Single-nucleotide polymorphisms can cause different structural folds of mRNA

Ling X. Shen*, James P. Basilion, and Vincent P. Stanton, Jr.

Variagenics, Inc., 60 Hampshire Street, Cambridge, MA 02139

ABSTRACT Single-nucleotide polymorphisms (SNPs) are the most common type of genetic variation in man. Genes containing one or more SNPs can give rise to two or more allelic forms of mRNAs. These mRNA variants may possess different biological functions as a result of differences in primary or higher order structures that interact with other cellular components. Here we report the observation of marked differences in mRNA secondary structure associated with SNPs in the coding regions of two human mRNAs: alanyl tRNA synthetase and replication protein A, 70-kDa subunit (RPA70). Enzymatic probing of SNP-containing allelic fragments of the mRNAs revealed pronounced allelic differences in cleavage pattern at sites 14 or 18 nt away from the SNP, suggesting that a single-nucleotide variation can give rise to different mRNA folds. By using phosphorothioate oligodeoxyribonucleotides complementary to the region of different allelic structures in the RPA70 mRNA, but not extending to the SNP itself, we find that the SNP exerts an allele-specific effect on the accessibility of its flanking site in the endogenous human RPA70 mRNA. This further supports the allele-specific structural features identified by enzymatic probing. These results demonstrate the contribution of common genetic variation to structural diversity of mRNA and suggest a broader role than previously thought for the effects of SNPs on mRNA structure and, ultimately, biological function.

Single-nucleotide polymorphisms (SNPs) are single base-pair substitutions that occur within and outside genes (1–3). SNPs account for many well characterized human phenotypes, including disease susceptibility and resistance (4, 5) and drug response (6, 7). As a result, SNP discovery efforts have markedly expanded over the past 2 years. The frequency of SNPs varies between genomic regions and between coding and noncoding sequences. The extent of nucleotide diversity ranges from 0.0003 to 0.005; in other words, there are from 3 to 50 SNPs per 10 kilobases when two chromosomes are compared (for review, see ref. 2).

Genes containing one or more SNPs can give rise to two or more allelic forms of mRNAs. mRNAs containing different bases at SNP sites may vary in their interactions with cellular components involved in mRNA synthesis, maturation, transport, translation, or degradation. It has been documented that a number of single base-pair substitutions alter or create essential sequence elements for splicing, processing, or translation of human mRNA (1). These SNPs are associated with altered length and/or steady-state level of cytoplasmic mRNA. On the other hand, SNPs that do not affect RNA consensus and protein sequences have not been analyzed in detail. It is conceivable that such SNPs could also lead to phenotypic effects, most likely through non-consensus-dependent mechanisms.

A growing body of evidence shows that the folding of mRNA influences a diverse range of biological events such as mRNA splicing (8, 9) and processing (10–12), and translational control (13–16) and regulation (17–19). Because the structure of mRNA is determined by its nucleotide sequence and its environment, we were interested in examining whether the folding of mRNA could be influenced by the presence of SNPs. In this report, we describe the analysis of SNPs in the coding regions of two essential genes: a U/C transition at nucleotide 1,013 of human alanyl tRNA synthetase (AARS, GenBank Accession no. D32050) and a U/C transition at nucleotide 1,674 of human replication protein A, 70-kDa subunit (RPA70, GenBank Accession no. M63488). Neither SNP affects the amino acid sequence of the respective protein. An initial screen of lymphoblast cDNA from 36 unrelated individuals (V.P.S., unpublished results) revealed allele frequencies of 0.49 and 0.51 for the AARS U and C alleles (57% heterozygotes observed) and allele frequencies of 0.85 and 0.15 for the RPA70 U and C alleles (31% heterozygotes observed). Using structural mapping and structure-based targeting strategies, we show that the SNPs have marked effects on the structural folds of the mRNAs. The contribution of common genetic variation to structural diversity of mRNA has not been described previously. These results suggest that phenotypic consequences of SNPs could arise from mechanisms that involve allele-specific structural motifs in mRNA.

MATERIALS AND METHODS

***In Vitro* Synthesis and Purification of mRNA.** Human cDNAs for PCR synthesis of transcription templates were prepared from total RNA of Epstein–Barr virus transformed lymphoblasts (Coriell Cell Repositories, Camden, NJ) or tumor cell line NCI-H292 (American Type Culture Collection). The cDNAs were genotyped for SNPs in human AARS and RPA70 either by sequencing or by restriction fragment length polymorphism analysis. A second round of cDNA sequencing was performed to confirm the nucleotide sequence of AARS and RPA70 cDNAs. Fragments of human mRNA alleles (113- to 1,000-mers) varying in a single base were obtained by using *in vitro* transcription from PCR-derived DNA templates with MEGAshortscript or MEGAscript T7 kits (Ambion, Austin, TX). Crude transcripts were either purified on an 8% denaturing polyacrylamide gel or on a G-50 Sephadex spin column (Boehringer Mannheim), and the purified mRNA was subsequently dephosphorylated by using calf intestinal alkaline phosphatase (Amersham Pharmacia) and 5′-end-labeled with ^{32}P by using [γ-^{32}P]ATP (DuPont/NEN) and T4 polynucleotide kinase (Amersham Pharmacia).

Enzymatic Mapping of Allelic mRNA Structures. Partial digestion of 1 pmol of 5′-^{32}P-labeled mRNA by nuclease S1 or RNase T1 (Amersham Pharmacia) was carried out at 37°C for

---

5 min in 10 μl of buffer containing 10 mM Tris·HCl, pH 7.5/100 mM KCl/5 mM MgCl$_2$/5 μg of phenol-extracted glycogen. Before the addition of S1 or T1, the RNA was renatured in the assaying buffer by heating for 45 sec at 85°C and cooling for 5 min at room temperature followed by 10 min at 37°C. The digestion products were fractionated on either a 4, 6, or 8% denaturing polyacrylamide gel, and the cleavage patterns were visualized by using autoradiography. RNase T1 digestion at 55°C and alkaline hydrolysis at 85°C were done to generate sequence ladders for cleavage site identification. RNase T1 digestion at 55°C was done for 1 min in 15 μl of buffer containing 25 mM sodium acetate, pH 5.2/25 mM KCl/3 μg of yeast tRNA. Alkaline hydrolysis was done for 15 min in 15 μl of buffer containing 50 mM sodium bicarbonate, pH 9.2/1 mM EDTA/10 μg of phenol-extracted glycogen.

***Escherichia coli* RNase H Digestion Assay.** Phosphorothioate oligodeoxyribonucleotides (PS-oligos) were used in our studies because of their enhanced stability against nucleases (20). PS-oligos were chemically synthesized and desalted on C-18 cartridges by Synthetic Genetics (San Diego). One picomole of 5′-^{32}P-labeled mRNA was renatured in 12 μl of RNase H buffer (20 mM Tris·HCl, pH 7.5/10 mM KCl/5 mM MgCl$_2$/0.1 mM DTT/5% glycerol/10 μg of phenol-extracted glycogen) by heating for 1 min at 85 °C, and then cooling for 5 min at room temperature followed by 10 min at 37°C. *E. coli* RNase H (0.6 units, Amersham Pharmacia) was then added to the renatured mRNA and incubated for 5 min at 37°C. PS-oligos (5 pmol in RNase H buffer) were subsequently added, and the reaction mixture (15 μl total) was incubated for 20 min at 37 °C. The digestion products were fractionated on a 4% denaturing polyacrylamide gel, and the cleavage patterns were visualized by autoradiography. RNA levels were quantitated by using a Fluorescent Image Analyzer (Fujifilm Model FLA-2000).

Oligodeoxyribonucleotide-Directed Cleavage of Endogenous RPA70 mRNA in Human Cell Extracts. Human tumor cell lines HeLa and Calu-1 (American Type Culture Collection) were genotyped for SNPs in RPA70 cDNA by restriction fragment length polymorphism analysis and by sequencing. The genotyping results showed that HeLa and Calu-1 express only the U or the C base at the 1,674 U/C polymorphism, respectively. HeLa and Calu-1 cells were cultured in monolayers in minimum essential medium (GIBCO/BRL) and McCoy's 5A medium (Sigma), respectively, under 5%CO$_2$/95% air in a humidified incubator at 37°C. Both media were supplemented with 10% heat-inactivated fetal bovine serum (JRH Biosciences, Lenexa, KS), 100 units/ml penicillin and 100 μg/ml streptomycin (GIBCO/BRL), and 2 mM L-glutamine (GIBCO/BRL). For preparation of cell extracts, the cells were grown to 80–90% confluence, trypsinized, and washed twice with the complete medium as specified above and twice with PBS (GIBCO/BRL). Approximately 5×10^7 cells were lysed in a 15-ml tissue grinder with a Teflon pestle (Wheaton) in the presence of 1.5 ml of lysis buffer (20 mM Tris·HCl, pH 7.5/420 mM NaCl/2 mM MgCl$_2$/0.2 mM EDTA/1 mM DTT/0.5 mM PMSF/25% glycerol). The sodium concentration was chosen to facilitate the disruption of the nuclear membrane to release RNase H from nuclei (21). Up to 200 strokes were needed to completely lyse the cells, as checked by using a hemacytometer. The cell lysate was centrifuged at $16,000 \times g$ for 10 min at 4°C, and cell extracts were obtained by transferring supernatant to clean microcentrifuge tubes and freezing immediately at -80 °C. Total protein concentration of cell extracts was determined by using the Bradford assay (Bio-Rad). For PS-oligo-directed cleavage of endogenous mRNA reactions, cell extracts were diluted 3- to 4-fold with the lysis buffer to normalize the total protein concentration to 3 mg/ml for the two cell types. A 10× PS-oligo solution was subsequently added to the normalized cell extracts, and the reaction mixture was incubated for 40 min at 37°C. The incubation time and oligo concentration range were chosen to maintain the extent of mRNA cleavage below

100% for subsequent quantitation of intact mRNA. The level of intact RPA70 mRNA was determined by total RNA isolation and Northern blot hybridization by using procedures described previously (22, 23). Blots were analyzed with a Fluorescent Image Analyzer (Fujifilm Model FLA-2000). Northern probes specific for RPA70 mRNA were synthesized by random primed [α-^{32}P]dCTP labeling of a cDNA probe corresponding to nucleotides 1,519–2,080 of human RPA70 mRNA. To normalize the amount of total RNA loaded into individual gel wells for Northern analysis, the level of glyceraldehyde 3-phosphate dehydrogenase mRNA was also measured by first stripping the RPA70 mRNA probe and then reprobing the same Northern blot with a random-primed [α-^{32}P]dCTP-labeled cDNA probe corresponding to a 600-nt sequence of human glyceraldehyde 3-phosphate dehydrogenase (Stratagene).

RESULTS AND DISCUSSION

Characterization of SNP-Containing mRNA Structures. Two SNPs in the coding regions of two human mRNAs were analyzed in this study: a U/C transition at nucleotide 1,013 of AARS (GenBank Accession no. D32050) and a U/C transition at nucleotide 1,674 of RPA70 (GenBank Accession no. M63488). To investigate effects of these SNPs on mRNA structure, we first synthesized both allelic SNP-containing fragments of the two mRNAs by using *in vitro* transcription and then performed mRNA structural mapping by using structure-specific nuclease S1 and RNase T1.

Nuclease S1 cleaves both DNA and RNA in single-stranded regions (24). It is a useful structural probe for identifying large unpaired accessible regions (>4 nt) in RNA. Fig. 1*a* shows nuclease S1 mapping of a 113-nt fragment of AARS mRNA (nucleotides 951–1,063). Differences in S1 cleavage patterns were observed between the U and C alleles. Strong cleavages centered at nucleotide 999 of the U allele revealed the presence of a single-stranded region $\approx$14 nt upstream of the SNP. The majority of the species of the C allele lack such a single-stranded region, as shown by the absence of strong cleavage bands in the same region. Allelic differences in S1 accessibility were also observed at several other sites, including regions 980 and 1,032 (Fig. 1*a*). At higher S1 concentrations, both alleles exhibited similar cleavage patterns at sites 980 and 1,032, suggesting that the stronger cleavage intensities at these sites in the C allele may be due partly to the absence of a large unpaired region at position 999. Structural mapping by RNase T1 (data not shown), which identifies unpaired guanosine residues (24), corroborated the finding of substantial allelic structural differences within ±14 nt of the SNP. The enzymatic mapping experiments were also done on a 1,000-nt fragment (nucleotides 529–1,528) of AARS mRNA. Differential S1 digestion patterns of the two alleles identical to those at sites 980, 999, and 1,032 in Fig. 1*a* were observed with the 1,000-mer mRNA sequence (data not shown).

Pronounced allelic structural differences caused by a single-nucleotide variation in the mRNA sequence were also detected in a 544-nt fragment of RPA70 mRNA (Fig. 1*b*). The fragment, consisting of nucleotides 1,378–1,921, was centered about the 1,674 U/C polymorphism. Intense S1 cleavages were detected in the U allele around nucleotide 1,656, indicating the presence of a single-stranded region $\approx$18 nt upstream of the SNP. In contrast, no cleavage above background intensity was detected at the same site of the C allele. The allelic difference at this site was confirmed by structural probing with RNase T1 (data not shown). Differential S1 cleavage between the two alleles was also observed at nucleotide 1,664, where the U allele was more weakly cleaved than the C allele. These data thus defined a differential-structure motif between the two allelic forms of RPA70 mRNA, with a maximum allelic difference in structural accessibility in the region of nucleotide 1,656. In addition,
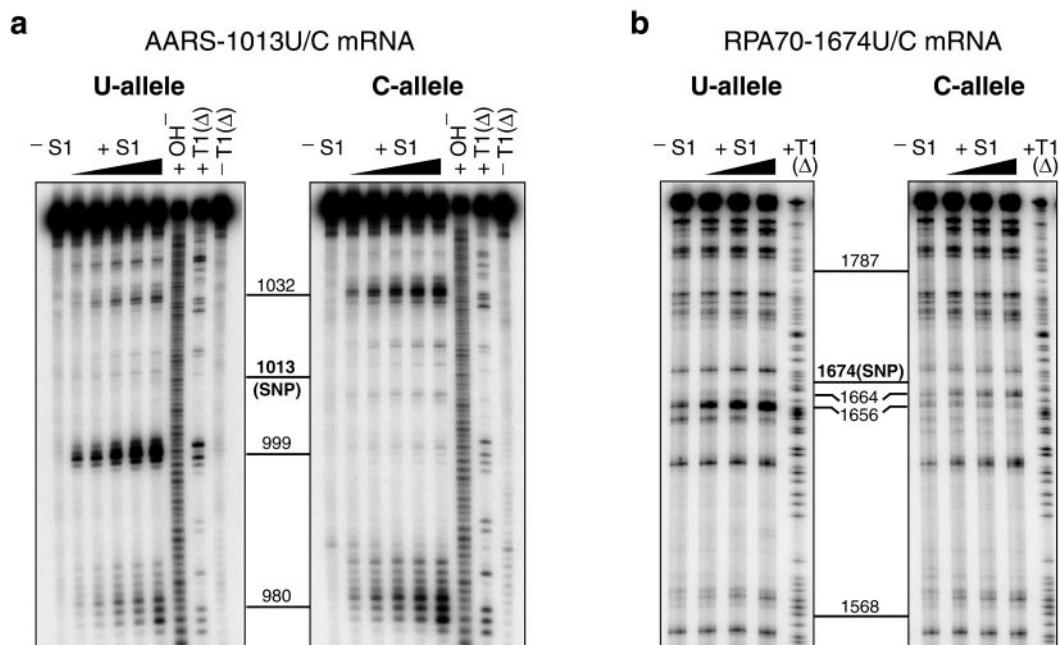
FIG. 1.   Nuclease S1 mapping of allelic mRNA structures. −S1, control reaction without S1 added; +S1, partial digestion by S1 at 37°C with increasing S1 concentration; +OH⁻, alkaline hydrolysis reaction; +T1(Δ), partial digestion by T1 at 55°C to create a G ladder to allow determination of cleavage sites; −T1(Δ), control reaction without T1 added. (*a*) Mapping of accessible single-stranded regions of the U and C alleles of a 113-mer fragment (nucleotides 951–1,063) of AARS mRNA. (*b*) Mapping of accessible single-stranded regions of the U and C alleles of a 544-mer fragment (nucleotides 1,378–1,921) of RPA70 mRNA. Polymorphic positions are indicated in boldface and labeled SNP.

the structural mapping data on both the RPA70 and the AARS mRNAs revealed that the polymorphic sites were relatively inaccessible by single-strand-specific nuclease S1. Proposed secondary-structure models of the polymorphic regions of the two pairs of mRNA alleles are shown in Fig. 2.

Because the SNP-dependent allelic structural differences described above were identified by using mRNA fragments in a purified system, it is critical to examine whether our findings of allele-specific structural features exist in the context of full-length mRNA in an environment better mimicking the physiological conditions in cells. To address this issue, we used a structure-based mRNA targeting strategy to monitor allelic structural accessibility of RPA70 mRNA in two different assay systems: a purified system containing mRNA fragments and a cell-mimicking system containing endogenous human RPA70 mRNA. We designed a series of PS-oligos complementary to the 1,656 region of RPA70 mRNA that contains the maximum allelic difference in structural accessibility as identified by enzymatic mapping. The target site at the region of 1,656 was chosen to span 20 or 15 nt that do not include the 1,674 U/C polymorphism. Thus, the target site was identical in nucleotide sequence between the U and C alleles and was predicted based on enzymatic mapping results to be accessible for PS-oligo binding in the U allele, and inaccessible in the C allele. Exogenous or endogenous RNase H, which cleaves RNA in RNA·DNA hybrids (25), was used to monitor the targeting efficiency and allele specificity of PS-oligos for the two alleles of RPA70 mRNA.

Allele-Specific Cleavage of RPA70 mRNA in a Purified System. We first used an *in vitro* assay with *E. coli* RNase H and the *in vitro*-transcribed 544-mer RPA70 mRNA fragments to examine the efficiency and allele specificity of structure-based targeting as compared with directly targeting the 1,674 U/C polymorphism. Fig. 3*a* shows nucleotide sequences of the PS-oligos designed to target either the RPA70 polymorphism (oligo **1**, U allele-specific; oligo **2**, C allele-specific) or the site of maximum differential structural accessibility in the 1,656 region (oligos **3–6**). SNP-targeting oligos **1** and **2** were previously optimized for allele discrimination in a cell transfection

assay (31). Each of the SNP-targeting oligos is perfectly matched only with one of the two alleles; it is singly mismatched with the other allele. Single-base mismatches formed between the SNP-targeting oligos and their respective nontargeted alleles contribute to the allele discrimination observed previously. Structure-targeting oligos **3** and **6** match perfectly with both alleles, and were designed to hybridize maximally with the single-stranded 1,656 region of the U allele, while avoiding the 1,664 region, which is slightly accessible in both alleles (Fig. 1*b*). Four and six additional nucleotides were added to the 5′-end of oligo **3** to produce oligos **4** and **5**, respectively, to facilitate folding into hairpin structures for enhanced allele discrimination (26).

Results of oligo-directed, site-specific cleavage of the 544-mer U and C alleles by *E. coli* RNase H are shown in Fig. 3 *b–c*. As predicted, the U allele that was more sensitive to S1 cleavage than the C allele in the 1,656 region was over 10-fold more efficiently cleaved by *E. coli* RNase H in the presence of structure-targeting PS-oligos **3–6**. Although the cleavage efficiencies with PS-oligos **1** and **2** that target the SNP at position 1,674 are greater for their respective matched alleles than for the mismatched alleles, allele discrimination cannot be accurately assessed because of uncertainties in the measurement of cleavage efficiencies below 1%. The polymorphic site targeted by oligos **1** and **2** was much less efficiently cleaved than the single-stranded 1,656 region in the U-allele targeted by oligos **3–6**. This is consistent with the difference in accessibility between the two sites as identified by S1 mapping.

Oligos **4** and **5** with predicted hairpin structures exhibited a 23% or 36% reduction, respectively, in the cleavage of both alleles as compared with their parent oligo **3**. Therefore, no enhancement in allele discrimination was obtained with oligos **4** and **5** under the experimental conditions.

Allele-Specific Cleavage of Endogenous RPA70 mRNA in Human Cell Extracts. We next performed oligo-directed cleavage of the full-length, endogenous RPA70 mRNA in whole-cell extracts prepared from two genotyped human cell lines, HeLa (1,674 U) and Calu-1 (1,674 C). Cell lysis conditions were chosen to allow the disruption of both the plasma
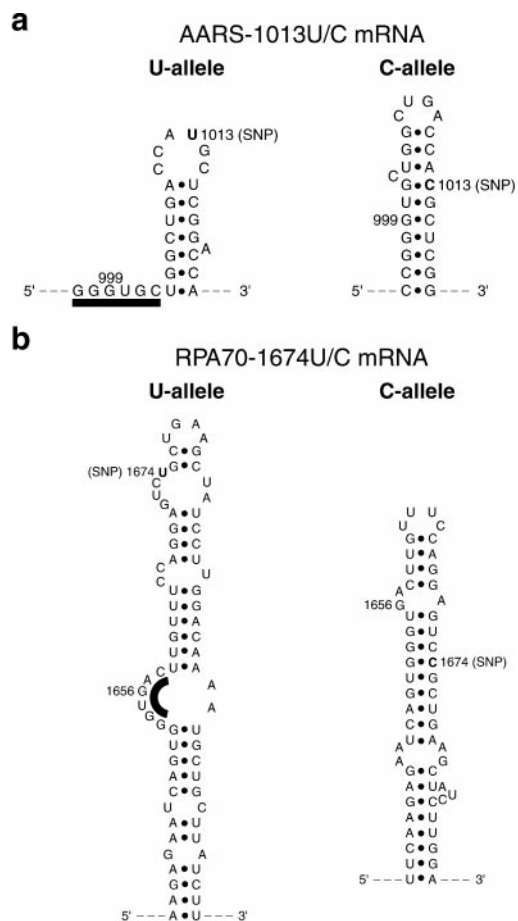
FIG. 2. Proposed secondary-structure models of the polymorphic regions of the pairs of mRNA alleles. The polymorphic bases are shown boldfaced and labeled SNP. The model building was assisted by the use of computer program MFOLD (M. Zuker, Washington University, St. Louis, MO). Only modeled regions supported by S1 and T1 mapping data are shown. (*a*) Proposed allelic structures of the polymorphic region of AARS mRNA. (*b*) Proposed allelic structures of the polymorphic region of RPA70 mRNA. The black bars highlight the highly accessible single-stranded regions.

FIG. 3. *E. coli* RNase H digestion of both alleles of the 544-mer fragment of the RPA70–1,674 U/C mRNA. (*a*) Nucleotide sequence of the PS-oligos used in the assay. Oligos **1** and **2** target the polymorphic site; the bases opposing the polymorphism in the mRNA are shown in boldface. Oligos **3–6** target the site of different allelic structures around nucleotide 1,656. The terminal residues of oligos **1**, **2**, **3**, and **6** are labeled according to the positions of the complementary nucleotides in RPA70 mRNA. Oligos **4** and **5** share the same target hybridization sequence as **3**, but with additional nucleotides (shown underlined) added to facilitate folding into hairpin structures for intended enhancement in allele discrimination. (*b*) *E. coli* RNase H digestion patterns of the 5′-^{32}P-labeled U and C alleles hybridized to the oligos in *a*. Cleavage sites were confirmed by running on the same gel RNase T1 digestion ladders for the two alleles (data not shown): they are centered at position 1,674 for oligos **1** and **2**, 1,650 for oligos **3–5**, and 1,654 for oligo **6**. (*c*) Comparison of percent cleavages of the U and C alleles calculated from *b*. Oligo-dependent cleavage bands were integrated as a percentage of total RNA in each lane with the overlapping background-cleavage bands subtracted. Variations in sample loading to each lane were internally corrected by measuring percent cleavage instead of absolute band intensity.

and the nuclear membranes in the absence of denaturants deleterious to the integrity of cellular proteins and nucleic acids. To control for cellular component-associated, nonspecific effects of PS-oligos in the two different cell types, we introduced an RPA70-specific but non-allele-discriminating 20-mer PS-oligo **7** to the assay to serve as a positive control, in addition to the use of SNP-targeting oligos **1** and **2**, which serve as reciprocal negative controls. Oligo **7** is complementary to the 2,230–2,249 sequence region of the 3′-untranslated region of human RPA70 mRNA; it was previously shown to target the 1,674 U and 1,674 C alleles with equally high efficiency in a cell transfection assay (31).

Fig. 4 compares the extent of site-specific cleavage of the 1,674 U and 1,674 C alleles of the endogenous human RPA70 mRNA in the presence of 100, 300, and 600 nM of either U or C allele-specific oligos **1**, **2**, **3**, or **6**, or oligo **7**, which targets both alleles. Both the full-length and one or two cleavage products were detected by using Northern blot hybridization. Because no *E. coli* RNase H was added to the reaction mixture, the site-specific cleavage of the mRNA is attributable to the presence of endogenous human RNase H in the cell extracts. Quantitative analysis of the extent of mRNA cleavage was done by measuring the full-length mRNA signal instead of the cleavage products. The latter did not hybridize with probes to the same extent because of probe design and were also likely
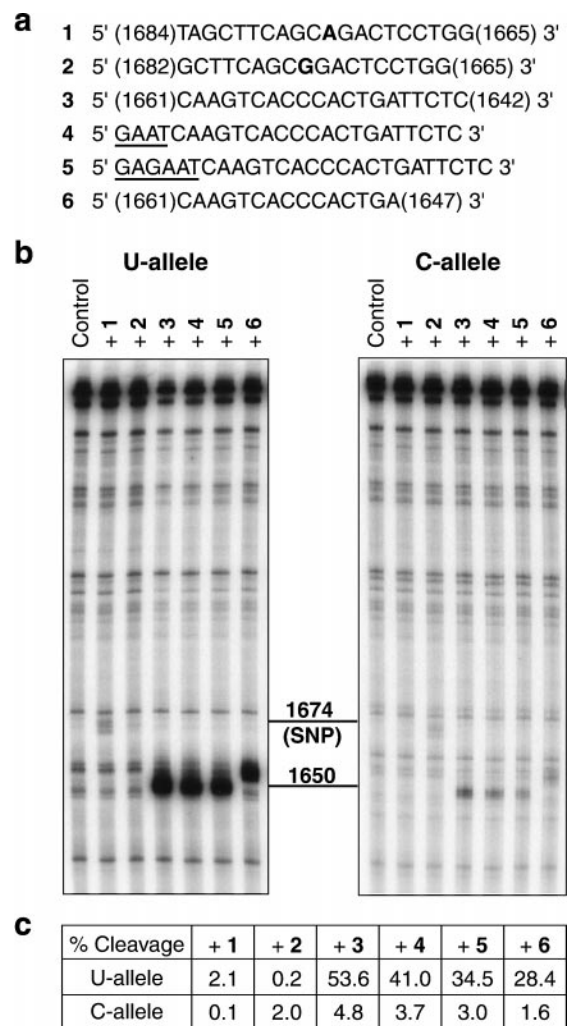
more susceptible to inhomogeneous degradation by RNases in the cell extracts.

The two cell types responded similarly to the addition of control oligo **7** over a concentration range of 100–600 nM (Fig. 4). In contrast, oligos **1**, **2**, **3**, and **6** displayed varying degrees of targeting efficiency and allele specificity qualitatively similar to the results obtained in the purified system. In the presence of 300 nM of 20-mer **1** or 18-mer **2** targeting the 1,674 U/C polymorphic site, the majority of the matched allele remained uncleaved, and the matched allele was equally or slightly more cleaved than the singly mismatched allele (Fig.
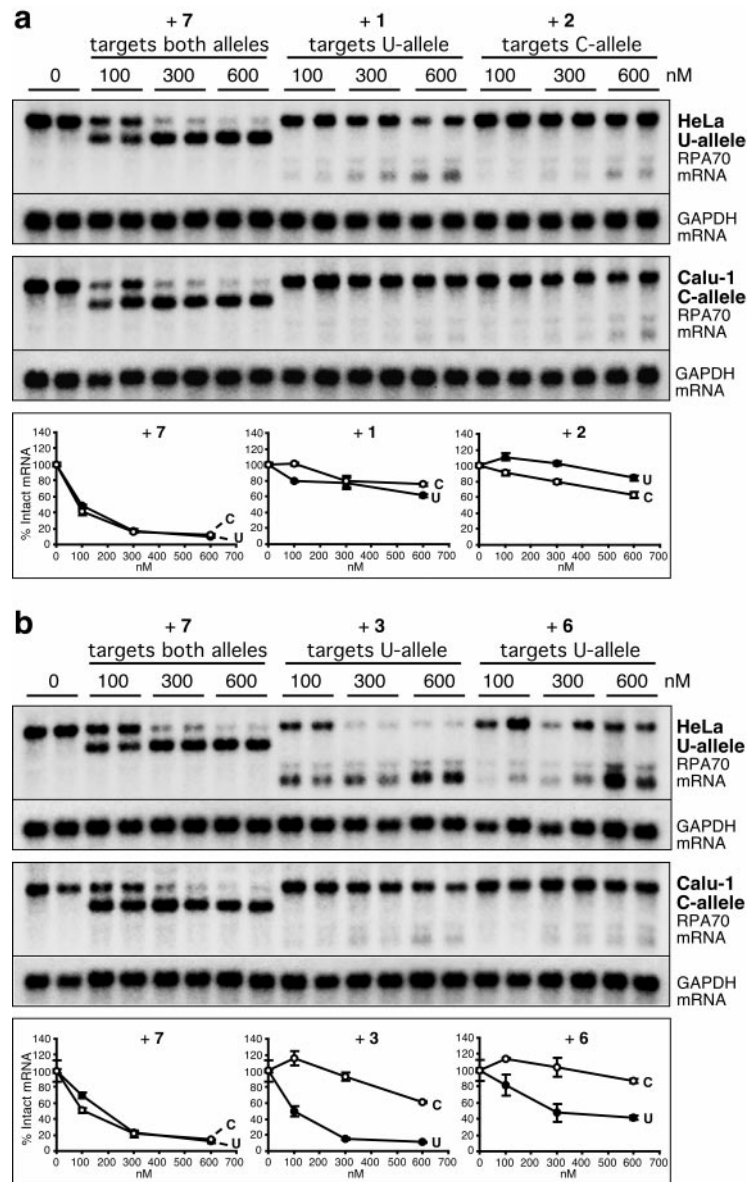
FIG. 4. Northern blot analysis of PS-oligo-directed cleavage of endogenous RPA70 mRNA in human cell extracts. The extracts were prepared from HeLa and Calu-1 cells that express only the 1,674 U or the 1,674 C allele of RPA70 mRNA, respectively. In addition to PS-oligos **1**, **2**, **3**, and **6** as described in Fig. 3, PS-oligo **7** (5′-TGGTCTGCAGTTAGGGTCAG-3′) was used in the assay as a non-allele-discriminating control oligo that targets nucleotides 2,230–2,249 of the 3′-untranslated region of RPA70 mRNA. (*a*) Targeting the 1,674 U/C polymorphic site by using oligos **1** and **2**. (*b*) Targeting the site of different allelic structures around nucleotide 1,656 by using oligos **3** and **6**. Reactions were done in duplicate. The averaged data points of the duplicates are plotted with standard deviations as error bars and are shown below the corresponding Northern blots. Intact RPA70 mRNA levels were normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) levels and expressed as a percentage of no-oligo control for each experiment.

4*a*). The lack of efficient cleavage with **1** and **2** can be explained by the inaccessibility of the 1,674 U/C polymorphic site as identified by using *in vitro* structural mapping. On the other hand, in the presence of 300 nM 20-mer **3**, which targets the single-stranded structural motif of the U allele, only 15% of the U allele remained uncleaved, whereas >90% of the nontarget C allele was intact (Fig. 4*b*). Such high efficiency and specificity of U allele targeting by **3** is consistent with an accessible single-stranded structure in the 1,656 region of the U allele and an inaccessible structure in the same region of the C allele (see Figs. 1*b* and 2*b*). Based on the data in Fig. 4, we conclude that under the experimental conditions, the allele discrimination obtained by targeting the different allelic structures of RPA70 mRNA with 20-mer **3** is 2- to 6-fold greater than that obtained by directly targeting the polymorphism with 20-mer **1** when the percentage of intact mRNA is evaluated. The 15-mer oligo **6**

also showed better allele discrimination than SNP-targeting oligos **1** and **2** at all three oligo concentrations, albeit less prominent than that of 20-mer **3**.

It is important to note that structure-targeting oligos **3** and **6** are completely complementary to their target site sequences, because the target sites do not contain the polymorphic bases. Thus, the pronounced allele discrimination by these structure-targeting oligos reflects a significant difference in target-site accessibility between the two alleles. This accessibility difference could be caused by either different structural folds of the two alleles of endogenous RPA70 mRNA, different contacts with cellular components, or both effects. The positive correlation we have observed between the results in the purified system and those in cell extracts strongly supports the argument that the pronounced allele discrimination obtained with the structure-targeting oligos is caused mainly by structural

differences between the two endogenous RPA70 mRNA alleles.

SNPs are the most abundant type of human DNA sequence variation. The fraction of SNPs associated with detectable differences in mRNA folding is yet to be discovered. Statistical studies of RNA secondary structures by Fontana *et al.* (27) predicted that structures derived from GC- or AU-rich sequences are much more sensitive to point mutation than those from sequences with uniform base distributions. The GC content of the 100 nucleotides flanking the AARS and RPA70 SNPs is 63% and 40%, respectively. However, data from the two examples described here are insufficient to provide a statistically significant assessment of the determinants of SNP-associated structural changes.

In conclusion, we have demonstrated that SNPs can cause significant structural variation in allelic forms of human mRNA. The structural differences occur both at and in regions flanking the SNP. The structural differences allow allele-specific cleavage of mRNA with PS-oligos complementary to a site of maximum differential structural accessibility that is independent of single-base discrimination at the polymorphic site. These findings illustrate how SNPs can affect mRNA structure and suggest mRNA-structure-dependent mechanisms by which SNPs can cause allele-specific biological consequences. Further elucidation of the contribution of common genetic variation to structural and functional diversity of mRNA will provide insight into fundamental mechanisms of human phenotypic variation and facilitate studies of disease susceptibility and drug response (4–7). The results described here should also be heeded by researchers interested in RNA as a therapeutic target; the possible effect of genetic variation on target heterogeneity should be considered for target characterization and drug design. Finally, structure-based mRNA targeting may be applied therapeutically to diseases where allele-specific RNA down-regulation would provide a therapeutic benefit. Nucleic acid-based inhibitors could be used for such purposes, however, small-molecule inhibitors (28–30) have more attractive pharmaceutical properties, including enhanced cellular uptake and possibly improved allele discrimination.

1. Cooper, D. N., Krawczak, M. & Antonarakis, S. E. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, eds. Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D. (McGraw–Hill, New York), Vol. 1, pp. 259–291.
2. Chakravarti, A. (1999) *Nat. Genet.* **21**, Suppl., 56–60.
3. Wang, D. G., Fan, J.-B., Siao, C.-J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., *et al.* (1998) *Science* **280,** 1077–1082.
4. Lander, E. S. (1996) *Science* **274,** 536–539.
5. Collins, F. S., Guyer, M. S. & Chakravarti, A. (1997) *Science* **278,** 1580–1581.
6. Weber, W. W. (1997) *Pharmacogenetics* (Oxford Univ. Press, New York).
7. Kleyn, P. W. & Vesell, E. S. (1998) *Science* **281,** 1820–1821.
8. Coleman, T. P. & Roesser, J. R. (1998) *Biochemistry* **37,** 15941–15950.
9. Côté, J. & Chabot, B. (1997) *RNA* **3,** 1248–1261.
10. van Gelder, C. W. G., Gunderson, S. I., Jansen, E. J. R., Boelens, W. C., Polycarpou-Schwarz, M., Mattaj, I. W. & van Venrooij, W. J. (1993) *EMBO J.* **12,** 5191–5200.
11. Allain, F. H.-T., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D. & Varani, G. (1996) *Nature (London)* **380,** 646–650.
12. Vasserot, A. P., Schaufele, F. J. & Birnstiel, M. L. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 4345–4349.
13. Pelletier, J. & Sonenberg, N. (1987) *Biochem. Cell Biol.* **65,** 576–581.
14. Wang, C., Sarnow, P. & Siddiqui, A. (1994) *J. Virol.* **68,** 7301–7307.
15. ten Dam, E. B., Pleij, C. W. A. & Bosch, L. (1990) *Virus Genes* **4,** 121–136.
16. Shen, L. X. & Tinoco, I., Jr. (1995) *J. Mol. Biol.* **247,** 963–978.
17. Hentze, M. W., Caughman, S. W., Rouault, T. A., Barriocanal, J. G., Dancis, A., Harford, J. B. & Klausner, R. D. (1987) *Science* **238,** 1570–1573.
18. Casey, J. L., Koeller, D. M., Ramin, V. C., Klausner, R. D. & Harford, J. B. (1989) *EMBO J.* **8,** 3693–3699.
19. Addess, K. J., Basilion, J. P., Klausner, R. D., Rouault, T. A. & Pardi, A. (1997) *J. Mol. Biol.* **274,** 72–83.
20. Shaw, J.-P., Kent, K., Bird, J., Fishback, J. & Froehler, B. (1991) *Nucleic Acids Res.* **19,** 747–750.
21. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. (1983) *Nucleic Acids Res.* **11,** 1475–1489.
22. Peppel, K. & Baglioni, C. (1990) *BioTechniques* **9,** 711–713.
23. Brown, T. & Mackey, K. (1997) in *Current Protocols in Molecular Biology*, eds. Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J. & Struhl, K. (Wiley, New York), Vol. 1, pp. 4.9.1–4.9.16.
24. Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P. & Ehresmann, B. (1987) *Nucleic Acids Res.* **15,** 9109–9128.
25. Hostomsky, Z., Hostomska, Z. & Matthews, D. A. (1993) in *Nucleases*, eds. Linn, S. M., Lloyd, R. S. & Roberts, R. J. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 341–376.
26. Tyagi, S., Bratu, D. P. & Kramer, F. R. (1998) *Nat. Biotechnol.* **16,** 49–53.
27. Fontana, W., Konings, D. A. M., Stadler, P. F. & Schuster, P. (1993) *Biopolymers* **33,** 1389–1404.
28. Werstuck, G. & Green, M. R. (1998) *Science* **282,** 296–298.
29. Mei, H.-Y., Cui, M., Heldsinger, A., Lemrow, S. M., Loo, J. A., Sannes-Lowery, K. A., Sharmeen, L. & Czarnik, A. W. (1998) *Biochemistry* **37,** 14204–14212.
30. Hamy, F., Felder, E. R., Heizmann, G., Lazdins, J., Aboul-ela, F., Varani, G., Karn, J. & Klimkait, T. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 3548–3553.
31. Basilion, J. P., Schievella, A. R., Burns, E., Rioux, P., Olson, J. C., Monia, B. P., Lemonidis, K. M., Stanton, V. P., Jr., & Housman, D. E. (1999) *Mol. Pharmacol.*, in press.