Tracking COL1A1 RNA in Osteogenesis Imperfecta: Splice-defective Transcripts Initiate Transport from the Gene but Are Retained within the SC35 Domain

Carol Johnson,* Dragana Primorac,‡ Monique McKinstry,§ John McNeil,* David Rowe,‡ and Jeanne Bentley Lawrence*

*Department of Cell Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01655; and ‡Department of Pediatrics, University of Connecticut Health Center, Farmington, Connecticut 06030

Abstract. This study illuminates the intra-nuclear fate of COL1A1 RNA in osteogenesis imperfecta (OI) Type I. Patient fibroblasts were shown to carry a heterozygous defect in splicing of intron 26, blocking mRNA export. Both the normal and mutant allele associated with a nuclear RNA track, a localized accumulation of post-transcriptional RNA emanating to one side of the gene. Both tracks had slightly elongated or globular morphology, but mutant tracks were cytologically distinct in that they lacked the normal polar distribution of intron 26. Normal COL1A1 RNA tracks distribute throughout an SC-35 domain, from the gene at the periphery. Normally, almost all 50 COL1A1 introns are spliced at or adjacent to the gene, before mRNA transits thru the domain. Normal COL1A1 transcripts may undergo maturation needed for export within the domain such as removal of a slow-splicing intron (shown for intron 24), after which they may disperse. Splice-defective transcripts still distribute thru the SC-35 domain, moving ~1–3 μm from the gene. However, microfluorimetric analyses demonstrate mutant transcripts accumulate to abnormal levels within the track and domain. Hence, mutant transcripts initiate transport from the gene, but are impeded in exit from the SC-35 domain. This identifies a previously undefined step in mRNA export, involving movement through an SC-35 domain. A model is presented in which maturation and release for export of COL1A1 mRNAs is linked to rapid cycling of metabolic complexes within the splicing factor domain, adjacent to the gene. This paradigm may apply to SC-35 domains more generally, which we suggest may be nucleated at sites of high demand and comprise factors being actively used to facilitate expression of associated loci.

Key words: osteogenesis imperfecta • RNA transport • RNA splicing • nuclear structure • collagen

Introduction

Knowledge of how mRNA transcription, processing and selective export are integrated into the dense structure of the interphase nucleus is important for understanding both normal mRNA metabolism and defects in many genetic diseases. Work presented here investigates the interrelationship of RNA metabolism and nuclear structure in normal and mutant cells, by examining the fate of COL1A1 transcripts in osteogenesis imperfecta (OI)1 Type I where a heterozygous mutation impacts the splicing and, as shown here, the transport of this mRNA. In so doing, this study also illuminates two fundamental but debated aspects of nuclear structure, concerning the relationship of SC-35 domains to pre-mRNA metabolism, and the existence of posttranscriptional RNA tracks.

OI is caused by mutations in the collagen type I genes that result in a wide range of phenotypes (Byers and Steiner, 1992). mutant RNA studied here from a patient with OI type I (patient 054) was found to carry a G-A transition in the splice site of intron 26, identical to a mutation reported earlier from another OI Type I patient (053) (Stover et al., 1993). The mutant RNA retains intron 26, which also introduces a frame-shift and downstream nonsense codons. The mutant transcripts are detected in the nucleus but not the cytoplasm (this study and Genovese et al., 1989; Genovese and Rowe, 1987). In the heterozygous state this mutation results in lower levels of Type I collagen and a mild disorder characterized by nondeform-
ing recurrent bone fractures, typical of OI Type I (Byers and Steiner, 1992). However, other mutants in which defective COL1A1 RNA is expressed in the cytoplasm can cause a much more severe (OI Type III) or lethal (OI Type III) disease (reviewed in Howe, 1991; Kuvianiemi et al., 1991; Byers and Steiner, 1992). Hence, in OI, as in many diseases, the question of whether mutant RNA will be transported from the nucleus and expressed in the cytoplasm is critical to the clinical outcome.

Nuclear retention of a defective RNA could result from either a lack of factors that promote export or the presence of factors that prevent nuclear release of transcripts, or both (reviewed in: Goldfarb, 1989; E1liott et al., 1994; Izaurralde and Mattaj, 1995; Nakielny et al., 1997; see also J armolowski et al., 1994). A connection between pre-mRNA splicing and nuclear-cytoplasmic transport had been suggested previously (Legrain and Rosbash, 1989; Chang and Sharp, 1989), and a very recent study provides detailed demonstration of such a link (Luo and Reed, 1999). Hence, retention of intron 26 might disrupt nuclear transport of mutant transcripts. It was also possible that this COL1A1 mutant RNA would be rapidly degraded due to non-sense mediated decay which studies of some mammalian mRNAs suggested previously (Legrain and Rosbash, 1989; Xing et al., 1993). Recent findings from several labs have heightened interest in discrete domains of abundant metabolic factors. While transcription clearly occurs throughout the nucleoplasm (previously cited references and Jackson et al., 1993), statistical analysis of ~20 endogenous genes/RNAs has shown that a substantial subset of genes expressing pre-mRNA localize with high frequency (70–99%) in the periphery of an SC-35–rich domain (Xing et al., 1993, 1995; Moen et al., 1995a; Smith et al., 1999). One interpretation might be that the accumulated factors were recruited from distant storage sites to a randomly located active gene. However, a variety of observations led us to suggest a more complex structural arrangement, whereby some active chromosomal loci are intimately associated with domains of abundant metabolic factors (Lawrence et al., 1993; Smith et al., 1999). This hypothesis is not in itself inconsistent with the possibility of the domain interior being devoid of new transcripts, especially since many (pre)-mRNAs that associate with domains are detected at the periphery, not necessarily throughout the domain (e.g., Xing et al., 1993).

A more puzzling case is presented by several individual (pre)-mRNAs, such as COL1A1, which are consistently seen throughout the interior of an SC-35–domain (Xing et al., 1995; Smith et al., 1999), as well as by the observation of adenovirus transcripts within the IGC (Bridge et al., 1996). A fundamental question then becomes, are some fluorescent domains purely stores of inert splicing factors and other accumulated factors on nascent transcripts of a highly expressed gene? Or, might individual SC-35–domains contain both a disproportionate excess of splicing factors as well as specific newly synthesized RNA species? For COL1A1, the idea that transcripts might move from a gene positioned at the periphery into the SC-35 domain is closely related to another debated concept, regarding postranscriptional RNA tracks. We use the term RNA track to denote a localized accumulation of detached transcripts that extends beyond and to one side of the gene, illuminating an early step in the RNA export path. The key point is that an RNA track contains detached transcripts in transit from the gene, as distinct from a tree of nascent transcripts still attached to the gene. The term track is not intended to imply a prefixed linear or vectorial structure along which transcripts must travel (analogous to a train track). Tracks seen for different RNA species in different cells may have quite distinct structural bases; for example tracks for several cellular mRNAs are less elongated and tightly localized than tracks of some viral RNAs (Lawrence et al., 1989; Xing et al., 1995; Lampel et al., 1997; Melck et al., 2000). Though not seen for all (pre)-mRNAs (discussed in...
Xing and Lawrence, 1993), RNA tracks extending from one side of the gene, hypothesized to contain posttranscriptional transcripts, have now been reported from multiple labs for several cellular and viral RNA(s) (previously cited references and Xing et al., 1993; Huang and Spector, 1991; Rap et al., 1991). In contrast, for other RNA(s) only random dispersal through the nucleolus was observed (e.g., Zachar et al., 1993; Politz et al., 1998; Singh et al., 1999; reviewed in D aneholt, 1999). Consequently, despite major differences in the RNAs and cell-types studied, there has remained debate as to whether putative posttranscriptional tracks of released transcripts are really just Christmas trees of nascent transcripts attached to a highly decondensed active gene (e.g., Zachar et al., 1993; K ramer et al., 1994).

Here detailed investigations of normal and splice-defective COL1A1 transcripts define the intranuclear fate of mutant transcripts in OI patient cells, but also provide a critical test of the interpretation of RNA tracks. To extend our analyses we devised a strategy to examine the cumulative distribution of the 50 introns in COL1A1 RNA, providing a fuller understanding of the relationship of both the RNA tracks and SC-35 domains to the splicing and transport of normal and mutant COL1A1 RNA.

Materials and Methods

Cell Culture, Fractionation, and Immunofluorescence

Fibroblast cultures were previously established from dermal biopsies of A E, an 11-year-old male (cell line 054), and a 13-year-old female (cell line 053; Rowe et al., 1985). These and WI38 human diploid lung fibroblasts (ATCC) were grown in DM E, high glucose with 10–20% FCS, 2 mM l- Glut, and 1% Pen/Strep. Due to senescence of the primary cell cultures, we are currently attempting to obtain more patient 054 and 053 fibroblasts.

Nuclear and cytoplasmic fractions were prepared as described previously for analysis of mRNA in patient fibroblasts (Genovese and Rowe, 1987; Stover et al., 1993). In brief, cultures were scraped into 10 ml of cold PBS, centrifuged, and resuspended in 15 ml of lysis buffer (10 mM Tris, pH 7.5, 10 mM KC l, 1.5 mM MgCl2) containing 0.25% Triton X-100. The pellet was disrupted by Dounce homogenization, and the suspension centrifuged for 10 min at 2,200 g. The supernatant was disrupted by Dounce homogenization, and the suspension centrifuged for 10 min at 2,200 g. The supernatant was then reextracted in lysis buffer with homogenization, and following centrifugation combined with the first extract. The nuclear pellet was re-suspended in 1x extraction buffer and subsequent steps of RNA isolation followed the SDS proteinase K protocol (Genovese et al., 1984).

RNase Protection and DNA Sequencing

Indirect RNase protection was done by hybridizing nuclear mRNA with antisense sscDNA transcribed from H F404, a probe specifically designed to protect RNA from exons 19–43 (Genovese et al., 1989). A filter RNase A digestion the RNA fragments were electrophoresed in 1% agarose with formaldehyde, blot ted onto nitrocellulose, and hybridized (Genovese et al., 1989). Control cells were incubated in DRM for 0, 6, 9, and 12 h followed by extraction of total RNA. Three plates of confluent cells were used for each time point. COL1A1 mRNA levels were examined by RNA protection and quantitated with a Molecular Dynamics PhosphorImager. Multiple plates of 054 cells were incubated in 25 mM DRM for 0, 4.5, and 9 h.

Fluorescence In Situ Hybridization and Immunofluorescence

Cells were prepared for fluorescence in situ hybridization (FISH) as previously described (Lawrence et al., 1989; Carter et al., 1991; Johnson et al., 1991). In brief, cells were washed in cytoskeletal buffer (CSK; F ey et al., 1986) extracted with 0.5% Triton X-100, and fixed in 4% paraformaldehyde 1x PBS (pH 7.4) for 8 min. Human collagen 1A1 DNA clones included cosm id C G103 containing the entire 18-kb human gene (Barsh et al., 1994), a 3' 1.8-kb partial cDNA clone, pH677 (ATCC), and a 5.2-kb full-length cDNA probe (ATCC). COL1A1 intron 26-specific probes were generated by PCR. Two oligonucleotide probes, specific to intron 24, were end-labeled with biotin. Probes, including the PCR generated intron 26, were nick translated with digoxigenin-11-dUTP or biotin-16-dUTP (Boehringer). 2.5 mg/ml of probe was lysylated with 100 U of E. coli rT rNA, Salmon Sperm DNA and Human Cot 1 DNA, resuspended in 100% formamide and heat denatured before addition of hybridization buffer to achieve 50% formamide and 2x SSC. For hybridization to RNA, probe was applied directly, but for DNA hybridization cellular DNA was denatured in 70% formamide for 2 min at 70°C. A sequential procedure was required to visualize RNA and DNA simultaneously (Xing et al., 1995). This involved initial hybridization and detection of RNA only, followed by fixation of that signal before proceeding with 0.07 N NaOH denaturation to hydrolyze RNA and denature cellular DNA. The DNA was then hybridized with a probe of identical sequence but different label. Biotinylated probe was detected with FITC-avidin (Boehringer) and digoxigenin probe with rhodamine-conjugated anti-digoxigenin (Boehringer). The SC-35 C-35 in 50% formamide was performed using a mouse monoclonal antibody against the spliceosome assembly factor SC-35 (Sigma-Aldrich; Fu and Maniatis, 1990), which may also detect another antigen that colocalizes to the same nuclear domains (Fu et al., 1992).

Exon Suppression Hybridization

To visualize the collective distribution of intron-containing transcripts within a given gene, a method was developed to simultaneously target all intron sequences while excluding exon sequences. The exon suppression hybridization technique, involves competing the hybridization of full-length genomic probe with excess unlabeled full-length cDNA. This involves a large excess of unlabeled cDNA in the hybridization reaction, after a prehybridization with unlabeled cDNA. 2 mg of full-length COL1A1 cDNA (cut to 50–200 bp) was prehybridized to nonadeninated WI-38 cells with probe mix also containing 10 mg of each of salmon sperm DNA, E. coli rT rNA, and human Cot 1 DNA for 2–16 h at 37°C in 50% formamide and 2x SSC. A ferter three rinses, cells were hybridized a second time with 2.5 mg/ml of heat-denatured, digoxigenin-labeled full-length collagen 1A1 genomic probe, containing 4 mg of unlabeled full-length COL1A1 cDNA. A ferter rinsing, the collective intron signal was detected with rhodamine anti-digoxigenin antibody. Control samples were treated identically, except that the labeled probe was a full-length cDNA, (to ensure that no exon sequences were detected). For colocalization of collective-intron signal with the cDNA track, an initial hybridization used labeled cDNA probe, before proceeding with the exon suppression protocol.

Microscopy and Image Analysis

All results were observable directly through the microscope, without image processing, and digital imaging was used to capture representative images and for microfluorometric measurements. A Zeiss A xipion microscope was equipped with a triple-bandpass epifluorescence filter (Chroma Technology, Inc.) and a custom-made color filter wheel, such that three colors can be viewed simultaneously with no optical shift. Digital images were captured with a Photometrics series 200 CCD camera (Photometrics, Ltd.). Morphometry and microfluorometry programs (G. W. Hannaway...
Results

Molecular Characterization of COL1A1 RNA in O54 Cells

Cells from patient O54 with OI Type I were investigated to determine the specific mutation in the COL1A1 gene, using procedures previously reported for OI Type I patient O53 (Stover et al., 1993). As described below, results showed a splice-junction defect that proved identical to that previously seen for patient O53. Initial evidence of a retained intron in patient O54 was obtained by indirect RNase protection performed on nuclear and total RNA extracts in patient and control cells (Fig. 1 A). The primary band at 1.8 kb, seen in all lanes, is the fully protected RNA, colinear with the ssDNA probe derived from Hf404, which encompasses exons 19–43 (Stover et al., 1993). However, in the nucleus of patient O54 (lane 1), a 1.4-kb and a weak 0.4-kb band were also present, indicating that one of the RNA species in these heterozygous cells contains an insertion in the region encompassing Hf404. Note that sometimes there was a very faint band barely detectable at ~1.5 kb in normal cells (Fig. 1 A, lane 3, arrow); this represents a small fraction of nuclear transcripts that still retain intron 24, which is removed late in the splicing process (see below).

Precise identification of the mutant insertion was obtained by first synthesizing a probe by RT/PCR of nuclear RNA from cell strain O53 using oligonucleotide primers derived from exons 24 and 28 (Stover et al., 1993). The 552 nt cRNA probe was then used for direct RNase protection (Fig. 1 B). Hybridization to nuclear and cytoplasmic RNA extracts in both patient and control cells produced bands at 209 and 90 nt, indicative of the fully spliced exons 24–26 and 27, 28 from the normal allele. Only in the nuclear fraction of O54 cells was there a distinct 442 nt band (Fig. 1 B, lane 3). This band (slightly smaller than the full probe due to polylinker sequences in the cRNA) is consistent with full protection by a mature RNA containing a retained intron 26. This band is absent from the cytoplasmic RNA of patient cells, as well as from the nuclear and cytoplasmic RNA of control cells. As exemplified here, the relative amount of the two lower bands in the O54 nuclear RNA was often significantly less than the intron containing band, suggesting that this larger transcript may accumulate within the nucleus. Genomic DNA isolated from O54 cells was cloned and sequenced (not shown), which showed the identical mutation reported earlier for patient O53 (Stover et al., 1993), a G to A transition in the first position donor site of intron 26, which is associated with retention of intron 26.

Failure of the mutant COL1A1 RNA transcript to appear in the cytoplasm could result from impaired export to the cytoplasm or rapid degradation. To distinguish between these two possibilities, the rate of disappearance of normal and mutant transcripts from the nucleus of cell strain O54 was examined after inhibition of RNA synthesis with 25 mM DRB (5,6-dichloro-1-β-D-ribofuranosyl benzimidazole), a specific RNA polymerase II inhibitor (Zandomeni et al., 1986). Control analysis using direct RNase protection by hybridization of cellular RNA with Hf404, which encompasses COL1A1 exons 19–43. Nuclear mRNA from heterozygous patient O54 contains the primary 1.8-kb band, as well as a 1.4-kb and a weak 0.4-kb band, indicating the presence of an insertion (lane 1). Nuclear mRNA from a control patient (411), which lacks an intron retention mutation (lane 2). Total RNA from control cells (lane 3). Note the presence of a faint 1.5-kb band in normal cells suggesting late removal of intron 24 (see results). (B) Direct RNase protection. A 552-bp probe encompassing exons 24–28 and

Figure 1. (A) Indirect RNase protection of cellular RNA with Hf404, which encompasses COL1A1 exons 19–43. Nuclear mRNA from heterozygous patient O54 contains the primary 1.8-kb band, as well as a 1.4-kb and a weak 0.4-kb band, indicating the presence of an insertion (lane 1). Nuclear mRNA from a control patient (411), which lacks an intron retention mutation (lane 2). Total RNA from control cells (lane 3). Note the presence of a faint 1.5-kb band in normal cells suggesting late removal of intron 24 (see results). (B) Direct RNase protection. A 552-bp probe encompassing exons 24–28 and
protection of total mRNA with the same exon 24–28 probe used above confirmed that the DRB conditions adequately inhibited new COL1A1 mRNA production, and suggested a cytoplasmic mRNA half life of 9 h, similar to other reported values (Sandmeyer et al., 1981). The half-life of COL1A1 nuclear RNA in DRB-treated cells was then examined in 054 cells at 0, 4.5, and 9 h. If nuclear transport is normal and cytoplasmic degradation rapid, then equivalent half-lives in the nucleus for normal and mutant transcripts would be predicted. As illustrated in Fig. 1 C, the nuclear half-life of the mutant COL1A1 transcripts is atypically long relative to the normal, with substantial mutant RNA present in nuclei up to 9 h, when normal transcripts were no longer detectable. Similarly, in a second DRB experiment (not shown), it took 2.5 h of DRB treatment for the normal RNA to degrade 50% versus 4.0 h for the mutant RNA, and the ratio of mutant to normal nuclear transcripts increased from 2.8 at 0 h DRB to 5.2 after 9 h DRB. More experiments would be necessary for precise determination of nuclear half-life, but these results show that mutant transcripts are longer lived in the nucleus than the normal, consistent with a defect in RNA transport. Although the mutant transcripts contain nonsense codons, they are clearly not under-represented and rapidly degraded in the nucleus (see Discussion).

**Normal and Mutant Transcripts Form Tracks in Nuclei of OI Cells**

Having investigated the molecular basis for the patient disease, we next undertook a direct visual examination of both normal and mutant COL1A1 alleles to address where within nuclear structure a transport block might occur. This involved a series of molecular cytological analyses using procedures optimized for in situ detection of nuclear RNAs and/or DNA (see Materials and Methods). Results are based on analysis of hundreds of cells viewed directly through an epifluorescence microscope, with digital imaging used to capture representative photographs. In addition, we developed for this study an approach using morphometry and microfluorimetry to objectively quantify cytological results in large numbers of cells.

The impact of the intron retention mutation on COL1A1 nuclear RNA distribution must be understood relative to RNA from the normal gene, hence this is briefly considered first, some aspects of which were published previously (Xing et al., 1995). Collagen nuclear RNA, detected with either a cDNA or genomic probe, is generally distributed in two prominent nuclear accumulations, with weak fluorescence often visible throughout the nucleoplasm. As previously reported, this weaker fluorescence appears as tiny punctate spots at the margin of sensitivity, possibly reflecting a broadly dispersed distribution of individual molecules (Xing et al., 1995). The prominent RNA accumulations are not linear but typically have somewhat elongate or elliptical morphology (Fig. 2, A, B, F, and H), between 0.5–3 μm in longest dimension. In contrast, hybridization to the entire collagen gene (~18 kb) produces a point signal of ~0.2 μm. Using a procedure that allows simultaneous hybridization of a full-length gene and its RNA in two distinct colors (Xing et al., 1995), the gene consistently localizes to one end of the RNA signal, which extends well beyond it (Fig. 2 A).

To understand the implications of these results, it is important to consider that any distance resolvable by fluorescence microscopy is large from a molecular perspective, given the 0.1–0.2-μm limit of resolution. Hence, the COL1A1 RNA accumulations are substantially larger than the full-length gene (which may well be smaller than the ~0.2-μm signal it generates, see Fig. 6 B legend). Hence, we will refer to these RNA accumulations as tracks, defined previously to indicate a posttranscriptional RNA accumulation that extends beyond and to one side of the gene (Lawrence et al., 1989; Xing et al., 1993; Xing and Lawrence, 1993). Further evidence is considered below that these tracks represent detached transcripts in transit from the gene, as distinct from nascent transcripts on the gene (a tree). COL1A1 RNA tracks are more globular than the highly linear viral RNA tracks in some cell types (Lawrence et al., 1988), and normal COL1A1 RNA tracks do not extend to the nuclear envelope (Xing et al., 1995).

A priori it was not known whether mutant transcripts would also form tracks like the normal RNA. For example, if the splice-junction defect blocks transcripts directly at or on the gene, then we would predict that RNA from the mutant allele would not form a track-like accumulation, but a smaller spot, likely brighter than the gene signal but of similar dimension. As illustrated in Fig. 2, B and C, COL1A1 transcripts visualized with a 1.8-kb cDNA probe in both WI38 and 054 cells, generally showed two elongated nuclear RNA tracks. Digital morphometry of RNA accumulations in a large sample (Fig. 3, A and B) indicated that track size was not markedly or consistently different in normal and mutant cells. However, in 054 cells the data generally are shifted slightly to the right, and there was more variability in size and intensity of RNA from the two alleles. This provided a preliminary suggestion that the mutant tracks might contain more RNA, but because there is also some variability between normal alleles, it could not be confidently concluded on this basis alone that mutant and normal tracks were distinguishable. The fact that RNA tracks in 054 cells were clearly not smaller or less intense than in WI38 cells indicates that the splicing mutation does not cause an accumulation of RNA directly at the gene, since both RNA signals extended well beyond the size of the gene signal (see below). Nor does the splice defect apparently result in rapid degradation of transcripts from the mutant allele, consistent with the above RNase Protection and half-life analysis, since substantial RNA accumulations were present with both alleles.

With regard to RNA distributions elsewhere in the nucleus, there were no atypical accumulations of COL1A1 RNA apparent at the nuclear envelope, around or within the nucleolus, or at other sites throughout the nucleoplasm. In some mutant 054 cells there were multiple very weak patches of fluorescence signal in the nucleoplasm, often near the mutant allele (see Fig. 2 J), although there was also sometimes nucleoplasmic signal in normal cells (discussed above). However, in all cases the most clear and localized accumulations occurred in the highly concentrated primary RNA tracks. Further analysis was required to determine definitively if the mutant tracks differed from the normal.
Figure 2. Fluorescent in situ hybridization using COL1A1 genomic, cDNA and intron 26 specific probes in normal and patient 054 fibroblasts. (A) Collagen transcripts (red) detected with a digoxigenin-labeled full-length genomic probe extend beyond and to the side of the gene, detected with the same probe sequence labeled with biotin (green). (B and C) Nuclear RNA tracks detected with a 1.8-kb 3’ end cDNA probe in (B) W138 and (C) 054 cells. (D and E) Intron 26 RNA in (D) W138 and (E) 054 cells. A arrow indicates small intron 26 signal associated with normal allele. (F, G, H, I, and J) Colocalization of intron 26 (red) and the cDNA track (green, overlap of red and green appears yellow). (F and H) Small intron 26 signal (arrows) localizes to the end of the cDNA track in W138 cells. An enlargement of the tracks (H) shows that intron 26 has been removed from a portion of the track. (G and J) In 054 cells, the normal cDNA track (arrow) is polar with respect to intron 26, whereas the other track is not, thus enabling identification of the mutant allele. (I) Higher magnification view of the retained intron 26 (red) completely colocalizing with the cDNA track from one allele in heterozygous mutant cells (green). (J) Illustrates patchy secondary signal sometimes observed in 054 cells, often near the mutant allele (see text). Bars, (A, H and I) 1.3 μm; (B, C, D, E, and F) 5 μm; (G and J) 4 μm.
RNA Tracks from the Mutant Allele Lack Polarity for Intron 26

Possible differences between the normal and mutant RNA accumulations might be more obvious from the distribution of intron 26, which is retained in the mutant primary transcript. As illustrated in Fig. 2 D, hybridization of an intron 26-specific probe to normal WI-38 cells shows signal in two small foci of similar size and intensity, clearly detectable above background in most but not all cells. In contrast, the intron 26 accumulations within heterozygous 054 nuclei differ markedly, with one a compact focus and the other a larger track (Fig. 2 E). Fig. 3, C and E, summarizes quantitation of areas from over 100 intron signals. In heterozygous 054 cells, 45% of intron accumulations were >5 μ², in contrast to only 1% in WI-38, and there was a much greater difference in the areas between the two alleles per nucleus for 054 versus WI-38 (5.8 versus 0.73 μ², respectively). Similar differences were observed in cells from patient 053 (not shown). These data suggest that the larger intron track, unlike that seen in normal cells, could be from the mutant allele. As illustrated in Fig. 2 E, G, and J, the small normal intron focus was not brighter than the intron signal throughout the mutant track, as expected if the intron focus with the normal allele represents primarily introns in pre-mRNA (i.e., rather than a pile-up of excised introns). As previously suggested, excised introns...
may disperse widely (Zeitlin and Efratiadis, 1984; Xing et al., 1993), however, a diffuse distribution of small COL1A1 introns is difficult to discern definitively from background.

In normal fibroblasts, simultaneous detection of the cDNA and intron probe in two different colors has shown the intron 26 detected only at one end of the larger track defined by cDNA hybridization (Fig. 2, F and H). This polar distribution was interpreted to indicate splicing, revealed by a distinct spatial arrangement of unspliced and spliced transcripts as they progressed further from the gene (Xing et al., 1993, 1995). If this spatial polarity is indeed indicative of splicing, then in 054 cells, tracks from the mutant allele should lack this polarity. Over 100 RNA tracks were examined qualitatively and quantitatively. In WI-38 cells, both intron accumulations formed tiny foci towards one end or one side of the cDNA track; rarely was any trace of signal detectable throughout the rest of the cDNA track. In contrast, the two RNA tracks within each 054 nucleus were clearly different, with one exhibiting the normal polar distribution of intron 26 and the other showing no such polarity. Instead, a bright intron 26 signal uniformly overlapped one entire cDNA track in patient cells (see Fig. 2, G, I, and J).

The percent of the cDNA track containing colocalized intron signal (% overlap) was calculated using digital microfluorimetry (Fig. 4, A and C). In 054 cells, 84% of cells had a single RNA track with >80% overlap, whereas no control tracks fell in this range. This difference was further demonstrated in Fig. 4, D and E, in which the two alleles in each cell were separated into those with the larger or smaller percent overlap. Unlike the WI-38 cells, the marked bimodal distribution for 054 cells clearly reflects a heterozygous cell with one normal and one splice-detective allele. These results establish that normal and mutant COL1A1 tracks can be distinguished by direct microscopy or quantitative digital imaging, within as little as a single cell. Moreover, the nonpolar nature of RNA tracks from the splice-detective allele proves that the polarity normally observed reflects splicing and distinct morphological distributions of spliced and unspliced transcripts, for intron 26.

**COL1A1 Introns Are Predominantly Spliced at One End of the RNA Track**

A n important potential implication of the above results is that the splice-detective RNA does not back up directly at the gene, but forms a track of molecules that we suggest are largely released transcripts, reflecting an early step in RNA export. Therefore, it is important to firmly establish whether normal RNA tracks are simply accumulations of nascent transcripts (a tree) on a highly distended gene, or, alternatively, are comprised largely of detached transcripts in movement from the gene (a track), perhaps progressing from less to more mature transcripts. The distribution of different intron sequences provides insights into the structural basis of the track and its relationship to splicing. The Christmas tree model would predict a distribution of introns along the length of the track in accordance with position of sequences within a presumably highly distended gene. (Such an intron distribution was recently reported for nascent transcript trees along the huge dystrophin gene [Smith et al., 1999].) The restriction of intron 26 to one extremity of the COL1A1 cDNA track, as well as the small size of the COL1A1 transcript, is most consistent with the alternative model of posttranscriptional RNA tracks (see Fig. 6); however, this one intron could be atypical of the numerous introns in this highly complex primary transcript.

To investigate where splicing occurred for the various 50 COL1A1 introns and visualize their distribution presented a technical challenge. A hybridization strategy was devised that made it possible to visualize essentially all intron-containing RNA from a given gene simultaneously, while excluding signal from exons (see Materials and Methods). This exon suppression technique involves competing hybridization of full-length genomic probe with excess unlabeled full-length cDNA, such that only intron RNA sequences are available to hybridize labeled probe. The collective intron thus reveals the summative distribution of the 50 COL1A1 introns, which can then be analyzed relative to either the cDNA track or other nuclear structures. Control samples confirmed that competition completely suppressed exon signal, as no specific signal was observed when the probe was a labeled cDNA (Fig. 5 C). In contrast, in the same experiment samples hybridized with full-length genomic probe consistently resulted in a bright focal signal, indicating that this derived from introns exclusively. To examine this collective intron signal relative to the cDNA track in distinct colors, hybridization with a biotin-labeled cDNA probe was followed by exon suppression hybridization with digoxigenin-labeled genomic probe (see Materials and Methods).

Examination of hundreds of normal cells in repeated experiments revealed a small bright focus of collective intron signal. This signal overlapped the cDNA track, but was clearly smaller and generally concentrated at one end or the periphery of the cDNA accumulation, similar to intron 26 (see Fig. 5, A and B). Remarkably, the size of the collective intron signal was not significantly different from that for intron 26, nor was the extent of overlap with the cDNA track (compare Fig. 3, C with D and Fig. 4, A with B). Given that the signal will include some excised introns as they disperse from the splicing site, the intron-containing pre-mRNA may be even more tightly localized than it appears here. The collective intron area averaged 1.24 μm² (n = 80) as contrasted with 4.1 μm² for the cDNA. Only rarely was any trace of collective intron signal discernible throughout the cDNA track. Since the complexity of the probe used will be less efficient in the detection of any one intron, we cannot conclude that all introns have been removed. We can conclude that almost all introns are removed in a tightly localized spot, consistent with cotranscriptional splicing, with transcripts in the track predominantly spliced.

These findings counter the interpretation of the track as merely nascent transcripts on a distended gene, and support the model of intact transcripts in an early stage of export (Fig. 6, model). A key point is that the COL1A1 transcripts do not disperse equally in all directions, possibly reflecting vectorial movement. Remarkably, the small focal concentration of intron 26 RNA, visualized simultaneously relative to the gene, can be resolved as slightly but consistently concentrated to one side of the gene (see Fig. 6).
Figure 5. Analysis of intron RNA distributions within COL1A1 RNA tracks and SC-35 domains, in normal WI38 cells. Exon-suppression hybridization was used to investigate the collective intron distribution (see Materials and Methods). (A and B) Colocalization of collective intron (red) and cDNA (green). Most of the introns colocalize in a tight focus at the edge of the cDNA track. Only occasionally is the intron signal elongated or diffuse (B, lower right). (C) cDNA tracks were completely competed away in control exon suppression experiments (compare with Fig. 2, B, F, and H). Colocalization of collective intron (red, D and E) or cDNA (red, F and H) with SC-35 domains (green). (D and E) The collective intron signal is concentrated at the edge of SC-35 domains. Relatively little intron RNA is detected throughout the central portion of the domain despite the comparatively high concentration of SC-35. Graphs show the fluorescence intensity for pixels along the line shown in blue (E). Note the sharp increase in SC-35 fluorescence at the boundary of an SC-35 domain. (F and H) cDNA tracks (red) fill all or most of the SC-35 domains with cDNA levels mirroring SC-35 (H). (G) Colocalization of intron 26 (red) and intron 24 (green). Intron 24 is removed late, and is retained throughout much of the RNA accumulation or track from which intron 26 has already been removed. (I) Colocalization of intron 26 RNA (red) and the full-length COL1A1 gene (green). Signal from intron 26 RNA partially overlaps the gene, but in 20 randomly chosen examples, was consistently displaced to one side. Bar, (A, C, D, and I) 4 μm; (B and E, right most image) 1.3 μm; (E, left most image) 1 μm; (E, center image) 1.6 μm; (F and G) 3 μm; (H) 1 μm.
This further supports the model (Fig. 6) and suggests the intriguing possibility that from the onset transcripts emanate asymmetrically from the gene. Importantly, the initial movement from the gene into the rest of the track does not require complete splicing, since intron 26-containing transcripts in OI cells also form a track.

**Splicing at the Periphery of an SC-35 Domain with Essentially Spliced Transcripts throughout the Domain**

If RNA molecules within the track are not localized because they are bound to the gene, what is the structural basis for these RNA accumulations? A key to understanding this lies in the previous demonstration that in normal fibroblasts COL1A1 RNA overlaps discrete intra-nuclear domains enriched in SC-35, and that the gene consistently positions at the edge of this domain (Xing et al., 1995). The above results showing collective intron signal at one end of the track suggest indirectly that splicing of all of these COL1A1 RNA introns may also occur at the periphery of the SC-35-rich domain. This point bears on a pivotal question: whether the domain comprises just diffuse factors that are accumulated via binding to COL1A1 introns, or whether the domain is a larger distinct entity, with which the gene associates and into which more mature transcripts enter after transcription and splicing are essentially complete.

To address this directly the collective intron RNA signal was investigated relative to the SC-35 domain, and, for comparison, a similar analysis of cDNA was done. Fig. 5, D and F, illustrates the marked differences in the extent of overlap for each. Collective intron RNA occupied a clearly smaller, more peripheral part of the SC-35 domain. In contrast, the cDNA signal was not restricted to the periphery but occupied much more of the discrete SC-35-rich region, having a somewhat similar morphology. Notably, the factors within most of the domain are in substantial excess of the intron-containing transcripts. Fig. 5, E and H, shows this by measuring the relative distributions of collective intron or cDNA and SC-35 signal across the domain in a series of cells. The histograms show the raw fluorescence signal in individual pixels along the line indicated. The

![Figure 6](https://example.com/fig6)

**Figure 6. Early steps in nuclear transport of COL1A1 RNA.** models A and B, summarize results presented here and their interpretation whereas Model C is more hypothetical. (A) Schematic of normal and mutant COL1A1 RNA tracks (intron 26 in red and cDNA in green) and their progression through the nucleus. An early step in the export of both normal and mutant COL1A1 transcripts involves movement from the gene (blue) into an SC35 domain (dark gray, step 1). Both properly spliced lower left track) and splice-defective RNA (upper track) accumulate within the domain, but the mutant transcripts fail to undergo the next step involving release from the domain and exit to the rest of the nucleoplasm (step 2). The fully spliced normal mRNA then traverses the nucleoplasm to the nuclear pore (step 3), which may involve more random dispersal of the RNA. A later or final step in nuclear export would involve RNA translocation through the nuclear pore (step 4). (B) Model showing the RNA track as a posttranscriptional accumulation of predominantly spliced transcripts in an early stage of export. The collective intron RNA (red) forms a focal signal at one end of the larger RNA track detected by cDNA (green), which in turn is much larger than the gene (blue) positioned at one end. The gene itself is at the limits of resolution, and thus may be smaller than the 0.2-μm fluorescent signal it generates. On a molecular scale, these transcripts have moved an appreciable distance relative to the gene. For COL1A1 RNA the track resides within an SC35 domain (not necessarily the case for all RNA tracks). (C) Hypothetical model showing a functional rationale for coupling the completion of mRNA maturation and release for mRNA export with the recycling/preassembly of factors within SC-35 domains. Because the RNA metabolic machinery for processing and transcription requires the interaction of such a large number of different factors, their concentration at specific sites and preassembly into components of a spliceosome or holoenzyme would facilitate recycling and maintenance of a high rate of splicing. The RNA template is synthesized at the periphery of the domain, and once released from factors within the domain would be released for export and the factors actively recycled at that site for rapid reuse. The closer the completion of processing occurs to putative sites of recycling the more efficient the total process. Further, if the time required to recycle splicing factors is long relative to the time they actually are engaged on the RNA, then sites of high splicing activity might be associated with large accumulations of factors not directly bound to unspliced RNA but engaged in dynamic recycling.
peak fluorescence for collective intron signal drops sharply in the periphery of the domain, whereas the SC-35 fluorescence remains uniformly high throughout, as does the cDNA signal. Note that the SC-35 domain itself has discrete boundaries defined by a sharp increase in concentration, as does the COL1A1 RNA track. These results demonstrate that the majority of the introns within the collagen gene are removed at the domain periphery and that transcripts entering the central regions are already predominantly spliced. Further, since the vast majority of introns have already been removed, the bulk of the factors that form the domain are not immediately bound to either splice-junction sequences of COL1A1 introns or to excised introns. As considered further below, the domain appears as a distinct entity that contains both COL1A1 mRNA and excess factors, beyond those directly engaged on COL1A1 introns.

**Mutant Transcripts Enter SC-35 Domains, but Their Exit Is Impeded**

If intron 26 is normally removed before transcripts enter the SC-35 domain, then it is possible that a defect in splicing could prevent the RNA from distributing through the domain and cause an abnormal RNA accumulation outside it. Alternatively, if intron removal is necessary for transcripts to exit the SC-35-rich domain, then mutant transcripts might accumulate within this region. In fact, in 054 cells mutant transcripts accumulated within SC-35 domain as do normal transcripts (not shown). Transcripts carrying a defect in splicing, therefore, extend into the SC-35 domain, as do the normal transcripts. As noted above, results suggested there might be more variability of track size in the heterozygous mutant cells (Fig. 3 B), raising the possibility that mutant tracks contain more transcripts. To determine the relative numbers of transcripts from the mutant and normal allele that accumulate within the track/ domain (irrespective of whether they are fully spliced) the size and total fluorescence of cDNA signal for each was quantitated and compared within individual 054 cells. Although absolute measures of transcript number are less reliable, FISH does provide a means to determine the relative number of transcripts with each allele in each cell. Because our analysis relies on comparison of normal and mutant alleles within the same nucleus, technical variations are minimized.

As established above, mutant and normal tracks could be discriminated based on the intron/cDNA overlap. Hence, the cDNA signal associated with each allele was measured to indicate the relative transcript number, and then the intron/cDNA overlap was used to distinguish the normal versus the mutant allele. This strategy was used to generate objective comparisons in a large sample of normal versus mutant tracks within each individual nucleus. The average area of the thus identified mutant track was 7.5 μ², significantly larger than the normal RNA track, which was 5.2 μ² in 054 and 5.1 μ² in WI-38. Mutant transcripts also were more concentrated within the track and throughout the domain. In ~80% of cells the mutant track had higher total intensity of cDNA signal. The total fluorescence signal in normal and mutant RNA tracks within each cell was expressed as a ratio. The mean ratio of fluorescence mRNA from the mutant versus normal RNA track in 054 cells was 2.1:1, with some mutant tracks having much larger increases. No such significant difference was seen for RNA tracks in WI-38 cells, where the average ratio was close to one. The average twofold increase in mutant RNA within the domain is consistent with other observations, including the 2.8:1 ratio of mutant to normal transcripts by RNase protection, and the change in nuclear ratio of COL1A1 to 1A2 from 2:1 to 4:1 (Genovese and Rowe, 1987). We do not rule out that there could be some additional very dispersed mutant transcripts, however, the higher RNA accumulations in the mutant track and domain indicate that improperly spliced transcripts are impeded from exiting this region. Since nuclear RNA accumulations disperse and must be regenerated after each cell division (e.g., Carter et al., 1991; Clemson et al., 1996), the average mutant RNA accumulation within the track might otherwise be even greater. Hence, cells that have not recently divided may be those with the largest mutant RNA accumulations, which sometimes appear to spill into adjacent domains.

We conclude that mutant COL1A1 transcripts accumulate to abnormal levels within the track, with an average twofold increase, and some mutant tracks showing substantially more. Furthermore, the site where this mutant RNA accumulates is in the SC-35 domain.

**Complete Splicing May Be Rate-limiting for Transcript Release from the SC-35 Domain**

An impediment to transit from the SC-35 domain could result directly from failure of the remaining splice-junction sequence to be released from splicing factor complexes within the domain, or could result less directly because some subsequent step necessary for further transport of mutant transcripts does not occur. The finding that RNA from normal COL1A1 alleles also forms some accumulation, a track within the domain, indicates that some slow or rate-limiting step occurs there normally; this may become more rate-limiting for the intron-retention mutant. If the normal RNA were completely spliced before accumulating within the domain, then this would suggest that some step other than splicing must limit the rate of export of normal RNA from the domain. Hence, it became important to address whether the COL1A1 RNA within the track and domain was indeed completely spliced. We addressed this by examining the distribution of intron 24, which RNase protection analyses suggested was removed substantially later than most other introns, including intron 26 (as may be apparent in Fig. 1, lane 3; Genovese et al., 1989; Rowe, D., unpublished observations). To determine if the temporal difference in splicing of this intron resulted in a different spatial distribution of intron 24, introns 24 and 26 were examined simultaneously in two distinct colors in normal WI-38 cells. As shown in Fig. 5 G, the distribution of intron 24 is much more extensive than intron 26; it is not confined to a tight focal accumulation as is intron 26 (and collective intron), but is detected as a larger accumulation with a globular or elliptical morphology. Hence, the late splicing of this intron is reflected in its morphological distribution, indicating that it is not all spliced at the periphery of the track and domain. Since the small (88 bp) intron 24 represents ~1% of the total-intron signal, the later splicing of this intron is not inconsistent with the find-
ing that the vast majority of COL1A1 introns are removed at the domain periphery, likely cotranscriptionally. These results directly show that COL1A1 transcripts continue to mature as they transit away from the gene, with at least this one intron apparently split posttranscriptionally at some distance from the gene. Moreover, this finding makes clear that, although transcripts entering the domain are overwhelmingly spliced, completion of 100% of intron splicing could still be the rate-limiting step in export of the RNA from the domain.

**Discussion**

This study advances understanding of the relationship between nuclear structure and pre-mRNA metabolism, in both normal cells and a specific genetic disease. Results identify a previously undefined step in nuclear transport, the point at which mutant transcripts become entrapped in O1 patient cells. Mutant transcripts do not accumulate at the nuclear pore or nucleolus, or directly on the gene, but ~1–3 μm from the gene within a specialized region of the nucleoplasm highly enriched in numerous RNA metabolic factors. Our results suggest that export of COL1A1 mRNA can be broken down into several steps, which include: (a) movement from the gene into the adjacent SC-35 domain, generating an RNA track to one side of the gene, (b) passage (of normal transcripts) from the domain into the rest of the nucleoplasm, (c) traversal through the nucleoplasm to the nuclear envelope, and (d) translocation through the pore to the cytoplasm (see Fig. 6 A). The significance of the RNA distributions observed depend largely on our observations and interpretations of the structural basis for RNA tracks and the SC-35 domain, hence these points will be discussed in more detail first.

**Tracks: an Accumulation of Predominantly Spliced and Released Transcripts Emanating to One Side of the Gene**

Perhaps due to the differences between various models studied (see Introduction), there has remained some resistance or confusion as to whether posttranscriptional RNA tracks really exist, or whether localized RNA accumulations are merely nascent transcripts still attached to the gene. However, the differences seen for different RNA s are not contradictory, but simply represent the complex biological realities, related to the differences in the relative rates of different steps in RNA metabolism (discussed in Shopland and Lawrence, 2000). Whether a track of posttranscriptional transcripts is seen at all will be a function not just of the RNA’s complexity, but likely reflects the extent to which processing and transport keeps pace with completion of transcription (Smith et al., 1999).

Because the point is critical to our overall conclusions, we briefly summarize evidence that COL1A1 RNA tracks are not just nascent transcripts on the gene, but primarily spliced, detached transcripts at an early step in export (Fig. 6 A). (a) The full-length gene signal is much smaller (~0.2–0.3 μm) than the RNA signal (~1–3 μm in longest dimension), and consistently positions to one side. (b) Intron 26 localizes with the gene, at one end of the larger track (detected with full-length cDNA). In contrast, signal with a 3’ cDNA illuminates the larger track (as with full-length cDNA). (c) The splice-defective RNA track shows no polarity of intron 26, proving that the polarity does reflect a distinct cytological arrangement of spliced and unspliced transcripts. (d) An important point is that the collective intron RNA signal (representing ~50 introns from across the gene) colocalize in a small spot (with intron 26, at or near the gene). If the track represented nascent transcripts on distended DNA that escaped our detection, the introns would distribute differently through the track reflecting their location within the gene (as previously seen for nuclear RNA from the large dystrophin gene, Smith et al., 1999). (e) Intron 24 is not restricted to a spot-like signal at one end of the track, consistent with other evidence that this intron is spliced slowly.

The accumulation of RNA to one side of the gene indicates that a slow or rate-limiting step in RNA transport occurs within the nucleoplasm, ~1–3 μm from the gene (discussed below). COL1A1 RNA tracks do not extend to the nuclear envelope, but into a discrete region or SC35 domain (Xing et al., 1995). Upon leaving the track or domain (Fig. 6 A, step 2), transcripts may either diffuse randomly or move rapidly in some more directed process to the nuclear pore (step 3). A discussed previously, involvement of structure or diffusion in RNA transport are not mutually exclusive possibilities (Xing et al., 1993; Daneholt, 1999; Shopland and Lawrence, 2000). We do not rule out entirely that transcripts move in all directions from the gene but accumulate only to one side. However, the finding that even the intron 26 RNA signal is asymmetric relative to the gene strongly suggests that COL1A1 transcripts, possibly even as they are transcribed, may emanate vectorially to one side of the gene.

**New Transcripts and Excess Factors within Domains**

The track of COL1A1 transcripts extends into one of the largest SC-35 domains in the nucleus. A straightforward a priori explanation would be that splicing factors bind to the many introns of this complex pre-mRNA accumulated on the gene. However, several observations suggest that the domain contains, in addition to new transcripts, abundant excess factors not immediately bound to the 50 COL1A1 splice-junction sequences. Especially important, almost all COL1A1 introns have already been removed at the domain periphery, with essentially spliced transcripts accumulated throughout the SC-35-rich region. This supports that the COL1A1 domain is not just diffuse factors bound to unspliced transcripts, but a distinct (though related) structural entity, as suggested by earlier observations (Mcgen et al., 1995b; Xing et al., 1995). As we have discussed previously, formation of the domain may be initiated due to the high demand for splicing factors and/or the presence of a slow-splicing intron. However, we suggest that the structural basis for the track of COL1A1 RNA becomes the SC-35 domain, not just visa versa. SC-35 domains in general, like IGC, are discretely bordered inter-chromatin structures which frequently reside between chromosome territories (Fakan and Puvion, 1980; Carter et al., 1993; Visa et al., 1993; Zirbel et al., 1993) and are highly resistant to detergent extraction and nuclear fractionation (e.g., Blencowe et al., 1994, reviewed in
Nickerson et al., 1995; Stein and Berezney, 1996). Similarly, the COL1A1 RNA accumulation sits just outside its parent chromosome territory and is preserved after detergent or matrix extraction (Clemson et al., 1996). We conclude that COL1A1 RNA enters a specific region of the nucleoplasm, an inter-chromatin domain containing both excess factors and newly synthesized transcripts.

The Intra-nuclear Fate of Splice-defective COL1A1 Transcripts in OI

Concerning mutant transcripts in OI, the first point demonstrated is that transcripts retaining intron 26 do not rapidly degrade within the nucleus but accumulate there. Despite downstream nonsense codons that might trigger nonsense-mediated decay within the nucleus (see Willing et al., 1996), these transcripts have a longer nuclear half-life than normal. Although we cannot rule out that any nonsense decay occurs, the retained intron may impede further transport thereby circumventing nonsense-mediated decay, which may either be a cytoplasmic process for COL1A1 RNA (as for yeast mRNAs), or may occur at the nuclear envelope (M. Aquat; Jacobson and Peltz, 1996; Primorac et al., 1999).

Importantly, the splice-defective RNA does not accumulate directly at or on the gene, but can detach and take the first step into the domain. However the next step, transit from the domain, is blocked or slowed substantially (Fig. 6A, step 2). Although we sometimes see low levels of dispersed transcripts in neighboring SC35 domains, our results and very recent results of Custodio et al. (1999) generally agree that splice-defective RNA does not accumulate abnormally at the nuclear envelope, nucleolus or randomly thru the nucleoplasm, but accumulates closer to the gene. However, Custodio et al. (1999) forward a significantly different interpretation: that transgenic mutant RNA becomes trapped on the gene itself. Because the small globin transgene is present in 14 tandem copies, it would likely not be feasible to examine it with the detailed resolution achieved here for COL1A1 DNA and RNA (spliced and unspliced), and the relationship of globin transgenes to splicing factor domains was not examined. It is also possible that the globin transgenes with just two introns and integrated at an exogenous site may behave differently than complex endogenous COL1A1 RNA.

Some checkpoint for COL1A1 RNA export apparently occurs within the SC35 domain. We suggest that normal COL1A1 RNA undergoes one or more maturation steps there, and that splice-defective mutant transcripts do not efficiently pass this point. We show that intron 24 is spliced slowly within the domain, a slow-splicing intron may cause the RNA accumulation within the SC35-rich region. Completion of splicing may be just one step required for efficient transit, since factors or a holoenzyme within the domain could also be involved in packaging or screening the mRNA for export. Recent work shows that spliced transcripts are packaged in a specific nucleoprotein complex before export (Luo and Reed, 1999), and some factors involved (Blencowe et al., 1994) localize to SC-35 domains, as does an RNA helicase implicated in RNA transport (Ohno and Shimura, 1996), and hsp 90 transcripts containing no introns (Jolly et al., 1999).

We have considered the possibility that normal COL1A1 tracks comprise defective transcripts that have similarly failed to pass this putative checkpoint. However, this is implausible because it would suggest an error rate of ~50%, (normal tracks have about half as many transcripts), and because mutant transcripts have a longer nuclear half-life. Further, the dystrophin gene (74 introns) might be expected to have a higher error rate, yet no RNA track or overlap with domains is seen (Smith et al., 1999).

A Rationale for Linking Factor Recycling within SC-35 Domains with Release of mRNA for Export

We propose that within the COL1A1 domain several functions related to RNA metabolism are spatially and functionally integrated in a way that facilitates their efficiency and coordination. The hypothetical model in Fig. 6C depicts such a concept, suggesting that processing complexes within the domain are actively engaged in recycling, closely associated with mRNA maturation and release for RNA export. In a process that requires interaction of >100 components, the domain may serve to concentrate RNA metabolic factors together to facilitate rapid recycling/preassembly into spliceosomal subunits, to keep up with the high demands of this (and possibly other) highly active genes.

The further from the domain completion of splicing and disassembly occurs, the more disconnect and inefficient in the reassembly and reuse. This suggests a functional rationale why COL1A1 RNA splicing and export may be linked to a site of factor recycling. This cycle, in turn, is likely controlled by phosphorylation (see Y. Eakley et al., 1999). The COL1A1 domain could provide a more general paradigm for SC-35 domains. Although this remains to be determined, these domains are indistinguishable by numerous immuno-biochemical criteria, including enrichment with poly A RNA (Carter et al., 1991), and may correspond to IGC tightly juxtaposed with perichromatin fibrils (Xing et al., 1995; M. elcak et al., 2000). The abundance of 30-nm particles in IGC suggests some large preassembled complexes in these structures (Fakan and Puvion, 1980; Reed et al., 1988). Our hypothesis incorporates evidence from uridine labeling studies that domains contain a disproportionate excess of factors, but such studies do not conclusively show these regions are devoid of new transcripts. In addition to our findings, Bridge et al. (1996) reported adenovirus transcripts in IGC, this group and Puvion's discussed a possible role in export (V. et al., 1993; see also M. elcak et al., 2000), and recently one study reported that speckles do incorporate BrUTP (Wei et al., 1999). In addition to the complexities of extrapolating uridine label to pre-mRNA (M. et al., 1995b; Clemson and Lawrence, 1996), results here reveal another important factor: 80% of label would be removed from COL1A1 RNA before it enters the domain. If spliced mRNA's destined for export enter domains, and excised introns do not degrade immediately but release to the inter-domain space (Q. et al., 1992; Xing et al., 1993), then uridine label will underestimate numbers of transcripts in the domain relative to the nucleoplasm. Hence, SC-35 domains could more generally be sites of rapid factor use/recycling intimately associated with sites of high demand, rather than comprise less active factors with distant relation to RNA metabolism. Our suggestion that domains con-
tain active factors in rapid recycling is more consistent with recent data from Pfair and Mistle (2000) and K ruhlak et al. (2000), showing that factors within domains turnover with rates similar to those in the nucleolus (Shapold and Lawrence, 2000).

In either case, results here show that an SC-35 domain is important in screening and entrapping mutant COL 1A 1 RNA , which makes a critical difference in outcome between the mild Type I phenotype of patient 054 and severe phenotypes of OI Types II and III.

We thank Lindsay Shapold for critical reading of the manuscript and we thank K elly Smith and Lindsay Shapold for useful discussions.

This publication was made possible by National Institutes of Health grants G M 49254 and G M 53234 to J. Lawrence and A R 30426 to D. Rowe.

The contents are solely the responsibility of the authors and do not necessarily reflect the views of the granting agency.

Submitted: 16 D ecember 1999
R evised: 27 J une 2000
A ccepted: 28 J une 2000

References


