–1272
5. Dohlman, H. G., and Thorner, J. (1997) *J. Biol. Chem.* **272,** 3871–3874
6. Berman, D. M., Wilkie, T. M., and Gilman, A. G. (1996) *Cell* **86,** 445–452
7. Berman, D. M., Kozasa, T., and Gilman, A. G. (1996) *J. Biol. Chem.* **271,** 27209–27212
8. Tesmer, J. J. G., Berman, D. M., Gilman, A. G., and Sprang, S. R. (1997) *Cell* **89,** 251–261
9. Coleman, D. E., and Sprang, S. R. (1999) *J. Biol. Chem.* **274,** 16669–16672
10. Hepler, J. R., Berman, D. M., Gilman, A. G., and Kozasa, T. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 428–432
11. Zeng, W., Xu, X., Popov, S., Mukhopadhyay, S., Chidiac, P., Swistok, J., Danho, W., Yagaloff, K. A., Fisher, S. L., Ross, E. M., Muallem, S., and Wilkie, T. M. (1998) *J. Biol. Chem.* **273,** 34687–34690
12. Chatterjee, T. K., Liu, X., Davisson, R. L., and Fisher, R. A. (1997) *J. Biol. Chem.* **272,** 12122–12131
13. Chatterjee, T. K., Eapen, A. K., and Fisher, R. A. (1997) *J. Biol. Chem.* **272,** 15481–15487
14. Chatterjee, T. K., and Fisher, R. A. (2000) *J. Biol. Chem.* **275,** 24013–24021
15. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp, P. A. (1986) *Annu. Rev. Biochem.* **55,** 1119–1150
16. Bonen, L. (1993) *FASEB J.* **7,** 40–46
17. Koken, M. H. M., Reid, A., Quignon, F., Chelbi-Alix, M. K., Davies, J. M., Kabarowski, J. H. S., Zhu, J., Dong, S., Chen, S.-J., Chen, Z., Tan, C. C., Licht, J., Waxman, S., De The, H., and Zelent, A. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 10255–10269
18. Yano, T., Nakamura, T., Blechman, J., Sorio, C., Dang, C. V., Geiger, B., and Canaani, E. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 7286–7291
19. Scully, R., Chen, J., Ochs, R. L., Keegan, K., Hoekstra, M., Feunteun, J., and Livingston, D. M. (1997) *Cell* **90,** 425–435
20. Jin, Y., Xu, X. L., Yang, M.-C. W. Wei, F., Ayi, T.-C., Bowcock, A. M., and Baer, R. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12075–12080
21. Mu, Z. M., Chin, K. V., Liu, J. H., Lozano, G., and Chang, K. S. (1994) *Mol. Cell. Biol.* **14,** 6858–6867
22. Guiochon-Mantel, A., Savouret, J., Quignon, F., Delabre, K., Milgrom, E., and De The, H. (1995) *Mol. Endocrinol.* **9,** 1791–1803
23. Dyck, J. A., Maul, G. G., Miller, W. H., Chen, J. D., Kakizuka, A., and Evans, R. M. (1994) *Cell* **76,** 333–343
24. Weis, K., Rmabaud, S., Lavau, C., Jansen, J., Carvalho, T., Carmo-Fonseca, M., Lamond, A., and Dejean, A. (1994) *Cell* **76,** 345–356
25. Carvalho, T., Seeler, J.-S., Ohman, K., Jordan, P., Pettersson, U., Akusjarvi, G., Carmo-Fonseca, M., and Dejean, A. (1995) *J. Cell Biol.* **131,** 45–56
26. Doucas, V., Ishov, A., Romo, A., Juguilon, H., Weitzman, M., Evans, R., and Maul, G. (1996) *Genes Dev.* **10,** 196–207
27. Sullivan, P. M., Petrusz, P., Szpirer, C., and Joseph, D. R. (1991) *J. Biol. Chem.* **266,** 143–154
28. Vellard, M., Sureau, A., Soret, J., Martinerie, C., and Perbal, B. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 2511–2515
29. Shimizu, A., and Honjo, T. (1993) *FASEB J.* **7,** 149–154
30. Fujieda, S., Lin, Y. Q., Saxon, A., and Zhabg, K. (1996) *J. Immunol.* **157,** 3450–3459
31. Park, J.-G, Muise, A., He, G.-P., Kim, S.-W., and Ro, H.-S. (1999) *EMBO J.* **18,** 4004–4012
32. Gowen, L. C., Avrutskaya, A. V., Latour, A. M., Koller, B. H., and Leadon, S. A. (1998) *Science* **281,** 1009–1012
33. Easton, D. F., Bishop, D. T., Ford, D., and Crockford, G. P. (1993) *Am. J. Hum. Genet.* **52,** 678–701
34. Terada, Y., Tatsuka, M., Suzuki, F., Yasuda, Y., Fujita, S., and Otsu, M. (1998) *EMBO J.* **17,** 667–676
35. Cliby, W. A., Roberts, C. J., Cimprich, K. A., Stringer, C. M., Lamb, J. R., Schreiber, S. L., and Friend, S. H. (1998) *EMBO J.* **17,** 159–169
36. Snow, B. E., Antonio, L., Suggs, S., Gutstein, H. B., and Siderovski, D. P. (1997) *Biochem. Biophys. Res. Commun.* **233,** 770–777
37. Snow, B. E., Hall, R. A., Krumins, A. M., Brothers, G. M., Bouchard, D., Brothers, C. A., Chung, S., Mangion, J., Gilman, A. G., Lefkowitz, R. J., and Siderovski, D. P. (1998) *J. Biol. Chem.* **273,** 17749–17755
38. Mao, J., Yuan, H., Xie, W., Simon, M. I., and Wu, D. (1998) *J. Biol. Chem.* **273,** 27118–27123