Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 7365–7369, July 1994
Genetics

Targeted insertions of two exogenous collagen genes into both alleles of their endogenous loci in cultured human cells: The insertions are directed by relatively short fragments containing the promoters and the 5′ ends of the genes

(homologous recombination/transcription-driven gene insertion)

ARUPA GANGULY, SARA SMELT, REEMA MEWAR, ANDRZEJ FERTALA, ALEKSANDER L. SIERON, JOAN OVERHAUSER, AND DARWIN J. PROCKOP*

Department of Biochemistry and Molecular Biology, Jefferson Institute of Molecular Medicine, Jefferson Medical College, Philadelphia, PA 19107

Contributed by Darwin J. Prockop, March 14, 1994

ABSTRACT **Previous studies demonstrated that type II procollagen is synthesized by HT-1080 cells that are stably transfected with constructs of the human *COL2A1* gene that contain the promoter and 5′ end of either the *COL2A1* gene or the human *COL1A1* gene. Since the host HT-1080 cells were from a human tumor line that synthesizes type IV collagen but not type II or type I procollagen, the results suggested that the constructs were integrated near active enhancers or promoters. Here, however, we demonstrate that a 33-kb construct of the *COL2A1* gene containing a 5′ fragment from the same gene was inserted into both alleles of the endogenous *COL2A1* gene on chromosome 12, apparently by homologous recombination by a nonconservative pathway. In contrast, a similar construct of the *COL2A1* gene in which the 5′ end was replaced with a 1.9-kb fragment from the 5′ end of the *COL1A1* gene was inserted into both alleles of the locus for the *COL1A1* gene on chromosome 17. Therefore, targeted insertion of the gene construct was not directed by the degree of sequence homology. Instead, it was directed by the relatively short 5′ fragment from the *COL1A1* gene that contained the promoter and the initially transcribed sequences of the gene. After insertion, both gene constructs were expressed from previously inactive loci.**

Targeted gene insertion, whereby an exogenous gene is inserted into the locus of an endogenous gene, generally occurs by homologous recombination. The process has been extensively studied in bacteria and fungi (see refs. 1–6) and more recently in cultured mammalian cells (7–16). The frequency of targeted gene insertion through homologous recombination can be increased by increasing the homologous length of the invading DNA, by cleavage of the invading or the target DNA (12), and by transcription of the targeted gene (1, 5, 6, 15, 17).

Several observations indicate that expression of recombinant collagen genes in stably transfected cells (18–21) is critically dependent on the presence in the host cells of a series of collagen-specific post-translational enzymes (22). In some of our initial experiments (21), we transfected mouse NIH 3T3 cells with a hybrid construct of the *COL2A1* gene in which the 5′ end of the gene was replaced with a 1.9-kb fragment from the *COL1A1* gene. A small fraction of the stably transfected cells synthesized human type II procollagen, apparently because the *COL1A1* promoter of the construct was active in the NIH 3T3 cells, which synthesize mouse type I procollagen but not type II procollagen. More recently, we found (23) that the same *COL1A1/COL2A1* hybrid construct could be used to synthesize recombinant type II procollagen by stable transfection of HT-1080 cells, a human tumor cell line that synthesizes type IV collagen but not type I procollagen or type II procollagen. Similar results were obtained with a construct of the *COL2A1* gene in which the 1.9-kb fragment from the *COL1A1* gene was replaced with the promoter and 5′ end of the *COL2A1* gene. The results suggested that the constructs were expressed because they had integrated near active enhancers or promoters. Here, however, we demonstrate that the construct containing the promoter and 5′ fragment of the *COL2A1* gene was inserted into both alleles of the *COL2A1* gene on chromosome 12, and the construct containing the promoter and 5′ fragment of the *COL1A1* gene was inserted into both alleles of the *COL1A1* locus on chromosome 17.

MATERIALS AND METHODS

DNA Constructs and Cell Transfection. The *COL2A1* minigene and the *COL1A1/COL2A1* hybrid gene (Fig. 1) were assembled in cosmid clones as described previously (21, 24). Mouse NIH 3T3 cells and human kidney tumor cells (HT-1080; American Type Culture Collection CCL 121) were grown under standard conditions and cotransfected (25) with the calcium phosphate precipitation method with the procollagen gene constructs cleaved with *Sal* I but not separated from the vector together with a 1:200 mixture of *Bam*HI-linearized plasmid (26) containing a neomycin-resistance gene with mouse metallothionein promoter, pDMMTNeo. In a typical experiment, 10^7 cells were transfected, and 200–300 clones of G418-resistant NIH 3T3 or HT-1080 cells were obtained. The medium proteins were assayed by Western blot analysis with rabbit anti-human procollagen II antibodies (21) and secondary antibodies of anti-rabbit IgG conjugated to alkaline phosphatase (Promega). Genomic DNA was isolated on a workstation (Genepure 341; Applied Biosystems), and Southern blot analyses were carried out, using DNA fragments labeled with [^{32}P]dCTP by using a random priming kit (Stratagene). To measure gene copy number, the relative intensities of the bands were assayed on a PhosphorImager (Molecular Dynamics).

Cosmid Cloning and Analyses of the Cosmids. Genomic DNA was partially digested with *Mbo* I, dephosphorylated, and inserted into the *Bam*HI site of a cosmid vector (SupercosI; Stratagene), using the Gigapack II XL packaging extract and NM554 host cells (Stratagene) to generate about 5×10^5 independent clones. The clones were screened with a 1.9-kb *Sal* I/*Sph* I fragment from the promoter of the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: FISH, fluorescent *in situ* hybridization; *COL1A1*, gene for proα1(I) chains of type I procollagen; *COL2A1*, gene for proα1(II) chains of type II procollagen.

*To whom reprint requests should be addressed.

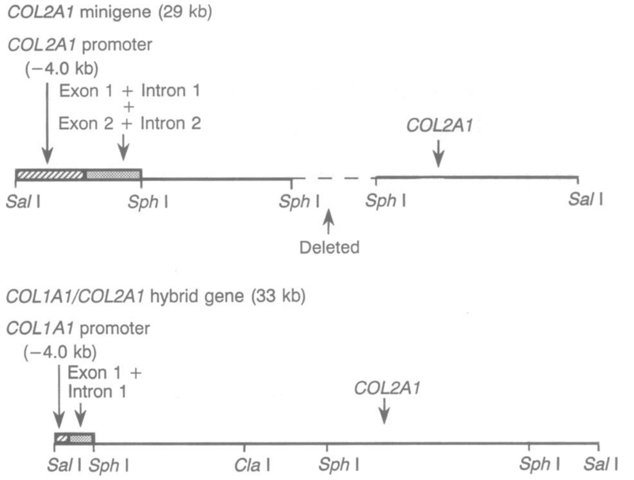

FIG. 1. Two constructs of the *COL2A1* gene. (*Upper*) The *COL2A1* minigene contained 4 kb of the *COL2A1* promoter and most of the *COL2A1* gene except for an internal deletion of 5 kb that facilitated assays for the presence and expression of the gene (24). The 3′ end was truncated so that it contained only the first two of the three polyadenylylation signals (21). (*Lower*) The *COL1A1/COL2A1* hybrid gene contained 476 bp of the human *COL1A1* promoter together with exon 1 (222 bp) and most of intron 1 (1223 bp of a total of 1453 bp) of the *COL1A1* gene (21). The *COL1A1* fragment was linked to 29.5 kb of the human *COL2A1* gene through the second intron in the *COL2A1* gene (23, 24).

COL1A1 gene. The inserts in positive clones were mapped by digestion with *Sac* I, *Pvu* II, and *Eco*RI and probing the digests with (*i*) the 1.9-kb *Sal* I/*Sph* I fragment from the 5′ end of the *COL1A1* gene (Fig. 1); (*ii*) the 3.5-kb *Sph* I/*Sal* I fragment from the 3′ end of the *COL2A1* gene (Fig. 1); (*iii*) a 2-kb *Eco*RI/*Bgl* II fragment from the mouse metallothionein promoter of the neomycin-resistance gene; and (*iv*) a 2-kb *Bam*HI/*Eco*RI fragment from the ampicillin-resistance gene present in the plasmid vector. Partial sequencing (Cyclist kit; Stratagene) was carried out with 21-mer oligonucleotide primers with sequences of the T3 promoter, the T7 promoter, and part of the *COL1A1* promoter (AGL-5; 5′-GTCCTCCTG-GCTGTGCCCCA).

Fluorescent *in Situ* Hybridization (FISH). Purified DNAs from cosmids containing either the *COL2A1* gene or the *COL2A1* minigene were labeled with biotin-14-dATP (Bio-Nick; GIBCO/BRL). The probe mixture, preparation of slides, denaturation, hybridization, and subsequent immunocytochemical detection were as described by Tkachuk *et al.* (27). Photographs were taken with Kodak Gold 100 ASA film.

Assays of Transcription with Nuclear Extracts. Assays of transcription with nuclear extracts (28) of HT-1080 cells were performed with a commercial kit (HeLa nuclear extract *in vitro* transcription system; Promega). About 100 ng of each DNA was used as template in an incubation system of 25 μl that contained 7 μl of nuclear extract. Samples were analyzed by electrophoresis on a denaturing 6% polyacrylamide gel and autoradiography.

RESULTS

Transfection and Screening of HT-1080 Cells Synthesizing Recombinant Type II Procollagen. As reported previously (23, 24), synthesis and secretion of human type II procollagen were observed in HT-1080 cells stably transfected with a neomycin-resistance gene and either the *COL2A1* minigene or the *COL1A1/COL2A1* hybrid gene construct (Fig. 1). Only about 2% of the neomycin-resistant cells synthesized the protein, but all produced relatively large amounts of the protein (Table 1). In contrast, about 88% of NIH 3T3 fibroblasts stably transfected with the *COL1A1/COL2A1* hybrid gene secreted type II procollagen but only about 3% produced relatively large amounts of the protein.

Southern Blot Analysis of the *COL2A1* and *COL1A1* Loci. Two clones of HT-1080 cells expressing the *COL2A1* minigene were assayed for *Sac* I fragments that were present near the middle of the endogenous *COL2A1* gene but absent from the *COL2A1* minigene because of the internal deletion (Figs. 1 and 2). The expected fragments were found in control cells but not in cells expressing the minigene (Fig. 2). However, assays for a 4.3-kb *Eco*RI fragment present only in the 3′ end of the endogenous gene and a 3.8-kb *Eco*RI fragment present in both genes indicated that the 3′ end of the gene was intact. Also, there was no increase in copy number of a shared 3.8-kb fragment (Table 2). Therefore, the *COL2A1* minigene was inserted into both alleles of the endogenous gene, apparently by homologous recombination through a nonconservative pathway.

HT-1080 cells expressing the *COL1A1/COL2A1* hybrid gene were analyzed with a Southern blot assay to detect unique restriction fragments that were present at the 5′ end of the endogenous *COL1A1* gene but were not present in the exogenous gene construct (Fig. 3). A 3.8-kb *Sac* I fragment found in the promoter region of the endogenous *COL1A1* gene was absent from two transfected cell lines (HTS and HTN) expressing the hybrid gene (Fig. 3 *Lower Left*). Also, a 1.4-kb *Pvu* II fragment found in the same region of the endogenous *COL1A1* gene was absent from the cells expressing the hybrid gene (Fig. 3 *Lower Right*). Similar results were obtained with two additional lines expressing the hybrid gene (not shown). Hybridization of the same filters with a probe for an internal 5.4-kb *Eco*RI fragment of the *COL2A1* gene demonstrated that the observations were not explained by artifactual differences in migration of the same DNA fragments (not shown). The results indicated, therefore, that the 5′-flanking sequences of both alleles of the endogenous *COL1A1* gene had been disrupted in all four clones expressing the *COL1A1/COL2A1* hybrid gene.

Table 1. Expression of the *COL2A1* gene in stably transfected cells

		Stably transfected clones		
Host cells	Invading human gene	No. clones assayed	% low expression*	% high expression*
Human HT-1080 cells	*COL2A1* minigene	100	0	2.0†‡
	COL1A1/COL2A1 hybrid gene	314	0	1.9§¶
Mouse 3T3 fibroblasts	*COL1A1/COL2A1* hybrid gene	116	85	3.4

*Clones were screened by Western blot assays of culture media with polyclonal antibodies specific for procollagen II (21) and equal amounts of protein loaded in each lane. The intensities of the signals from high-expressing clones were 3- to 5-fold greater than those from clones expressing low levels (for illustrative data, see refs. 21 and 23).
†Data from Sieron *et al.* (24).
‡Two of two clones tested had targeted insertion into the *COL2A1* locus (see Fig. 2 and Table 2).
§Data from Fertala *et al.* (23).
¶Four of four clones tested had targeted insertion in the *COL1A1* locus (see Fig. 3 and Table 2).

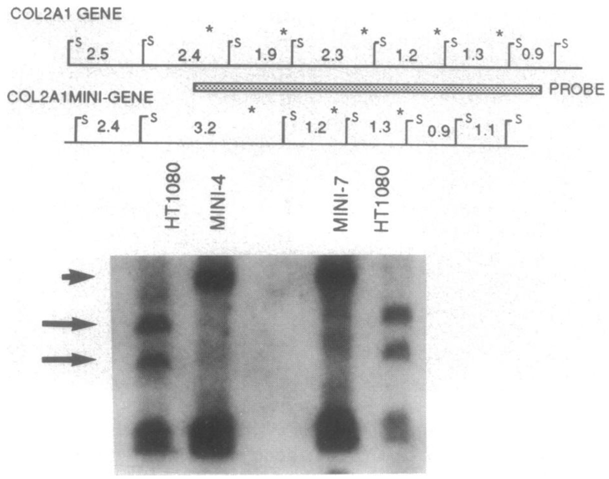

FIG. 2. Southern blot analysis of two clones of cells expressing the *COL2A1* minigene. (*Upper*) Restriction map of the endogenous *COL2A1* gene and the *COL2A1* minigene. The *Sac* I restriction sites are indicated and fragment sizes are in kb. Asterisks indicate fragments that hybridized with the probe that consisted of a 10-kb *Eco*RI fragment from the middle of the *COL2A1* gene (Fig. 1). (*Lower*) Southern blot of *Sac* I digest of genomic DNA from nontransfected HT-1080 cells and two cloned cell lines expressing the *COL2A1* minigene (MINI-4 and MINI-7). In the nontransfected HT-1080 cells, the expected 1.9-kb *Sac* I fragment is seen together with a broad band containing both the 2.3- and 2.4-kb *Sac* I fragments. The smaller *Sac* I fragments of 0.9–1.3 kb comigrated as a broad band at the bottom of the gel.

Analysis of a Cosmid Clone from a Cell Line Expressing the *COL1A1/COL2A1* Hybrid Gene. Because the results suggested an unusual mode of insertion of the *COL1A1/COL2A1* hybrid gene, cosmid clones containing the inserted gene were isolated from the HTN line. One clone of interest (11A1) contained an insert of about 38 kb (Fig. 4). Sequencing of about 400 bp from the 3′ end of the insert identified sequences that were present in the 3′ end of *COL1A1/COL2A1* gene construct. Sequencing of about 400 bp from the 5′ end of the insert identified *Alu* sequences that were not distinctive. However, a *Sac* I/*Sac* I fragment of 3 kb from near the 5′ end of the insert specifically hybridized to a 3.8-kb *Sac* I/*Sac* I fragment that was found in the 5′-flanking region of the endogenous *COL1A1* gene but that was not present in the original *COL1A1/COL2A1* gene construct. Sequencing with an antisense primer for the 5′ end of the *COL1A1/COL2A1* gene construct together with additional Southern blot analyses (not shown) demonstrated that the insert consisted of (*i*) a fragment from the 5′-flanking sequences of the endogenous *COL1A1* gene that was not present in the gene construct used to transfect the cells; (*ii*) a single copy of the plasmid containing the neomycin-resistance gene in the reverse orientation; (*iii*) a junction region that contained 30 bp from the

Table 2. FISH analysis of transfected HT-1080 cells

Invading human gene	Cell line*	Copy no.	Chromosomal locations: Collagen loci	Chromosomal locations: Other†
COL2A1 minigene	MINI-4	2‡	Both *COL2A1* alleles	One site on C group
	MINI-7	2‡	Both *COL2A1* alleles	One site on E group
COL1A1/COL2A1 hybrid gene	HTN	2§	Both *COL1A1* alleles	One site on Ch 3
	HTS	≈10§	Both *COL1A1* alleles	One site on C group

*Cell lines MINI-4 and MINI-7 were from the same transfection experiments. The other two lines were from independent transfections.

†Additional signals on chromosome 3 (Ch 3) in the HTN line were from additional integrations of the neomycin-resistance gene only (see text).

‡Total copies of the *COL2A1* gene as assayed by detection of a 4.3-kb *Eco*RI fragment present only at the 3′ end of the endogenous *COL2A1* gene and a 3.8-kb *Eco*RI fragment present in both the endogenous and exogenous gene. The observed ratio was 1:1.

§Copies of the exogenous gene only as assayed by the 4.3-kb *Eco*RI fragment present only in the endogenous *COL2A1* gene and a unique 6.6-kb fragment generated by an *Eco*RI site at the 3′ end of the exogenous *COL2A1* gene and another apparently in linked sequences from the cosmid vector (see Fig. 4). The observed ratio was 1:1. A 1:1 ratio was also observed with a 1.9-kb *Sal* I/*Sph* I probe (Fig. 1) that detected a 9-kb *Eco*RI fragment from the 5′ end of the inserted gene and a second *Eco*RI fragment of about 20 kb from the endogenous gene (Fig. 4).

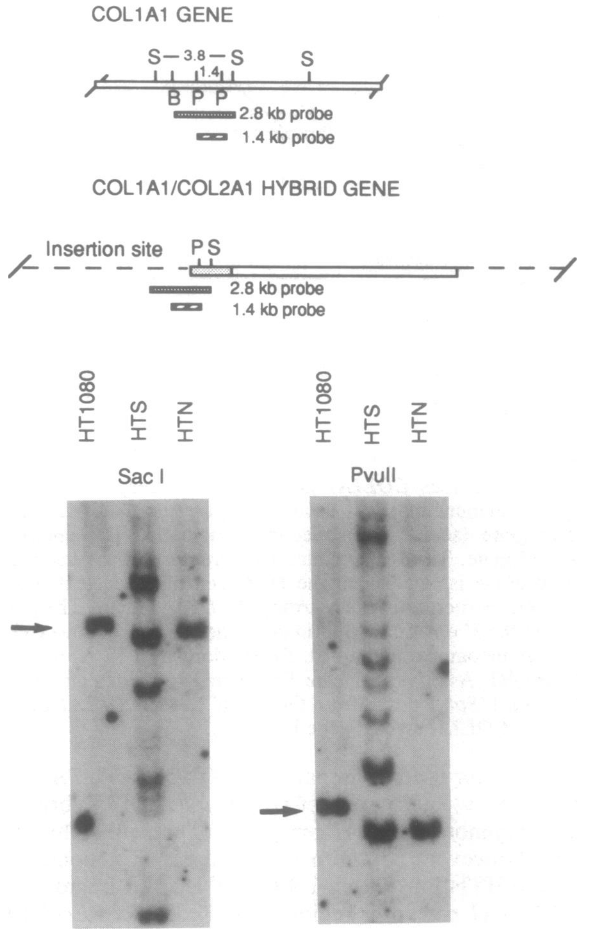

FIG. 3. Southern blot analysis of two cloned cell lines expressing *COL1A1/COL2A1* hybrid gene. (*Upper*) Restriction map for the endogenous *COL1A1* gene and the *COL1A1/COL2A1* hybrid gene. S, *Sac* I; B, *Bam*HI; P, *Pvu* II. (*Lower Left*) The 2.8-kb *Bam*HI/*Sac* I probe detected a 3.8-kb *Sac* I fragment from the endogenous *COL1A1* gene in nontransfected HT-1080 cells but not in two clones expressing the *COL1A1/COL2A1* hybrid gene. In the HTN line with a copy number of 2 (Table 2), the probe hybridized only with a slightly shorter *Sac* I fragment spanning one *Sac* I site in the displaced cellular flanking sequences and a second *Sac* I site at the 3′ end of the neomycin-resistance gene (see Fig. 4). In the HTS line with a copy number of about 10 (Table 2), the probe hybridized with additional fragments from multiple copies of the exogenous genes. (*Lower Right*) The 1.4-kb *Pvu* II probe detected a 1.4-kb *Pvu* II fragment in nontransfected cells but not in cells expressing the hybrid gene construct. The shorter *Pvu* II fragment in the HTN line was generated by one *Pvu* II site in the displaced flanking sequences and a second in the 3′ end of the pDMMT gene (Fig. 4).

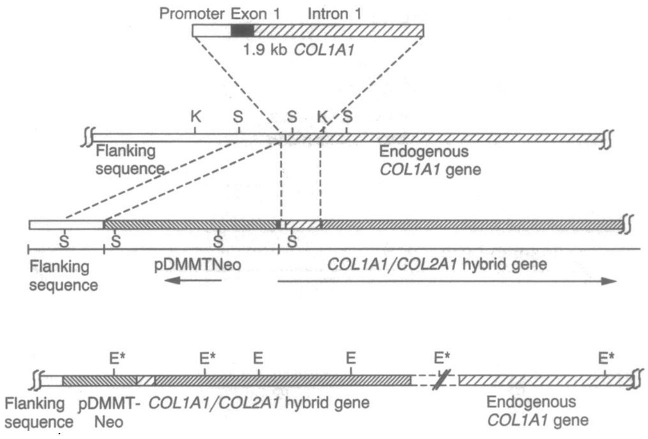

FIG. 4. First (top) schematic. The 1.9-kb fragment from the 5′ end of the endogenous *COL1A1* gene, which corresponds to the *Sal* I/*Sph* I fragment that was used to prepare the *COL1A1/COL2A1* hybrid gene (see Fig. 1). Second schematic. The endogenous *COL1A1* gene. Third schematic. The structure of the insert in the cosmid clone isolated from the HTN cell line (see text). Bottom schematic. Structure of the rearranged *COL1A1* locus. The junction between the 3′ end of the hybrid gene and the endogenous *COL1A1* gene was incompletely defined. Restriction sites: K, *Kpn* I; S, *Sac* I; E, *Eco*RI. Asterisks indicate *Eco*RI fragments detected with the 1.9-kb *Sal* I/*Sph* I probe from the 5′ end of the *COL1A1* gene (see *COL1A1/COL2A1* hybrid gene in Fig. 1).

cosmid vector used to prepare the *COL1A1/COL2A1* hybrid gene; and (*iv*) most of the *COL1A1/COL2A1* hybrid gene. The endogenous *COL1A1* gene was not present in the cosmid insert. However, Southern blot analysis of genomic DNA from the HTN line with a 1.9-kb *Sal* I/*Sph* I fragment from the *COL1A1* promoter region (Fig. 1) detected *Eco*RI fragments of 9 kb and about 20 kb of equal intensity. The 9-kb fragment spanned the junction between the neomycin-resistance gene plasmid and the hybrid gene (Fig. 4). The fragment of about 20 kb apparently arose from one *Eco*RI site at the 3′ end of endogenous *COL1A1* and a second *Eco*RI in cosmid vector sequences linked to the 3′ end of the hybrid gene construct (bottom schematic in Fig. 4).

FISH Analysis for the Integration Sites of the Exogenous Genes. FISH analysis on the cell line HTN demonstrated integration of the exogenous *COL1A1/COL2A1* hybrid gene into both copies of chromosome 17 and near the q21 locus of the *COL1A1* gene (Fig. 5). The integration into chromosome 17 was further confirmed by double-labeling for FISH analysis with specific α-satellite DNA probes (Oncor) for the centromere of chromosome 17 (not shown). As expected, there also was a signal at the endogenous locus of the *COL2A1* gene at 12q14. In addition, there was a prominent signal on one copy of chromosome 3 that probably arose from multiple integrations of the neomycin-resistance gene. This conclusion was confirmed by additional Southern blot assays with a probe consisting of the pDMMTNeo plasmid, which contains the neomycin-resistance gene (not shown). Similar results were obtained with the HTS line, which also expressed the *COL1A1/COL2A1* hybrid gene (Table 2). Two cell lines expressing the *COL2A1* minigene showed a signal on both copies of chromosome 12 near the q14 locus of *COL2A1* (Table 2) but no signal at the 17q22 locus of *COL1A1*. Additional signals were seen on one chromosome of the C group (chromosomes 6–12) in one of the cell lines (MINI-4) and on one chromosome of the E group (chromosomes 16–18) from a second cell line (MINI-7).

Transcription of the *COL1A1/COL2A1* Hybrid Gene by Nuclear Extracts. In transcription assays with nuclear extracts from HT-1080 cells, the construct of the *COL1A1/COL2A1* hybrid gene generated a full-length nuclear runoff

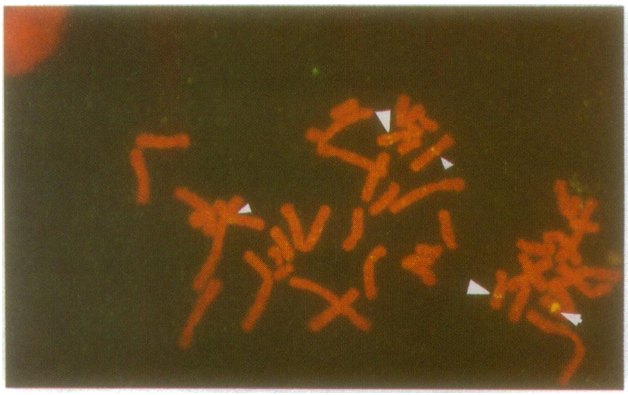

FIG. 5. FISH analysis of cells from the HTN cell line that expressed the *COL1A1/COL2A1* gene and that had a copy number of 2. The probe was a cosmid clone containing the *COL2A1* gene. Signal was detected in both alleles of chromosome 12 at about the q14 locus of the endogenous *COL2A1* gene (medium arrowhead) and on both copies of chromosome 17 at about the q21 locus of the *COL1A1* gene (large arrowhead). The signal in one copy of chromosome 3 (arrowhead with short shaft) was from inserts of the plasmid containing the neomycin-resistance gene (see text).

transcript (lanes 5 and 6 in Fig. 6). Also, the cosmid clone (11A1) isolated from the HTN cell line (Fig. 4) generated a full-length runoff transcript of the hybrid gene (lanes 3 and 4 in Fig. 6). However, no transcript was detected from the neomycin-resistance gene that was found in the same cosmid (Fig. 4). In contrast, a transcript was obtained when the neomycin-resistance gene was present in a 6-fold greater molar excess (lane 2 in Fig. 6).

Cytogenetic Analysis of Cloned Lines. High-resolution G-banding of the line HTN (Table 2) indicated there were 42–46 chromosomes per cell, and several structural alternations were consistently found: $5p^+$ (1qter→q42.1::5p15.3→5qter), and $11q^+$ (11pter→11q24.2::3q13.2→3qter), and iso13q (+13). In the line HTS, the chromosome number ranged between 40 and 46 in nine cells and between 90 and 93 in two cells. Several structural alterations were consistent in every metaphase analyzed: 5p+(?5qter→5p15.3::1q42.1→1qter or 15q22→15qter), $11q^+$ [der(11)t(3;11)(3q13.2;11q25)], and $8p^+$ [t(8;14)(8qter→8p22::14q13→14qter)]. Three cells

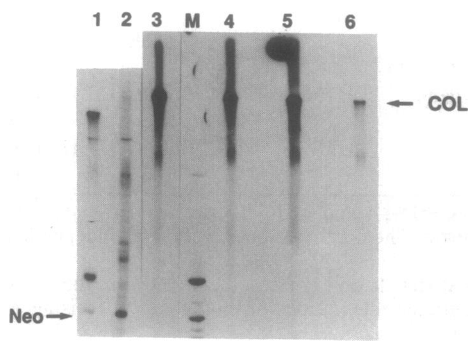

FIG. 6. Autoradiogram of *in vitro* transcription assay of the gene constructs, using HT-1080 nuclear extract. Lane 1, transcription of 100 ng of control DNA containing the cytomegalovirus (CMV) promoter and a CMV gene fragment (Promega); lane 2, transcript of 100 ng (20 fmol) of the pDMMTNeo plasmid; lane 3, transcript with 100 ng (3 fmol) of cosmid clone 11A1 obtained from the HTN cell line (see Fig. 4); lane M, DNA marker fragments; lane 4, duplicate of sample in lane 3; lane 5, transcript with 100 ng (3 fmol) of *COL1A1/COL2A1* hybrid construct; lane 6, second clone (100 ng) of the *COL1A1/COL2A1* hybrid construct. Arrows, full-length transcripts of the neomycin-resistance gene (Neo) and the *COL1A1/COL2A1* hybrid gene construct (COL).

had an iso13q, two cells an iso13p, and two cells a $13p^+$ which was probably a (15pter::13pter→13q22:).

DISCUSSION

In the experiments described here with the human tumor cells HT-1080, we observed both targeted insertion of two collagen gene constructs and expression of the gene constructs from previously inactive loci.

The yield of stably transfected HT-1080 cells was about 1 per 30,000 and, therefore, within the range of yields usually obtained with a neomycin-resistance gene and the protocol of calcium phosphate precipitation (25). The exogenous *COL2A1* minigene was inserted into the endogenous locus for the *COL2A1* gene on chromosome 12 in about 2% of the G418-resistant clones. Comparable frequencies of targeted gene insertion are commonly seen in experiments with embryonic stem cells, particularly if genes for both positive and negative selection are employed (see ref. 16). Insertion into both alleles of a locus is rare, but embryonic stem cells with an insert containing an antibiotic-resistance gene in one allele can be used to generate clones with insertions into both alleles by incubating the cells with increasing concentrations of the antibiotic (16). Therefore, the observations with the *COL2A1* minigene can be explained by targeted insertion into the *COL2A1* locus on chromosome 12 through homologous recombination by a nonconservative pathway, and then conversion of the second allele to homozygosity through gene conversion or other chromosomal rearrangement. It was unexpected, however, that the targeted insertion was directed to the *COL1A1* locus on chromosome 17 in stably transfected cells expressing a similar construct of the *COL2A1* gene in which the 5′ end was replaced with a relatively short fragment of the *COL1A1* gene. Only 1.9 kb of the *COL1A1/COL2A1* hybrid gene was homologous with the *COL1A1* locus, whereas 30 kb of its sequences were homologous with the *COL2A1* locus. Also, insertion of the hybrid gene disrupted the 5′-flanking sequences of the endogenous *COL1A1* gene, but the remainder of the endogenous gene was largely intact. Therefore, it is difficult to explain the selective targeting of the hybrid gene construct simply on the basis of homologous recombination. Instead, the results suggest the hypothesis that the gene was transcribed after it entered the nucleus and that the presence of a transcript or a transcriptionally active region targeted the construct for insertion into the *COL1A1* locus. This hypothesis is consistent with the observation that the *COL1A1/COL2A1* hybrid gene was transcribed by nuclear extracts of the HT-1080 cells efficiently and at a higher rate than the neomycin-resistance gene driven by the mouse metallothionein promoter. The hypothesis is also consistent with the observations that recombination of genes is increased by transcriptional activation of the genes in *Escherichia coli* (1), in yeast (4–6), and in some mammalian systems (15, 17).

The targeted insertion observed here probably reflects one or more fortuitous features of the gene constructs and the host cell line. The HT-1080 cell line used as a host was derived from a fibrosarcoma (29) and has been widely employed for experimental work (see refs. 30 and 31). The cell line has a modal number of 46 chromosomes and, as confirmed here, a minimal number of chromosomal rearrangements. None of the readily apparent chromosomal rearrangements involved the *COL1A1* and *COL2A1* loci. One of the most obvious characteristics of the HT-1080 cells, however, was that only a small fraction of the G418-resistant transfectants expressed the cotransfected *COL2A1* genes. Therefore, the gene constructs apparently were not expressed if integrated into any of a large number of sites from which the neomycin-resistance gene was expressed.

In addition to the targeted insertions, there were two unexpected observations here. One was that the two gene constructs were expressed from loci that were previously inactive. Apparently, the targeted insertions of the constructs either disrupted inhibitory cis-regulatory elements or produced more subtle conformational changes in the loci. The further unexpected observation was that the organization of the exogenous DNA sequences in the cosmid isolated from a transfected cell line (Fig. 4) was not consistent with the most common modes of gene insertion or replacement by homologous recombination (2, 3, 11, 12).

Obviously, it will be of interest to see whether strategies similar to those we have employed here can be used to target genes to other loci in the HT-1080 cells or in other mammalian cells.

The authors thank Dr. Margaret M. Aronson for cytogenetic analysis of the HT-1080 cells. They also thank Mary Vinson and Hazel Clouston for assistance in some of the work. The work was supported in part by National Institutes of Health Grants AR-38188 and AR-39740, and a grant from the Lucille P. Markey Charitable Trust.

1. Ikeda, H. & Matsumoto, T. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4571–4575.
2. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. (1983) *Cell* **33**, 25–35.
3. Cox, M. M. & Lehman, I. R. (1987) *Annu. Rev. Biochem.* **56**, 229–262.
4. Voelkel-Meiman, K., Keil, R. L. & Roeder, G. S. (1987) *Cell* **48**, 1071–1079.
5. Thomas, B. J. & Rothstein, R. (1989) *Cell* **56**, 619–630.
6. Grimm, C., Schaer, P., Munz, P. & Hohli, J. (1991) *Mol. Cell. Biol.* **11**, 289–298.
7. Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. (1985) *Nature (London)* **317**, 230–234.
8. Alt, F. W., Blackwell, T. K., Depinho, R. A., Reth, M. G. & Yancopoulos, G. D. (1986) *Immunol. Rev.* **89**, 5–30.
9. Jasin, M. & Berg, P. (1988) *Genes Dev.* **2**, 1353–1363.
10. Capecchi, M. R. (1989) *Science* **244**, 1288–1292.
11. Bollag, R. J., Waldman, A. S. & Liskay, R. M. (1989) *Annu. Rev. Genet.* **23**, 199–225.
12. Lin, F.-L., Sperle, K. & Sternberg, N. (1990) *Mol. Cell. Biol.* **10**, 103–112.
13. Hasty, P., Rivera-Perez, J., Chang, C. & Bradley, A. (1991) *Mol. Cell. Biol.* **11**, 4509–4517.
14. Shenkar, R., Shen, M. H. & Arnheim, N. (1991) *Mol. Cell. Biol.* **11**, 1813–1819.
15. Nickoloff, J. A. (1992) *Mol. Cell. Biol.* **12**, 5311–5318.
16. Mortensen, R. (1993) in *Current Protocols in Molecular Biology*, ed. Janssen, K. (Greene and Wiley, New York), pp. 9.15.1–9.17.2.
17. Kotani, H. & Kmiec, E. B. (1994) *Mol. Cell. Biol.* **14**, 1949–1955.
18. Schnieke, A., Dziadek, M., Bateman, J., Mascara, T., Habers, K., Gelinas, R. & Jaenisch, R. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 764–768.
19. Greenspan, D. S., Hoffman, G. G. & Lee, B. S. (1989) *J. Biol. Chem.* **264**, 20683–20687.
20. Olsen, A. S., Geddis, A. E. & Prockop, D. J. (1991) *J. Biol. Chem.* **266**, 19878–19881.
21. Ala-Kokko, L., Hyland, J., Smith, C., Kivirikko, K. I., Jimenez, S. A. & Prockop, D. J. (1991) *J. Biol. Chem.* **266**, 14175–15178.
22. Prockop, D. J. & Kivirikko, K. I. (1984) *N. Engl. J. Med.* **311**, 376–386.
23. Fertala, A., Sieron, A. L., Ganguly, A., Li, S.-W., Ala-Kokko, L., Anumula, K. R. & Prockop, D. J. (1994) *Biochem. J.* **298**, 31–37.
24. Sieron, A. L., Fertala, A., Ala-Kokko, L. & Prockop, D. J. (1993) *J. Biol. Chem.* **268**, 21232–21237.
25. Kingston, R. E., Chen, C. A. & Okayama, H. (1993) in *Current Protocols in Molecular Biology*, ed. Janssen, K. (Greene and Wiley and Sons, New York), pp. 9.1.1–9.1.9.
26. Law, M. F., Byrne, J. C. & Howley, P. M. (1983) *Mol. Cell. Biol.* **3**, 2110–2115.
27. Tkachuk, D. C., Pinkel, T., Kuo, W.-L., Weier, H. U. & Gray, J. P. (1991) *Gene Anal. Tech. Appl.* **8**, 67–74.
28. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475–1489.
29. Rasheed, S., Nelson-Rees, W. A., Toth, E. M., Arnstein, P. & Gardner, M. B. (1974) *Cancer* **33**, 1027–1033.
30. Perry, R. P., Kelley, D. E., Schibler, U., Huebner, K. & Croce, C. M. (1979) *Cell. Physiol.* **98**, 553–560.
31. Pihlajaniemi, T., Myllylä, R., Alitalo, K., Vaharo, A. & Kivirikko, K. I. (1981) *Biochemistry* **20**, 7409–7415.