Matrix Biology Vol. 14/1994, pp. 297–305

Molecular Basis of Nanomelia, a Heritable Chondrodystrophy of Chicken[1]

D. PRIMORAC*,[2] M. L. STOVER*, S. H. CLARK† and D. W. ROWE*

* Department of Pediatrics and Division of Rheumatic Diseases, Department of Medicine, The University of Connecticut Health Center, Farmington, Connecticut and
† Veterans Affairs Medical Center, Newington, Connecticut, USA

Abstract

Nanomelia is a recessively inherited connective tissue disorder of chicken affecting cartilage development. Other investigators have demonstrated that it involves low aggrecan production and diminished aggrecan mRNA levels. Based on genetic linkage studies showing a high likelihood that the mutation responsible for the nanomelic phenotype lay within the aggrecan gene, a series of experiments was performed to define the molecular basis of the trait. Aggrecan mRNA was present in the nucleus of the nanomelic chondrocyte but greatly reduced in the cytoplasmic compartment, a finding suggestive of a premature stop codon within the aggrecan transcript. Since no defect in mRNA splicing could be demonstrated by ribonucleasease protection studies, direct DNA sequencing was initiated by polymerase chain reaction of the mRNA and of genomic DNA. A stop codon was demonstrated at codon 1513, which is located in the eighth repeat of the chondroitin sulfate 2 domain of the large tenth exon. The mutation creates a unique *Bas*BI restriction site which readily distinguishes the mutant and wild-type alleles.

Key words: aggrecan, cartilage, chondrodystrophy, nanomelia.

Introduction

Naturally occurring mutations affecting limb formation in vertebrates are important models for studying a complex developmental process. Nanomelia is one such example and has illustrated the importance of the cartilage proteoglycan, aggrecan, to cartilage biology and limb development. This mutation was described by Landauer (1965) as a recessive trait of chicken associated with shortened limbs and beak, underdevelopment of the chest wall and death around 17–19 days of embryonic development. Defective synthesis of cartilage chondroitin sulfate (CS)[3] proteoglycan and the absence of proteolycan core protein were demonstrated in nanomelic embryos (McKeown and Goetinck, 1979; Argraves et al., 1981). With the cloning of chick aggrecan cDNA (Sai et al., 1986), it was shown that the content of aggrecan mRNA from nanomelic chondrocytes was approximately 6% of normal and that the size of the RNA transcript was normal (Stirpe et al., 1987). Subsequently, antibodies directed against the N-terminal hyaluronic acid-binding domain of aggrecan identified an immunoreactive product from nanomelic chondrocytes with a molecular mass of 300 kDa rather than the expected 370 kDa (O'Donnell et al., 1988). The authors postulated that nanomelia resulted from premature translational termination, although they pointed out that the immunoreactive product detected in nanomelic chondrocytes might re-

[1] Parts of this work were presented at the American Society of Bone and Mineral Research in Minneapolis, MN on September 30, 1992 and the East Coast Connective Tissue Society, Philadelphia, PA on March 20, 1993.

[2] Permanent address: University of Zagreb, School of Medicine at Split, 58000 Split, Croatia. This work was done in partial fulfillment of D. P's Ph. D. requirement at the University of Zagreb.

[3] Abbreviations used: CS, chondroitoin sulfate; RNase, ribonuclease; PCR, polymerase chain reaction; RT, reverse transcription; RSB, reticulocyte swelling buffer; SET, buffer containing SDS, EDTA and Tris; TAE, buffer containing Tris, sodium acetate and EDTA.

present another cross-reacting core protein and not necessarily an authentic but truncated aggrecan protein.

Before we began to define the molecular basis for this mutation, it was important to show that the mutation most likely lay within the aggrecan gene and not in another gene which acted in trans to regulate its activity. Thus, the study carried out by Velleman and Clark (1992) demonstrating genetic linkage of the nonomelia mutation to the aggrecan gene was fundamental to undertaking this project. First we demonstrated that the mutant gene was as transcriptionally active as the wild type since normal levels of the RNA product accumulated within the nucleus. However, its content within the cytoplasm was greatly reduced. Failure of cytoplasmic accumulation of a specific mRNA has been associated with a premature termination of translation secondary either to a point mutation causing premature termination or a frame shift mutation causing formation of a downstream termination codon (Maquat and Kinniburgh, 1985; Urlaub et al., 1989). The latter possibility might result from either an error of splicing or a base insertion or deletion. Ribonuclease (RNase) A protection experiments were performed to show that the aggrecan RNA within the nucleus was co-linear with normal cDNA, thus excluding the possibility of a splicing defect causing the abnormality. DNA sequencing from the 3′ end of the gene revealed a point mutation resulting in a premature stop codon within the highly repetitive CS2 binding domain. The same mutation has recently been reported by Li et al. (1993).

Materials and Methods

Cell culture

Fertile eggs were obtained from normal white leghorn breeding stocks and nanomelic breeding stocks maintained at the Storrs Agricultural Experimental Station and incubated for 17 days. Sterna from affected and normal embryos were dissected free of perichondrium and incubated in Ham's F12 medium (Gibco) containing 0.1% collagenase CLS2 (Worthington Biochemical) for 4 h. Liberated cells were plated at a density of 10^6 cells/100 mm petri plate in 9 ml of Ham's F 12 medium containing 20% fetal calf serum and incubated at 37 °C in humidified atmosphere at 7.5% CO_2.

Medium was changed every second day. The cells were split at a 1:2 ratio at 5–7 days of culture and grown to confluency for an additional 5–7 days. In a typical experiment, six nanomelic sterna were used to generate 12 100-mm petri plates of chondrocytes. Two days prior to harvesting cells for RNA extraction, the culture medium was supplemented with 50 μg/ml ascorbic acid.

RNA Isolation

Confluent chondrocytes were scraped from the petri dish with a rubber policeman into 10 ml of cold phosphate-buffered saline and collected by centrifugation for 3 min at 2000 × g in a 15-ml Falcon tube. The cell pellet was resuspended in 10 ml of cold reticulocyte swelling buffer (RSB: 10 mM Tris, pH 7.5, 10 mM KCl, 1.5 mM $MgCl_2$) containing 0.25% Triton X-100 and dispersed by Dounce homogenization using a B size pestle. After the cells were allowed to lyse on ice for 5 min, they were rehomogenized with 5–10 strokes and checked microscopically for extent of cell rupture and quality of nuclear morphology. Further Dounce homogenization was done if the cells were not adequately lysed. The isolated nuclei were collected in the same 15-ml Falcon tube by centrifugation for 5 min at 2000 × g. The supernatant was transferred to a 30-ml Corex tube (cytoplasmic extract) containing 1.1 ml of 10 × SDS, EDTA and Tris (SET) buffer (10.0% SDS, 0.5 M Tris HCl pH 7.5, 10 mM EDTA) and 110 μl proteinase K (10 mg/ml). The nuclear pellet was re-extracted in 5 ml of RSB containing 0.25% Triton X-100 with brief Dounce homogenization. The nuclei were again separated by centrifugation and extracted in 1 × SET containing 100 μl/ml proteinase K. The RNA was isolated from both fractions as previously described (Genovese et al., 1984). The yield of total RNA from 12 confluent plates of chondrocytes was 21–30 μg from the nuclear fraction and 200–300 μg from the cytoplasmic compartment.

Genomic DNA extraction

Minced leg muscle tissue (0.5–1.0 gm) was placed in a 10-ml Falcon tube containing 5.0 ml 2 × SET and 200 μg/ml of proteinase K and rocked over night at 55 °C. The solubilized tissue was extracted in an equal volume of phenol/chloroform/isoamyl alcohol (12:12:0.5 by volume), centrifuged 5 min and the aqueous portion re-extracted in an equal volume of chloroform:isoamyl alcohol (24:1 by volume). The aqueous portion was transferred to a fresh tube, adjusted with ½ volume 7.5 M amonium acetate, incubated at room temperature for 10 min and centrifuged at 10000 × g for 10 min. The supernatant was transferred to a fresh tube and layered with an equal volume of 100% EtOH. DNA was spooled from the solution with a sealed Pasteur pipette, washed twice with 70% EtOH, air dried and redissolved in 1.0 ml Tris-EDTA buffer, pH 8.0.

RNase protection

Three types of RNase protection were utilized to characterize the quantity and structure of the aggrecan RNA in normal and nanomelic chondrocytes. Direct RNase protection using short (~ 350 bp) probes was employed to quanti-

tate RNA, while longer (~ 1.7 kb) cRNA probes were used to screen for a skipped exon (see Fig. 1). In either case, the probe was cloned in Bluescript SK^+ (Stratagene) and transcribed with T_3 or T_7 RNA polymerase in the presence of [P^{32}]UTP to produce a uniformly labeled antisense cRNA probe. The probe was mixed with 6 µg of nuclear or cytoplasmic RNA, precipitated with ethanol, redissolved in 80% formamide, 0.4 M NaCl, 40 mM Hepes, pH 7.6 and 2 mM EDTA and incubated at 50 °C over night. The solution was digested with a mixture of T_2 and T_1 RNase as previously described (Lichtler et al., 1992). Each reaction mixture was phenol/chloroform extracted, ethanol precipitated and electrophoresed in an 8% polyacrylamide gel containing 8 M urea. RNase resistant bands were identified by autoradiography with Kodak X-OMAT AR film.

Indirect RNase protection was utilized to screen aggrecan RNA for an insertion (Genovese et al., 1989). Single-stranded antisense DNA probes were generated from the same Bluescript SK^+ clones by co-culture with 10^7-pfu/ml of VCSM13 helper phage (Veira and Messing, 1987). Once the infection has been established, cells harboring the helper were selected in 25 µg/ml kanamycin over night, followed by harvesting the single-stranded cDNA by polythylene glycol precipitation. The indirect RNase protection was performed as previously described using 4 µg RNA, 1 µg single-stranded cDNA and cloned RNase A at 0.05 µg/ml, 37 °C for 30 min (Genovese et al., 1989).

Polymerase chain reaction (PCR) of RNA and genomic DNA

Two µg of nuclear or cytoplasmic RNA was transcribed by reverse transcription (RT) to cDNA using 2 µg of a core protein-specific primer that extended from codon 2100 (5′ GGGTGGGTCTGTGCACGACACCGGGT)[4] in 50 mM Tris, pH 8.3, 7.0 mM KCl, 3 mM $MgCl_2$ and 10 mM DTT. The mixture was heated at 65 °C for 3 min, cooled and adjusted to contain 500 mM dNTPs and 500 U M-MLV RT (BRL). The incubation lasted for 1 h at 37 °C, followed by phenol/chloroform extraction and ethanol precipitation. The sample was dissolved in 20 µl water, and 1 µl was used in the PCR reaction.

The region encompassing the CS2 repetitive domain (codons 1363–1762) was PCR amplified using oligonucleotides that flank the region (codon 1327: 5′ GCCAACAGTTTCACAGGAG and codon 1735: 5′ GCTAACAGCAGGCAGCCCAG, see Fig. 1). The PCR mixture contained 50 mM KCl, 10 mM Tris, pH 8.3, 1.5 mM $MgCl_2$, 0.01% gelatin and 2.5 U Taq DNA polymerase (BRL) in a final reaction volume of 84 µl. The mixture was heated to 94 °C for 5 min., after which 16 µl of 1.25 mM dNTPs was added (Chow et al., 1992), and the sample was cycled 30 times at 94 °C for 15 s, 65 °C for 15 s and 72 °C for 90 s in the Perkin Elmer 9600 Thermocycler. The PCR products were visualized in a 1% agarose TAE gel stained with ethidium bromide. Genomic DNA was amplified with the same primers and reaction conditions as the aggrecan cDNA.

The PCR products were excised from 1% TAE agarose and extracted with a Qiaex solution (Qiagen, Chatsworth, CA). The purified DNA was cut with the restriction enzyme, *Sst* II, cloned into Bluescript SK II (Stratagene) and sequenced by the dideoxy method employing Sequenase (US Biochemical Corporation). The *Bsa*B1 restriction fragment polymorphism that resulted from the nonomelia mutation was detected in PCR-amplified genomic DNA using the same primers that flanked the CS2 repetitive domain (see above). The product was cut with the BsaB1 restriction enzyme for one h at 60 °C, electrophoresed in 5% acrylamide and visualized with ethidium bromide staining.

Results

Fig. 1 shows the composite structure of the gene and the probes that were generated from cDNA and genomic clones generously provided by Drs. Chandrasekaran and Tanzer (Tanaka et al., 1988; Chandrasekaran and Tanzer, 1992). The first nine exons have not been mapped in chicken but are assumed to exist based on similarity with the mouse and human aggrecan gene (Barry, 1990; Doege et al., 1990). Not included within the published cDNA sequence of the chicken is the alternatively spliced epidermal growth factor-like exon located within intron 10 (Baldwin et al., 1989; Doege et al., 1991; Grover and Roughley, 1993). The aggrecan gene is unusual in that exon 10 is an exceptionally large internal exon (2.8–3.9 kb depending on species). It contains two highly repetitive serine-rich regions that encode the site of CS attachment, designated the CS1 and CS2 domains (Dudhia et al., 1990; Neame, 1990).

The published work of Stirpe et al. (1987) and O'Donnell et al. (1988), in conjunction with the genetic linkage analysis by Velleman and Clark (1992), strongly suggested that in the nanomelic mutation a premature stop codon was present within the aggrecan gene and accounted for the low levels of RNA and the truncated immunoprecipitable protein product. To determine if the distribution of aggrecan RNA between the nuclear and cytoplasmic compartments was consistent with this possibility, we assessed its presence in both compartments by direct RNase protection using a XmnI subfragment of probe H which corresponds to exon 13 and 14 (Fig. 1). Fig. 2 shows the aggrecan cRNA probe of 380 nt (lane 9) that protects an mRNA fragment of

[4] The two published sequences of the chicken aggrecan gene differ in their nucleotide numbering because of different lengths of the 5′ untranslated region. Because they agree at the level of the translated protein sequence, codon number is used to localize sequence.

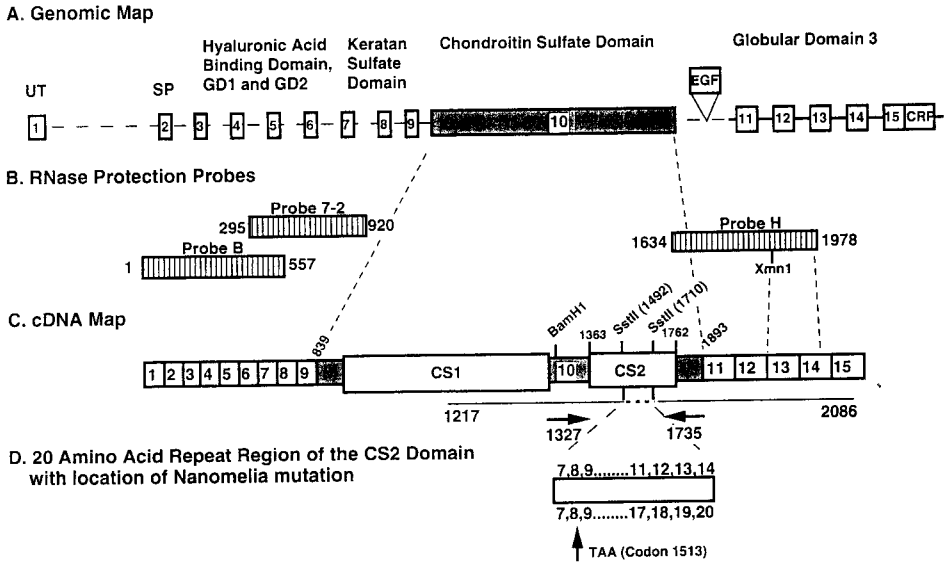

Fig. 1. Maps of the Aggrecan Gene and cDNA. (A) The tentative genomic map of the chicken aggrecan DNA is based on the known chicken sequence and on homology with the human and rat gene. The primary sequence of the chicken gene from exon 11 through 15 is known from genomic sequencing (Tanaka, 1988), and the large exon 10 was obtained by genomic sequencing (Chandrasekaran and Tanzer, 1992). Exons 1–9 were obtained by sequencing a sequencing a cDNA fragment obtained from a lambda library (Chandrasekaran and Tanzer, 1992; Li et al., 1993). The untranslated first exon (UT), signal peptide (SP), alternatively spliced epidermal growth factor-like (EGF) domain, complement regulatory protein-like (CRP) domain and three globular domains (GD) are illustrated in the diagram. (B) The three probes used in RNase protection experiments were derived from cDNA fragments that span exons 1–8 (probe B), another that extends from exon 6 through approximately 300 bp of exon 10 (probe 7–2) and a fragment that begins in the final 140 bp of exon 10 and covers exons 11–14 (probe H). A subfragment of probe H spanning exons 13 and 14 was produced to quantitate aggrecan mRNA by linearizing the plasmid at a *Xmnl* site prior to cRNA transcription with T3. (C) The cDNA map of chicken aggrecan mRNA based on the published sequence (Chandrasekaran and Tanzer, 1992; Li et al., 1993). The two repetitive domains that make up the CS binding domains (CS1: 10-amino-acid repeat and CS2: 20-amino-acid repeat) are highlighted within the large exon 10 that extends from codon 839 through 1893. The regions flanking the repeated domain that was sequenced (codons 1217 through 2086) is underlined, and the area where the inconsistencies were identified is underlined with a broken line. (D) The region within CS2 between the two *Sst* II sites (codons 1492 and 1710) contains the nanomelic mutation plus the six additional 20-amino-acid repeats not found in the original published sequence of this region (Chandrasekaran and Tanzer, 1992). The original repeat number prior to detection of the additional six repeats is illustrated above the shaded rectangle, and the corrected numbering is given below. The region was amplified with oligonucleotides shown by horizontal arrows that end at codons 1327 and 1735. The *Sst* II sites are located in repeats 7 and 12 of the published sequence and repeats 7 and 18 of the revised sequence. The nanomelic mutation is located in repeat 8 at codon 1513.

310 nt (arrow). As a control for the quantity of RNA used in the hybridization reaction, we employed a second probe from the C-terminal domain of type II collagen (Vuorio et al., 1982), since previous published work found that the quantity of type II collagen produced by nanomelic chondrocytes was normal (Sawyer and Goetinck, 1981). It generates a cRNA probe of 200 nt (lane 10) and protects two fragments of 150 and 100 nt (arrows). The signal strengths of the bands were significantly stronger in RNA obtained from normal cells (lanes 1 and 2 *vs.* lanes 7 and 8), probably because the nanomelic cells do not maintain their differentiated state as well as control cells. The film was re-exposed for shorter intervals to obtain comparable band densities between the control and nanomelic samples. The ratio of aggrecan to collagen bands in control cells is approximately 1:4 in both the nuclear and cytoplasmic compartments (lanes 3 and 4). In contrast, the 1:4 ratio that was present in the nuclear compartment of the nanomelic cells (lanes 5 and 7) was not maintained in the cytoplasmic compartment. At the shorter exposure, no aggrecan band was detectable in the cytoplasmic extract (lane 6). At the longer exposure (lane 8), a faint band migrating somewhat more slowly than the comparable nuclear band was observed. The band running just above the authentic aggrecan band is probably not related to aggrecan RNA because it is evident at equal intensity in both compartments of nanomelic cells (lanes 5–8) and at shorter exposures in normal cells (lane 3).

These data are consistent with diminished RNA accumulation secondary to premature translational termination such as an abnormality of splicing that results in a frame shifted transcript. To assess for possible abnormal splicing, nuclear RNA was obtained from unaffected embryos from a nanomelic mating assuming that ⅔ would be carriers and ⅓ would be normal. RNA from these presumptive heterozygotes and from wild type chicken chondrocyte

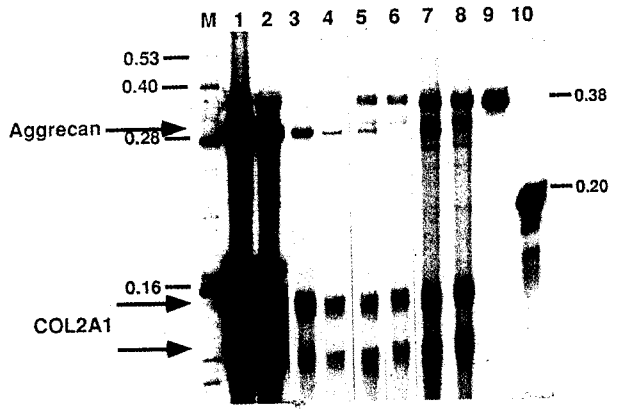

Fig. 2. Direct RNase protection of aggrecan RNA in the nuclear (N) and cytoplasmic (C) extracts from normal and nanomelic chondrocytes. Six μg of total RNA from each extract was protected with a *Xmnl*-truncated subfragment of probe H colinear with exons 13 and 14 of the aggrecan mRNA (see Fig. 1B). A type II collagen mRNA probe encoding the C-terminal propeptide was linearized with *Sau*3A. The RNA markers in Lane M are end-labeled; Lanes 1, 2, 7 and 8 were taken at the same X-ray film exposure, while Lanes 3, 4, 5 and 6 were taken at shorter exposures to illustrate the relative amounts of aggrecan and COL2A1 mRNA in the lanes. Lanes 1 and 2: N and C, normal; Lanes 3 and 4: N and C, normal; Lanes 5 and 6: N and C, nanomelic; Lanes 7 and 8: N and C, nanomelic. Lanes 9 and 10 are the input aggrecan and COL2A1 cRNA probes alone.

RNA was examined by direct and indirect RNase protection using probes which encoded exons 1 through 8 (probe B, Fig. 1), exons 6 through 10 (probe 7–2, Fig. 1) and exons 11 through 14 (probe H, Fig. 1). The results showed that all probes were fully protected by nuclear and cytoplasmic RNA from chondrocytes derived from wild type and normal embryos of a nanomelic mating (data not shown).

Having failed to find an abnormality of splicing, we searched the primary cDNA sequence for the presence of either a frame shift mutation or point mutation causing premature termination. RT/PCR sequencing from the 3′ end residues beginning at codon 2086 and from the 5′ end beginning at codon 1217 produced a sequence that was identical to the published sequence (Chandrasekaran and Tanzer, 1992) except for occasional polymorphic changes that would not account for the RNA phenotype of nanomelia. However, as the region containing the CS2 domain was entered and particulary when oligonucleotide primers from the CS2 repeat region were used to produce PCR fragments, numerous sequence inconsistencies, including stop codons, were encountered, although there was still great similarly to the published sequence. Sequence comparison with a cDNA obtained by Krueger et al. (1990) that was upstream of the CS2 domain was identical to ours and showed a similar loss of fidelity to the published sequence at its 3′ end.

Because we could not generate a reliable sequence from the repetitive CS2 domain with internal sequencing primers, we chose PCR oligonucleotides that flanked the domain (codon 1327 and codon 1735, Fig. 1C) with the intention of sequencing it with vector-derived primers. Based on the published sequence, the predicted size of this PCR fragment was 0.9 kb, but the observed fragment from genomic DNA of normal, nanomelic embryos and the plasmid originally used to sequence this domain was 1.25 kb (data not shown). Furthermore, the predicted 300 bp *Sst* II fragment (Chandrasekaran and Tanzer, 1992) which resides within the repetitive domain migrated at approximately 650 bp (codons 1492 and 1710) in all three samples of DNA (data not shown), thus identifying an apparent insertion within the *Sst* II fragment.

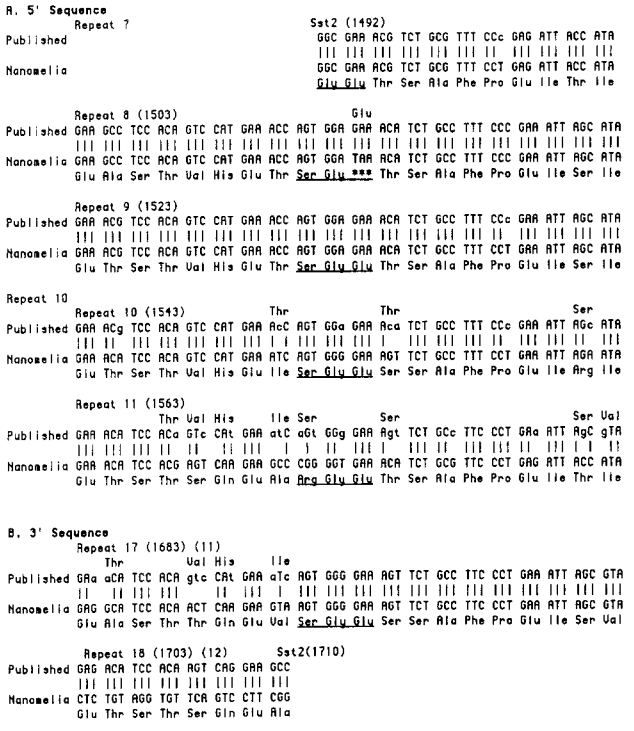

Fig. 3. Comparison of the DNA sequence of the ends of the *Sst* II fragment from nanomelia with the corresponding published sequence (Chandrasekaran and Tanzer, 1992). Genomic DNA from nanomelic embryos was PCR amplified using oligonucleotides that flank the *Sst* II sites as shown in Fig. 1. The analysis was performed on different nanomelic embryos and genomic DNA from normal chicken to exclude the possibility of a PCR artifact. (Panel A). The upstream *Sst* II cleavage site at codon 1492 is located within the seventh repeat. Sequence fidelity is reasonably preserved through repeat 9, begins to degenerate in repeat 10 and is completely lost in repeat 11. The nanomelic mutation is located within repeat 8. (Panel B) Sequence derived from the downstream *Sst* II cleavage site. The corrected repeat numbers, 17 and 18, are used while the original repeat numbers, 11 and 12, are given within the parentheses. Comparison of the sequences indicates that the 17th repeat of the corrected nanomelic sequence begins to correspond to the 11th repeat in the original sequence.

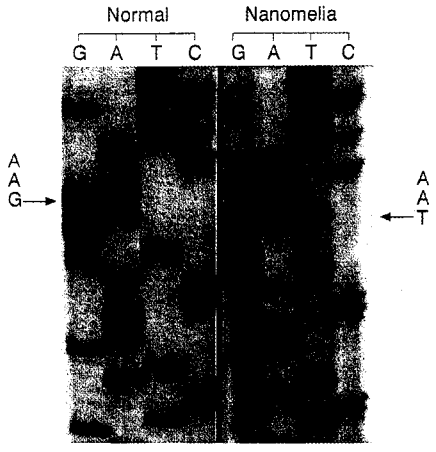

Fig. 4. DNA sequencing for the *Sst* II fragment within the CS2 domain that harbored the nanomelic mutation. The PCR derived fragment was subcloned into Bluescript KS^+ and sequenced from the vector primers.

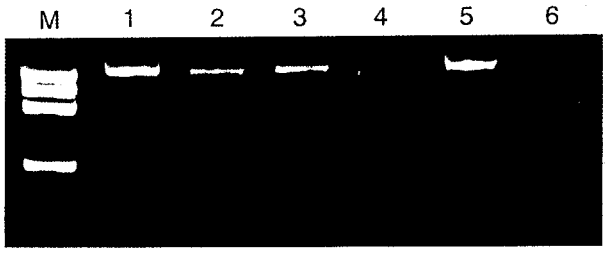

Fig. 5. *Bas*Bl restriction digest of the CS2 fragment obtained by PCR amplification of homozygous normal (lanes 1 and 2), heterozygous (lanes 5 and 6) and nanomelic embryos (lanes 3 and 4). Lanes 1, 3 and 5 are the uncut PCR product, while lanes 2, 4 and 6 were cut with *Bas*Bl. M, DNA markers.

The 657 bp *Sst* II fragment was cloned and its sequence compared with the published sequence (Chandrasekaran and Tanzer, 1992). Sequencing toward the 5′ end from codon 1710 to 1691 (Fig. 3, panel B) showed complete identity, but farther upstream of codon 1691 the fidelity with the published sequence was distinctly lost. Sequencing from codon 1492 at the 5′ end (Fig. 3, panel A) showed polymorphic changes at 1498 and 1538 and a G → T transversion, producing a Glu → stop codon, at 1513. In addition, conservative amino acid changes at 1550, 1534 and 1561 were found, followed by further degeneracy in fidelity to the published sequence. The results of the sequencing show that an additional six 20-amino-acid repeats are present which escaped analysis in the original sequence.

Once the basis of the difficulty in obtaining reliable sequence within the CS2 repeat domain was understood, the mutation causing the premature stop became apparent: the eighth repeat contained a GAA to TAA mutation (Figs. 4 and 1 D). No other alteration in sequence was uncovered within the nanomelic sequence from codon 1217 to 2086 that could account for the protein and RNA findings. Furthermore, the mutation generated a unique BasBl restriction site which could be demonstrated in genomic DNA from nanomelic embryos (Fig. 5, lane 4). Analysis of genomic DNA from phenotypically unaffected embryos collected from parents defined as nanomelic carriers showed, as expected, the presence of either one (Fig. 5, lane 6) or two normal alleles (Fig. 5, lane 2) that were resistant to cleavage by the enzyme.

Discussion

Our investigation into the molecular basis of nanomelia was influenced by our analysis of mutations of another extracellular matrix protein, type I collagen, in association with the human disease, osteogenesis imperfecta (Rowe, 1991; Byers and Steiner, 1992). In that heritable disorder, mutations which maintain the translational reading frame accumulate normal amounts of the mutant transcript in the cytoplasmic compartment, and corresponding levels of mutant protein are produced. In contrast, mutations which produce a premature stop of translation fail to accumulate the mutant transcript in the cytoplasm and lead to a 50% reduction of protein production (Barsh et al., 1982; Rowe et al., 1985; Willing et al., 1992; Stover et al., 1993). Thus, determination of the distribution of the mRNA between the nuclear and cytoplasmic compartments can provide an important starting point for mutation identification using total RNA if the mutation is contained within the cytoplasm, or using the nuclear RNA if the transcript is restricted to that compartment. This approach not only avoids the need for excessive DNA sequencing but provides a biological insight into the effects of the mutation on mRNA processing and translation.

The nanomelic chondrocytes demonstrated ample aggrecan mRNA in the nuclear compartment with little detectable accumulation in the cytoplasm. Structural analysis by RNase protection studies failed to show evidence of exon skipping or intron retention as a basis for a frameshifted transcript. We did not elect to use current methods for point mutation detection, such as chemical cleavage (Cotton, 1989) or single-stranded conformational polymorphisms (Sarkar et al., 1992) to localize the mutation as previous work by O'Donnell et al. (1988) using cell free translation indicated the likely domain harboring the stop codon. Not anticipated in the analysis of the suspect domain was the discrepancy uncovered between our data and the published sequence. As a result, the entire 3′ end of the cDNA was resequenced to ensure that the stop codon encountered at codon 1513 was the authentic mutation. This result was confirmed by the presence of the mutation-generated BasBl restriction site in homozygotes and heterozygotes for the

nanomelic phenotype. Our findings are in agreement with those of Li et al. (1993).

The omission of the six 20-amino-acid repeats of the CS2 domain in the published sequence only became apparent as sequencing through this region in the nanomelic mRNA revealed many inconsistencies. We experienced difficulty obtaining reliable sequence when primers derived from the CS2 domain were in the DNA sequencing reaction. Once the domain was subcloned and sequenced with primers external to the repeat region, the reason for the initial omission became clear. Since the revised sequence is present in both normal and nanomelic aggrecan RNA, it has no bearing on the nanomelic phenotype. The corrected sequence will be described elsewhere (Chandrasekaran et al., submitted) and is in agreement with the sequence of Li et al. (1993).

The molecular basis for the decreased cytoplasmic accumulation of aggrecan mRNA secondary to the presence of the premature termination codon is still not firmly defined. Multi-exon genes containing premature stops in other than the terminal exon fail to show accumulation of the mutant transcript in the cytoplasm either from a diminished half-life of the RNA within the cytoplasm (Baserga and Benz, 1988; Lim et al., 1992) or from failure of nuclear transport (Maquat and Kinniburgh, 1985; Cheng et al., 1990). Experiments to distinguish clearly between these two possibilities have been difficult to perform, and there may be alternative mechanisms based on the gene, cell type and mutation.

The aggrecan gene is unusual in that it has an internal exon of over 3000 bp. The current model of RNA splicing based on exon definition limits the size of an internal exon to less than 300 bp (Robberson, 1990). The fact that this gene has evolved a way to permit proper splicing of such a large internal exon may explain why more RNA containing the premature stop was detected by Northern analysis or translated into a truncated protein than has been observed for other genes with premature stop codons (Hamosh et al., 1991; Nakahashi et al., 1992; Koivisto et al., 1992). Although the importance of nanomelia is the insight it gives into the biology of aggrecan as related to cartilage development, the unusual size of exon 10 and its ability to defy the rules of exon definition may reveal how splicing occurs in a number of extracellular genes that have large internal exons (Hardingham and Fosang, 1992).

To date, a human homologue to nanomelia has not been defined, although several dominantly inherited dwarfing syndromes have been examined for DNA linkage to the aggrecan gene (Finkelstein et al., 1991). Recently, the cartilage matrix deficient mouse has been shown to contain a frame-shifted aggrecan transcript due to a 7-bp deletion in the 5′ coding region of the gene (Yamada, personal communication). Now that the phenotype and genotype of these two animal models of recessively inherited chondrodystrophy are appreciated, analysis for a homozygous null mutation in the corresponding human gene should focus on sporadic and probably lethal cases affecting the axial, facial and thoracic skeleton. In addition, molecular markers to identify the heterozygous state of nanomelia and the cartilage matrix deficient mouse can be used to test the hypothesis that chronic underproduction of aggrecan associated with one null aggrecan allele predisposes to degenerative joint disease. Since both animal models contain null mutations, it has not been possible to identify regions within the structural domains of the gene that delete function from the protein. At present, experiments of man or nature have not been identified to indicate that there are amino acid substitutions in the aggrecan gene that might act as dominant negative mutations as has been found in many structural proteins (Vikkula et al., 1993). However, there is an increasing realization that subtle mutations in the structural proteins expressed by chondrocytes may result in early onset degenerative joint disease (Nakata et al., 1993). This information, plus evidence that the heterozygote nanomelic or cartilage matrix deficient model have abnormal cartilage, would provide a rationale to examine human pedigrees with heritable osteoarthritis for genetic linkage to the aggrecan gene.

Acknowledgement

This work was supported by NIH Grant AR38933 and by the Medical Research Service, Veterans Affairs Medical Center.

References

Argraves, W. S., McKeown-Longo, P. J. and Goetinck, P. F.: Absence of proteoglycan core protein in the cartilage mutant nanomelia. *FEBS Lett 131:* 265–268, 1981.

Baldwin, C. T., Reginato, A. M. and Prockop, D. J.: A new epidermal growth factor-like domain in the human core protein for the large cartilage-specific proteoglycan. *J. Biol. Chem. 264:* 15747–15750, 1989.

Barry, F.: Genetic and structural organization of proteoglycans. *Biochem. Soc. Trans. 18:* 197–198, 1990.

Barsh, G. S., David, K. E. and Byers, P. H.: Type I osteogenesis imperfecta: A nonfunctional allele for proα1(I) chains of type I procollagen. *Proc. Natl. Acad. Sci. USA 79:* 3838–3842, 1982.

Baserga, S. J. and Benz Jr., E. J.: Nonsense mutations in the human β-globin gene affect mRNA metabolism. *Proc. Natl. Acad. Sci. USA 85:* 2056–2060, 1988.

Byers, P. H. and Steiner, R. D.: Osteogenesis imperfecta. *Ann. Rev. Med. 43:* 269–282, 1992.

Chandrasekaran, L. and Tanzer, M.: Molecular cloning of chicken aggrecan. *Biochem. J. 288:* 903–910, 1992.

Chandrasekaran, L. and Tanzer, M.: Molecular cloning of chicken aggrecan: Correction. *Biochem. J.* (submitted).

Cheng, J., Fogel-Petrovic, M. and Maquat, L. E.: Translation to near the distal end of the penultimate exon is required for normal levels of spliced triosephosphate isomerase mRNA. *Mol. Cell Biol. 10:* 5215–5225, 1990.

Chow, Q., Russell, M., Buck, D. E., Raymond, J. and Block, W.:

Prevention of pre-PCR mis-priming and primer dimerization improves low-copy number amplification. *Nucleic Acids Res. 20:* 1717–1723, 1992.

Cotton, R. G.: Detection of single base changes in nucleic acids. *Biochem. J. 263:* 1–10, 1989.

Doege, K., Saski, M. and Yamada, Y.: Rat and human cartilage proteoglycan (aggrecan) gene structure. *Biochem. Soc. Trans. 18:* 200–202, 1990.

Doege, K. J., Saski, M., Kimura, T. and Yamada, Y.: Complete coding sequence and deduced primary structure of the human cartilage large aggregating proteoglycan, aggrecan. *J. Biol. Chem. 266:* 894–902, 1991.

Dudhia, J., Fosang, A. J. and Hardingham, T. E.: Domain structure and sequence similarities in cartilage proteoglycan. *Biochem. Soc. Trans. 18:* 198–200, 1990.

Finkelstein, J. E., Doege, J., Yamada, Y., Pyeritz, R. E., Graham Jr., J. M., Moeschler, J. B., Pauli, R. M., Hecht, J. T. and Francomano, C. A.: Analysis of the chondroitin sulfate proteoglycan core protein (CSPGCP) gene in achondroplasia and pseudoachondroplasia. *Am. J. Hum. Genet. 48:* 97–102, 1991.

Genovese, C., Rowe, D. W. and Kream, B.: Construction of DNA sequences complementary to rat a_1 and a_2 collagen mRNA and their use in studying the regulation of type I collagen synthesis by 1,25 dihydroxy vitamin D. *Biochemistry 23:* 6210–6216, 1984.

Genovese, C., Brufsky, A., Shapiro, J. and Rowe, D.: Detection of mutations in human type I collagen mRNA in osteogenesis imperfecta by indirect RNase protection. *J. Biol. Chem. 264:* 9632–9636, 1989.

Grover, J. and Roughley, P. J.: Versican gene expression in human articular cartilage and comparison of mRNA splicing variation with aggrecan. *Biochem. J. 291:* 361–367, 1993.

Hamosh, A., Trapnell, B. C., Zeitlin, P. L., Montronse-Rafizadeh, C., Rosenstein, B. J., Crystal, R. C. and Cutting, G. R.: Severe deficiency of cystic fibrosis transmembrane conductance regulator messenger RNA carrying nonsense mutation R553X and W1316X in respiratory epithelial cells of patients with cystic fibrosis. *J. Clin. Invest. 88:* 1880–1885, 1991.

Hardingham, T. E. and Fosang, A. J.: Proteoglycans: Many forms and many functions. *FASEB J. 6:* 861–870, 1992.

Koivisto, L.-M., Turtola, H., Aaltoj-Setala, K., Top, B., Frants, R. R., Kovanen, P. T., Syvanen, A. C. and Kontula, K.: The familial hypercholesterolemia (FH)-North Karelia mutation of the low density lipoprotein receptor gene deletes seven nucleotides of exon 6 and is a common cause of FH in Finland. *J. Clin. Invest. 90:* 219–228, 1992.

Krueger Jr., R. C., Fields, T. A., Mensch Jr., J. R. and Schwartz, N. B.: Chick cartilage chondroitin sulfate proteoglycan core protein. *J. Biol. Chem. 265:* 12088–12097, 1990.

Landauer, W.: Nanomelia, a lethal mutation of the fowl. *J. Hered. 56:* 131–138, 1965.

Li, H., Schwartz, N. B. and Vertel, B. M.: cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and identification of a stop codon in the aggrecan gene associated with the chondrodystrophy, nanomelia. *J. Biol. Chem. 268:* 23504–23511, 1993.

Lichtler, A., Barrett, N. L. and Carmichael, G. G.: Simple, inexpensive preparation of T1/T2 ribonuclease suitable for use in RNase protection experiments. *Biotechniques 12:* 231–232, 1992.

Lim, S.-K., Sigmund, C. D., Gross, K. W. and Maquat, L. E.: Nonsense codons in human β-globin mRNA result in the production of mRNA degradation products. *Mol. Cell Biol. 12:* 1149–1161, 1992.

Maquat, L. E. and Kinniburgh, A. J.: A β° thalassemic β-globin RNA that is labile in bone marrow cells is relatively stable in HeLa cells. *Nucleic Acids Res. 13:* 2855–2867, 1985.

McKeown, P. J. and Goetinck, P. F.: A comparison of the proteoglycans synthesized in Meckel's and sternal cartilage from normal and nanomelic chick embryos. *Develop. Biol. 71:* 203–215, 1979.

Nakahasi, Y., Fujita, H., Taketani, S., Ishida, N., Kappas, A. and Sassa, S.: The molecular defect of ferrochelatase in a patient with erythropoietic protoporphyria. *Proc. Natl. Acad. Sci. USA 89:* 281–285, 1992.

Nakata, K., Ono, K., Miyazaki, J., Olsen, B. J., Muragaki, Y., Adachi, E., Yamamura, K. and Kimura, T.: Osteoarthritis associated with mild chondrodysplasia in transgenic mice expressing α1(IX) collagen chains with a central deletion. *Proc. Natl. Acad. Sci. USA 90:* 2870–2874, 1993.

Neame, P. J.: Domains in cartilage proteoglycans: Do they define function? *Biochem. Soc. Trans. 18:* 202–204, 1990.

O'Donnell, C. M., Kaczman-Daniel, K., Goetinck, P. F. and Vertel, B. M.: Nanomelic chrondrocytes synthesize a glycoprotein related to chondroitin sulfate proteoglycan core protein. *J. Biol. Chem. 263:* 17749–17754, 1988.

Robberson, B. L., Cote, G. J. and Berget, S. M.: Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell Biol. 10:* 84–94, 1990.

Rowe, D. W., Shapiro, J. R., Poirier, M. and Schlesinger, S.: Diminished type I collagen synthesis and reduced α1(I) collagen mRNA in cultured fibroblasts from patients with dominantly inherited (type I) osteogenesis imperfecta. *J. Clin. Invest. 71:* 689–697, 1985.

Rowe, D. W.: Osteogenesis Imperfecta. In: *Bone and Mineral Research*, ed. by Heersche, J. N. M. and Kanis, J. A., Elsevier, Amsterdam, 1991, pp. 209–241.

Sai, S., Tanaka, T., Kosher, R. A. and Tanzer, M. L.: Cloning and sequence analysis of a partial cDNA for chicken cartilage proteoglycan core protein. *Proc. Natl. Acad. Sci. USA 83:* 5081–5085, 1986.

Sarkar, G., Yoon, H.-S. and Sommer, S. S.: Screening for mutations by RNA single-strand conformation polymorphism (rSSCP): Comparison with DNA-SSCP. *Nucleic Acids Res. 20:* 871–878, 1992.

Sawyer, L. M. and Goetinck, P. F.: Chondrogenesis in the mutant nanomelia. Changes in the fine structure and proteoglycan synthesis in high density limb bud cell cultures. *J. Exp. Zool. 216:* 121–131, 1981.

Stirpe, N. S., Argraves, W. S. and Goetinck, P. F.: Chondrocytes from the cartilage proteoglycan-deficient mutant, nanomelia, synthesize greatly reduced levels of the proteoglycan core protein transcript. *Develop. Biol. 124:* 77–81, 1987.

Stover, M. L., Primorac, D., McKinstry, M., Lu, S. C. and Rowe, D. W.: Defective splicing of α1(I) collagen mRNA in nondeforming (Type I) osteogenesis imperfecta. *J. Clin. Invest. 92:* 1994–2022, 1993.

Tanaka, T., Har-El, R. and Tanzer, M. L.: Partial structure of the gene for chicken cartilage proteoglycan core protein. *J. Biol. Chem. 263:* 15831–15835, 1988.

Urlaub, G., Mitchell, P. J., Ciudad, C. J. and Chasin, L. A.: Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. *Mol. Cell Biol. 9:* 2868–2880, 1989.

Velleman, S. G. and Clark, S. H.: The cartilage proteoglycan deficient mutation, nanomelia, contains a DNA polymorphism in

the proteoglycan core protein gene that is genetically linked to the nanomelia phenotype. *Matrix 11:* 66–72, 1992.

Vieira, J. and Messing, J.: Production of single-stranded plasmid DNA. *Meth. Enzymol. 153:* 3–11, 1987.

Vikkula, M., Ritvaniemi, P., Vuorio, A. F., Kaitila, I., Ala-Kokko, L. and Peltonen, L.: A mutation in the amino-terminal end of the triple helix of type II collagen causing severe osteochondrodysplasia. *Genomics 16:* 282–285, 1993.

Vuorio, E., Sandell, L., Kravis, D., Sheffield, V. C., Vuorio, T., Dorfman, A. and Upholt, W.: Construction and partial characterization of two recombinant cDNA clones for procollagen from chicken cartilage. *Nucleic Acids Res. 10:* 1175–1192, 1982.

Willing, M. C., Pruchno, C. J., Atkinson, M. and Byers, P. H.: Osteogenesis imperfecta type I is commonly due to a COL1A1 null allele of type I collagen. *Am. J. Hum. Genet. 51:* 508–515, 1992.

Dr. D. W. Rowe, Department of Pediatrics, Room L5093; University of Connecticut Health Center, Farmington, Connecticut, USA.

Received November 12, 1993; accepted January 24, 1994.