Isolation of the Osteonectin Gene: Evidence That a Variable Region of the Osteonectin Molecule Is Encoded within One Exon

David M. Findlay,* Larry W. Fisher, Camille I. McQuillan, John D. Termine, and Marian F. Young

Bone Research Branch, National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland 20892

Received July 6, 1987; Revised Manuscript Received September 28, 1987

ABSTRACT: A complementary DNA clone for bovine osteonectin was used to isolate the osteonectin gene from two libraries of bovine genomic DNA fragments. Two overlapping clones were obtained whose relationship was determined by restriction mapping and sequence analysis. The two clones contain the entire osteonectin coding region spanning approximately 11 kilobases of genomic DNA. The coding region of the gene was determined, by electron microscopy and DNA sequencing, to reside in nine exons. In addition, there is at least one 5′ exon interrupted by an intron in the 5′-nontranslated sequence of the gene. Excluding this 5′ exon and the 3′-terminal exon, the exons are small and approximately uniform in size, averaging 130 ± 17 base pairs. Three of the exons at the 5′ end of the gene were sequenced and appear to encode discrete protein domains. For example, the putative exon 2 contains the coding region for the leader peptide of the molecule. The amino-terminal protein sequence was determined for osteonectin extracted from human, rabbit, and chicken bone and compared with those for bovine, mouse, and pig osteonectin. These data suggest that osteonectin is highly conserved between species, interspecies changes being seen primarily at the amino terminus of the protein and specifically in the region encoded by putative exon 3 in the bovine gene.

Osteonectin is a 32-kilodalton (kDa)[1] glycoprotein originally isolated from bone where it comprises a significant percentage of the noncollagenous protein (Termine et al., 1981a). Osteonectin is biosynthesized by osteoblast-like cells in culture (Gehron Robey & Termine, 1985), and it is incorporated into the cell layer/extracellular matrix (Gehron Robey & Fisher, 1987). In vitro studies have shown that osteonectin has a high affinity for calcium, hydroxyapatite, and type I collagen (Romberg et al., 1985) and is capable of facilitating mineralization of type I collagen in vitro (Termine et al., 1981b). In a different in vitro system, osteonectin can suppress the rate of hydroxyapatite crystal growth in solution (Romberg et al., 1986). It was therefore inferred that osteonectin was involved in mineralization of bone extracellular matrix. Recently, however, proteins from nonmineralizing tissues have been sequenced and shown to be highly homologous or identical with osteonectin from bone. These include a platelet osteonectin analogue (Stenner et al., 1986), a 43-kDa protein secreted by bovine aortic endothelial cells in culture (Sage et al., 1986), a basement membrane protein (BM40) first identified in the mouse EHS tumor (Dziadek et al., 1986), and a protein termed SPARC produced by mouse parietal endoderm (Mason et al., 1986a,b). In addition, M. F. Young et al. (personal communication) have used specific osteonectin antisera and osteonectin cDNA to identify osteonectin in the rapid proliferative phase of human and mouse differentiating uterine decidua and in peritumoral stroma cells in some tumors.

The presence of osteonectin in nonmineralizing tissues has caused a reexamination of its physiological role. It is possible that the protein functions differently in mineralized and nonmineralized tissues, where the complement of neighboring matrix proteins clearly differs. Alternatively, its mineral binding property may be secondary to some function common to all tissues where osteonectin is found. Studies of the regulation of osteonectin expression may ultimately produce clues to the biological function of the molecule. Evidence to date indicates that expression of the osteonectin gene is under tight regulation; in vivo, it appears that the molecule is present in areas of rapid tissue remodeling. That osteonectin may be under hormonal control is suggested by preliminary work using normal human bone cells (Gehron Robey et al., 1986) where, at least in culture, 1,25-dihydroxyvitamin D_3 is capable of down-regulating osteonectin expression. Ultimately, understanding the expression of osteonectin and its regulation at the genomic level by systemic factors depends upon the isolation of the genomic elements encoding the protein and exerting transcriptional control upon it. To this end, we previously made cDNA encoding the entire bovine osteonectin protein (Young et al., 1986) which, in the present work, enabled the first isolation of the bovine osteonectin gene. We report here the exon–intron structure and partial sequence of DNA clones containing the gene. We present new protein sequence data which provide evidence that in an otherwise highly conserved protein, interspecies changes appear clustered at the N-terminus of the osteonectin molecule and show that the majority of these changes potentially reside in a single exon as defined in the bovine gene.

Materials and Methods

Construction and Screening of Bovine Genomic Libraries. High molecular weight genomic DNA was extracted from 5-month fetal bovine liver as described (Maniatis et al., 1982), and 100 μg of DNA was cut into 9–50-kb fragments by digestion with 8 units of *Sau*3AI for 10 min at 37 °C. DNA was centrifuged through a 10–40% sucrose gradient containing 1 M NaCl, 20 mM Tris (pH 8.0), and 5 mM EDTA, and 10 μL was removed from 0.5-mL fractions for assessment by electrophoresis in 0.8% agarose gels (Maniatis et al., 1982). DNA fragments with lengths of 14–28 kb were pooled, dialyzed against 10 mM Tris (pH 7.5)/1 mM EDTA, concen-

[1] Abbreviations: kb, kilobase(s); bp, base pair(s); kDa, kilodalton(s); Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; PFU, placque-forming unit(s); Tricine, *N*-[tris(hydroxymethyl)methyl]glycine; HPLC, high-performance liquid chromatography; LDL, low-density lipoprotein.

trated with 1-butanol, and finally precipitated in 0.2 M sodium acetate and 2 volumes of ethanol. Approximately 2 μg of genomic DNA was ligated to 3 μg of *Eco*RI- and *Bam*HI-digested EMBL 3 vector (Promega) in the presence of 25 mM Tris (pH 7.4), 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.2 mM spermidine trihydrochloride, 0.5 mM ATP, and 200 units of T4 DNA ligase (New England Biolabs). Ligated DNA was packaged into phage particles using Giga-Pack Plus (Stratagene). After infection of LE392 cells, 1×10^6 PFU were obtained per microgram of DNA. Approximately 0.5×10^6 plaques were screened after being plated at a density of 20000 plaques/plate on 150-mm Luria broth plates. Osteonectin genomic clones were identified in the library by plaque hybridization using ^{32}P-labeled fragments (see text) of a full-length bovine osteonectin cDNA clone, λOn17 (Young et al., 1986), and a previously described hybridization technique (Maniatis et al., 1982). Similarly, a λCharon 30 bovine genomic library (a gift from D. Capon, Genentech) was screened after 0.75×10^6 plaques were plated.

Synthetic Oligonucleotides. The sequences of the oligonucleotides mentioned herein are as follows: 376, 3′CAGGTCCACCTTCATCCT5′; MB, 3′AACGGACTACTCGTTCTTCA5′; MY2, 3′GCACCGGCTCCACCGGCTC5′; MY9, 3′AAGAAAGAGGAAACGGAC5′; MY10, 3′TCCCGGACCTAGAAGAAA5′; 375, 3′CGGACGGACGGACGGTGACT5′.

Electron Microscopy. Heteroduplexes were prepared between recombinant genomic DNA in λCharon 30 arms (λOg6) and recombinant cDNA in λgt11 arms. DNA mixtures were denatured for 10 min at 85 °C in 60% formamide, 0.5 M NaCl, and 0.1 M Tricine (pH 8.0) and then incubated at 52 °C for 16 h. For R-loop analysis, heteroduplexes were prepared between λOg6 and wild-type λCharon 4A Phage DNA. After denaturation, the DNA mixtures were chilled on ice and then hybridized with bovine bone cell poly(A+) RNA at 52 °C for 16 h. Electron microscope analyses were carried out as described (Tilgham et al., 1978) with double-strand circular SV40 DNA added as an internal size standard.

Direct Agarose Gel Hybridization. Recombinant genomic restriction fragments were hybridized with ^{32}P-labeled oligonucleotides directly in 0.8% agarose gels, as described elsewhere (Miyada et al., 1985). Hybridization temperatures (T_h) were calculated according to the relationship

$$T_h = 4l_1 + 2l_2 - 5\ °C$$

for duplexes 11–20 bases long where l_1 is the number of G + C bases and l_2 is the number of A + T bases. Oligonucleotides were end-labeled using T4 polynucleotide kinase (Pharmacia), and the gels were treated to enable reprobing with oligonucleotides as described previously (Miyada et al., 1985).

Northern Analysis. A 3.5-μg aliquot of total RNA extracted from cultured bovine bone cells (Gehron Robey & Termine, 1985) or intact bovine liver was electrophoresed in a 1.2% formaldehyde–agarose gel. After transfer to nitrocellulose (Maniatis et al., 1982), a *Hin*dIII–*Kpn*I fragment of λOg6 (Figure 3) or intact recombinant λOg6 DNA was ^{32}P labeled by nick translation (Amersham Kit) and used to probe replicate blots. Hybridization conditions and filter washing, using conditions which were as described previously (Smith et al., 1975; Young et al., 1986), preceded autoradiography of the nitrocellulose filters.

DNA Sequence Analysis. *Hin*dIII–*Kpn*I (λOg6) or *Eco*RI (λOg2) restriction fragments of genomic DNA (Figure 3) were subcloned into M13 mp-19 vector and sequenced by the chain-termination method (Sanger et al., 1977). Primers were either 17-mer primers directed to the vector or synthetic oligonucleotides constructed against portions of the cDNA sequence, as described above. Both the Klenow fragment (DNA polymerase I) and reverse transcriptase enzymes were employed for sequencing although the latter yielded less ambiguous sequence data in regions of high G-C content.

RNA Sequencing. RNA was sequenced by a modification of the method of Geliebter et al. (1986). Five micrograms of poly(A+)RNA (in 5 μL) was heated at 65 °C for 5 min and quickly cooled on ice slush to denature the RNA secondary structure. After addition of 2.5 μL of 1 M KCl together with 10 ng (2.5 μL) of ^{32}P end-labeled oligonucleotide primer (MY10), annealing was accomplished by incubation at 50 °C for 1 h. The incubate was removed to room temperature where 5 μL of transcription buffer, of the following composition, was added: 125 mM Tris (pH 8.3 at 50 °C), 25 mM $MgCl_2$, 160 μg/mL actinomycin D, 0.1% (v/v) mercaptoethanol, and 1 unit/mL RNasin (Promega Biotec). Three microliters of this mixture was aliquoted to tubes containing, in a volume of 0.75 μL, all four deoxynucleotides (1.25 mM dATP, dCTP, and dTTP; 2.5 mM dGTP) and one of ddATP (0.47 mM), ddCTP (0.24 mM), ddGTP (0.35 mM), or ddTTP (0.7 mM). Sequencing reactions were initiated by the addition of 0.5 μL of AMV reverse transcriptase (Seikagaku) and carried out at 50 °C for 1 h.

Osteonectin Extraction and Amino Acid Sequencing. Bone from rabbit, human, and chick was sequentially extracted in guanidine hydrochloride plus 0.5 M EDTA essentially as detailed elsewhere (Termine et al., 1981b; Fisher et al., 1987). Purified lyophilized samples were dissolved in 1% trifluoroacetic acid and subjected to automated Edman degradation using an Applied Biosystems Model 470A gas-phase sequencer employing the standard "No Vac" program supplied by the manufacturer (Hewick et al., 1981). Phenylthiohydantoin derivatives were identified by HPLC on an IBM cyano column (Hunkapiller & Hood, 1983). The HPLC system used with this column consisted of a Perkin Elmer Series 4 liquid chromatograph, an LC-85B spectrophotometric detector equipped with a 1.4-μL flow cell, an LCI 100 computing integrator, and a Model 7500 computer employing Chrom III software. These analyses were supplied under contract by the University of California at San Diego.

Results

Isolation of Osteonectin Genomic Clones. Previously we have constructed λ phage clones containing full-length DNA inserts coding for bovine osteonectin (Young et al., 1986). Using cDNA fragments of one such clone (λOn17), we screened bovine genomic libraries by plaque hybridization. First, an available library (Genentec) was screened using as probe a ^{32}P-labeled 1.5-kb *Eco*RI osteonectin cDNA fragment, which comprises the DNA code for the C-terminus of the protein and 1000 bp of noncoding sequence (Figure 1). Four genomic clones were obtained after repeated rounds of screening. One of these, λOg6, was chosen for more complete characterization of the basis of preliminary mapping. This clone was found to lack approximately 200 bp of sequence found in the cDNA (out of ~2.0 kb of total cDNA size). To obtain genetic clones containing the 5′ end of the gene, the same library was rescreened with the most 5′ portion of the cDNA, a 212 bp *Eco*RI fragment which encodes the N-terminus of the molecule, and 54 bp of noncoding sequence. Because no positive plaques were found, a phage library of bovine genomic DNA was constructed as described under Materials and Methods. Of the four clones isolated by using the same 212 bp probe, one clone, λOg2, was found to contain the remainder of the coding sequence for osteonectin.

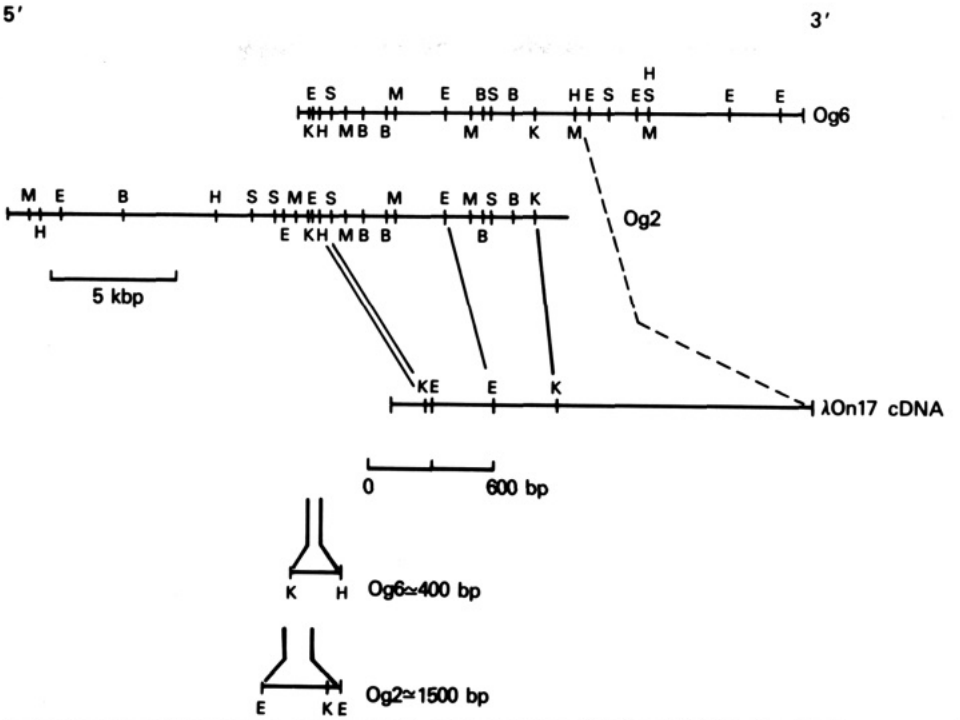

FIGURE 1: Restriction endonuclease cleavage map of osteonectin genomic clones λOg6 and λOg2. Restriction enzyme cleavage sites in the insert DNA are indicated as follows: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; M, *Sma*I; S, *Sst*I. The orientation of the genomic DNA is indicated, as is the relationship of the genomic DNA to a bovine osteonectin cDNA (λOn17). Two overlapping fragments from λOg6 and λOg2 which were subcloned for sequencing are indicated in expanded scale directly below their respective positions in each clone.

Identification of Osteonectin Genomic Clones by Northern Analysis. In order to show that λOg6 corresponds to the gene coding for an osteonectin messenger RNA bearing the expected size and tissue specificity, a Nothern blot hybridization was performed. Both the intact genomic clone and a *Hin*dIII–*Kpn*I restriction fragment of approximately 400 bp (see Figure 1) were ^{32}P labeled by nick translation and hybridized to RNA from bone or liver. In both cases (Figure 2), a band of hybridization was seen in the autoradiograph at ~2 kb, the size of bovine bone mRNA; comparatively very little hybridization, as observed in earlier studies, was seen with liver RNA (Young et al., 1986).

Mapping of the Genomic Clones. To determine the relationship between the genomic clones λOg6 and λOg2 and bovine osteonectin cDNA, the clones were mapped by using specific restriction endonuclease digestions. DNA from single and double endonuclease digests was subjected to agarose gel electrophoresis, transferred to nitrocellulose, and hybridized to specific ^{32}P-labeled fragments of osteonectin cDNA. Maps were confirmed by using the "Southern Cross" method (NEN). Figure 1 shows the restriction enzyme cleavage maps for the two genomic clones in relation to λOn17 cDNA. It can be seen that approximately 90% of the coding sequence for osteonectin is contained in 10 kb of λOg6, while the remainder of the protein sequence is coded for by clone λOg2. Clone λOg2 extends approximately 10 kb 5′ beyond the coding sequence of the osteonectin gene. To determine more precisely which 5′ DNA sequences were present in exon-containing fragments of the genomic clones, fragments of restriction enzyme digested DNA were separated by electrophoresis on agarose gels. The fragments were then probed within the gel using ^{32}P end-labeled synthetic oligonucleotides directed against the cDNA sequence. As shown in Figure 3, oligonucleotide 376 hybridized to fragments of both genomic clones whereas oligonucleotide MB hybridized exclusively to clone λOg2. Oligonucleotide MY10 hybridized to clone λOg2 only and is shown hybridized to a *Hin*dIII fragment from this clone.

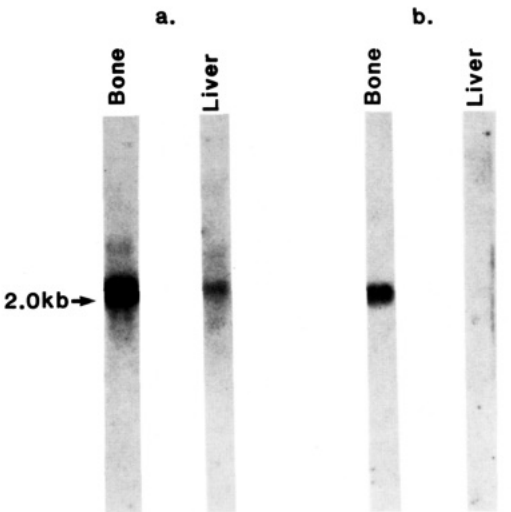

FIGURE 2: Northern hybridization. Total RNA was extracted from cultured bovine bone cells or intact bovine liver, and 3.5-μg aliquots were electrophoresed in 1.2% agarose gels and transferred to nitrocellulose. Replicate filters were hybridized to intact recombinant λOg6 DNA (a) or a *Hin*dIII–*Kpn*I fragment of λOg6 (b) labeled with ^{32}P. Approximately 10^7 cpm were used for hybridization. After being washed, filters were autoradiographed with 3-day exposure time.

These data confirmed that λOg6 and λOg2 share sequence found in the 212 bp cDNA *Eco*RI fragment but that only λOg2 encompasses the more 5′ sequence from this fragment.

Exon–Intron Structure of the Osteonectin Gene. To determine the exon–intron structure of λOg6, electron microscope vizualization of R-loop hybrids between bone cell poly(A+) RNA and the genomic λOg6 clone or heteroduplexes of λOn17 osteonectin cDNA and λOg6 osteonectin genomic DNA was performed. Both types of hybridizations were identical in appearance. Figure 4 shows a photomicrograph of a heteroduplex together with an interpretive diagram. The diagram shows approximately 10 kb of flanking DNA at the 3′ end of

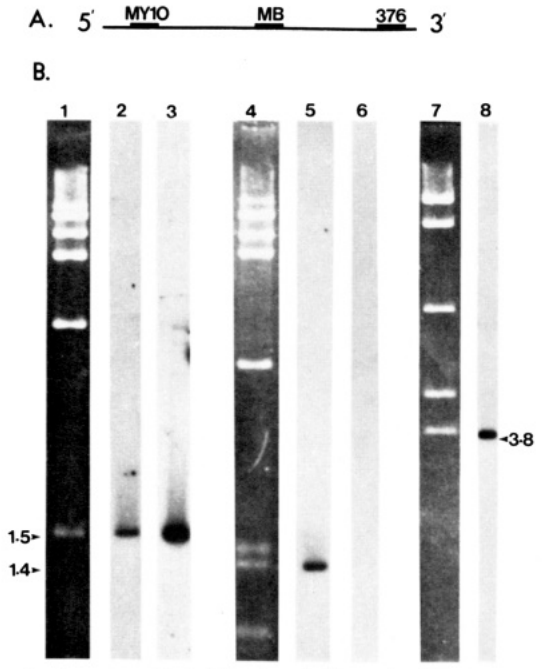

FIGURE 3: Agarose gel hybridization of oligonucleotides to osteonectin genomic clone fragments. Oligonucleotides complementary to portions of the 212 bp *Eco*RI fragment of the cDNA, as shown in (A), were hybridized to restriction enzyme fragments of λOg2 and λOg6 run in 0.8% agarose gels (B). Lanes 1, 4, and 7, ethidium bromide stained DNA fragments; lanes 2, 3, 5, 6, and 8, autoradiographs of gels after hybridization with ^{32}P-labeled oligonucleotides. Lane 1, λOg2 *Eco*RI fragments and following hybridization to ON 376 (lane 2) or ON MB (lane 3). Lane 4, λOg6 *Kpn*I/*Bam*HI fragments and following hybridization to ON 376 (lane 5) or ON MB (lane 6). Lane 7, λOg2 *Hin*dIII fragments and after hybridization to ON MY10 (lane 8). Size markers indicate lengths of fragments in kilobases.

Table I: Intron–Exon Structure of the Osteonectin Gene

	exon	intron
2	70[a]	~1200
3	63[a]	
4	87,[a] 131 ± 86	1571 ± 199
5	157 ± 44	1820 ± 142
6	177 ± 37	1823 ± 177
7	162 ± 39	1150 ± 63
8	173 ± 45	2142 ± 168
9	150 ± 17	508 ± 126
10	1188 ± 64	

[a] Sized by sequencing. All other sizes determined by electron microscopy; size in bp ± SD. Intron and exon numbers assigned assuming only one 5′-noncoding exon and intron.

the gene and a loop-out for the β-galactosidase gene which is located in the short arm of λgt11. Single-stranded loop-outs corresponding to introns, interrupting regions of double-stranded DNA (exons), are labeled 4–9, numbering from the 5′ end of the gene. The 5′ to 3′ polarity of the gene was established by hybridization of defined regions of cDNA to restriction enzyme-digested genomic DNA. Introns and exons were sized by analysis of several micrographs (Table I), with reference to double-stranded SV40 circles included as length standards. Exons 2, 3, and 4 were sized by sequence analysis. Intron and exon numbers are assigned assuming that there is only one intron and exon in the 5′-noncoding region (see below).

Sequence Analysis. Because of certain interesting features of the N-terminus of the osteonectin molecule, as discussed below, the three most 5′ exons were delineated by sequencing fragments from both clones. A *Hin*dIII–*Kpn*I fragment of λOg6 and an *Eco*RI fragment of λOg2 (see Figure 1), both

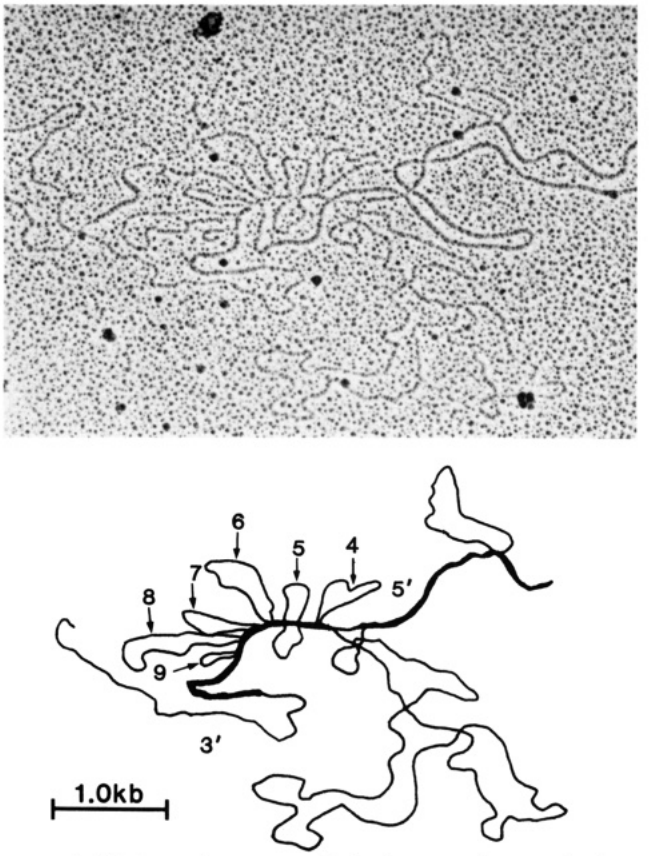

FIGURE 4: Electron microscope analysis of osteonectin genomic clone λOg6. Electron micrograph (upper panel) and interpretive diagram (lower panel) of a heteroduplex prepared between recombinant genomic DNA in λCharon 30 arms and recombinant osteonectin cDNA in λgt11 arms (λOn17). Introns are numbered 4–9 with respect to the entire gene, and the 3′ → 5′ polarity of the genomic DNA is noted.

of which hybridized with the 5′ end of the cDNA, were subcloned into M13 vector. Sequencing was by primer extension using both M13 sequencing primers and synthetic oligonucleotides (see Materials and Methods). The sequence (Figure 5) confirmed that the clones encode the bovine osteonectin gene. The λOg6 fragment was sequenced in the 5′ to 3′ direction from the indicated *Kpn*I site to an exon–intron boundary dividing the Asn-53 codon, beginning with the 3′-splice junction consensus dinucleotide GT (Wieringa et al., 1983). Much of this sequence was encompassed in the *Eco*RI fragment of λOg2, and sequencing 3′ to 5′ by priming with oligonucleotide 376 indicated that an exon–intron boundary occurred between the codons for Glu-23 and Val-24. The consensus sequence GA(C) marked the 5′-intronic boundary (Wieringa et al., 1983) between what were assigned exon 4 and intron 3. The use of oligonucleotides MY2 and MB established that exon 3 spanned the coding sequence of Gln-3 to Glu-23 inclusive. Again the 5′-intronic dinucleotide GA was found at the presumed intron–exon border. An imperfect amino acid repeat within the osteonectin protein was found encoded at the 3′ end of exons 3 and 4. Exon 2 was sequenced through the coding sequence for the putative "pre" sequence of osteonectin and 13 bases into noncoding sequence where an exon–intron boundary was found shortly before the 5′ end of the *Eco*Ri genomic fragment. To determine whether this apparent divergence was the result of incorrect sequence analysis at the extreme 5′ end of the cDNA, the fidelity of the cDNA sequence was confirmed by direct sequencing of poly(A+) mRNA from bovine bone cells. While cDNA and mRNA sequence analysis yielded identical data, divergence

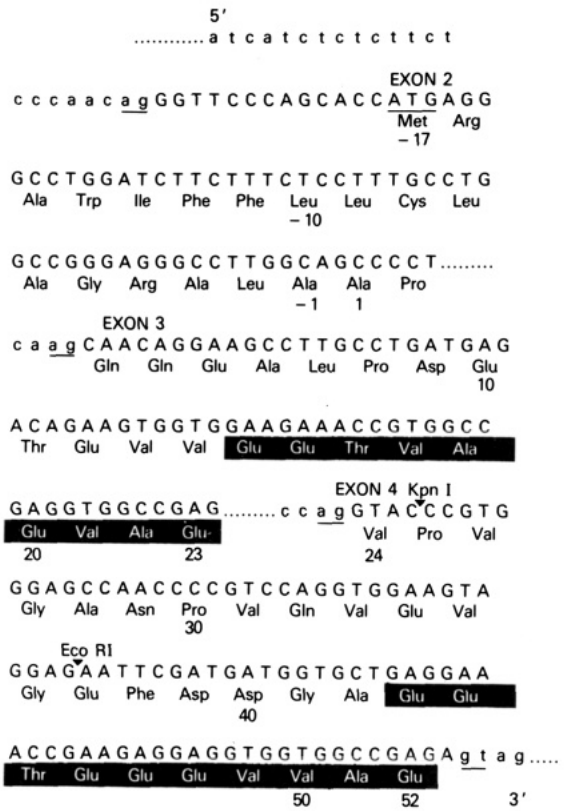

FIGURE 5: Nucleotide sequence of osteonectin genomic DNA. Partial sequences of a *Hin*dIII–*Kpn*I restriction fragment of λOg6 and an *Eco*RI restriction fragment of λOg2 (see Figure 3) were derived by primer extension. The λOg6 fragment was sequenced 5′ → 3′ from the *Kpn*I site indicated which is the 5′ extent of the *Hin*dIII–*Kpn*I fragment. The λOg2 fragment was sequenced 3′ → 5′ from the *Eco*RI site indicated which is the 3′ extent of the *Eco*RI fragment. Underlined are the splice consensus dinucleotides and the ATG codon indicating the start of translation, respectively. Lower case letters are used for intervening sequence and upper case for exons. Two imperfect amino acid repeats are shown boxed.

of genomic and cDNA sequences was seen after a GA dinucleotide (Figure 6).

Analysis of Osteonectin Protein Sequence. The full cDNA sequence and predicted amino acid sequence had previously been determined for bovine osteonectin (J. D. Termine et al., personal communication) and "SPARC", the presumed mouse osteonectin (Mason et al., 1986a). Comparison of these sequences showed that the two molecules were most different at their N-termini. It was therefore of interest to also compare the N-terminal sequences of osteonectin from a number of other species to look for correlation between the protein sequence and the genomic exon structure. Osteonectin was extracted from the bones of a variety of animal species, and N-terminal amino acid sequences were obtained. In the case of human, rabbit, and chick, purified bone osteonectin protein was extracted from bone as described. The pig sequence was reported by Domenicucci et al. (1987). These particular N-terminal amino acid sequences were chosen because they spanned the region encoded by exon 3. They are shown in Figure 7 relative to the amino acid sequence determined from bovine amino acid sequencing and cDNA sequencing (Young et al., 1986), with the coding domain of exon 3 in the bovine gene indicated.

Discussion

Two osteonectin genomic clones, λOg6 and λOg2, were isolated from bovine genomic libraries. The clones span approximately 10 kb 3′ downstream from the 3′ end of the cDNA to 10 kb 5′ upstream from the Met codon proposed to be the

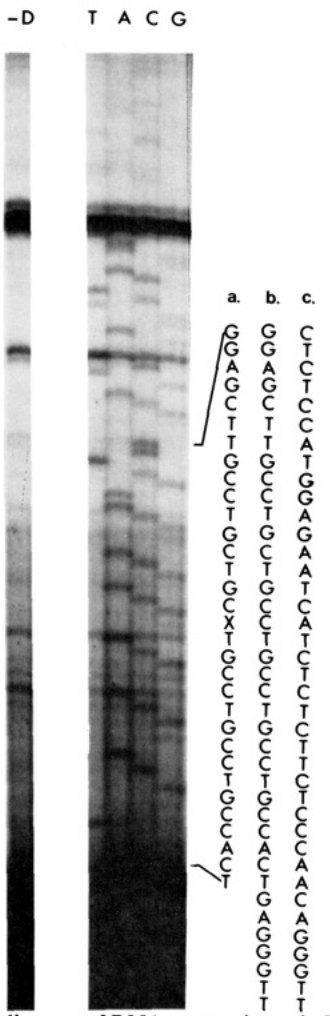

FIGURE 6: Autoradiogram of RNA sequencing gel. Polyadenylated bone cell RNA was sequenced as described under Materials and Methods. Sequencing was initiated with oligonucleotide primer MY10, complementary to the mRNA at codons for amino acids −16 to −11. Sequences shown correspond to (a) the complement of sequence read from the gel which matches (b), the corresponding sequence of the cDNA. Sequences are in the 3′ → 5′ orientation reading from bottom to top. (c) is the genomic DNA sequence determined with MY10 as primer with the *Eco*RI fragment of clone λOg2 described in Figure 3. Lane −D shows primer extension of the RNA in the absence of dideoxynucleotides.

start of translation. Both clones have been mapped. This and sequence analysis have specified precisely the extent of overlap between them. The osteonectin genomic clones were also characterized with respect to exon–intron structure. Clone λOg6 was shown by electron microscopy of heteroduplexes and R-loop analysis to contain seven exons and six introns. Apart from the most 3′ exon, which is likely to correspond to the 3′-noncoding sequence of the mRNA, as seen, for example, in the pyruvate kinase gene (Lonberg & Gilbert, 1985), exons are relatively uniform in size, being ~130 bp long. Since electron microscopic analyses are accurate to only about 50 bp, it remains possible that very small intervening sequences are also present in λOg6. Exons 2, 3, and 4 were directly sequenced by subcloning fragments from λOg6 and λOg2. There was no demonstrable cDNA hybridization with λOg2 fragments 5′ from exon 2. Similarly, oligonucleotide 375 directed against the 5′-encoding sequence of the cDNA did not hybridize with clone λOg2. Intron 1 is therefore predicted to be large, in excess of 9 kb, based on the extent of 5′ DNA flanking λOg2 from the end of exon 2 (see Figure 1). The fidelity of the cDNA 5′ sequence has been verified by direct

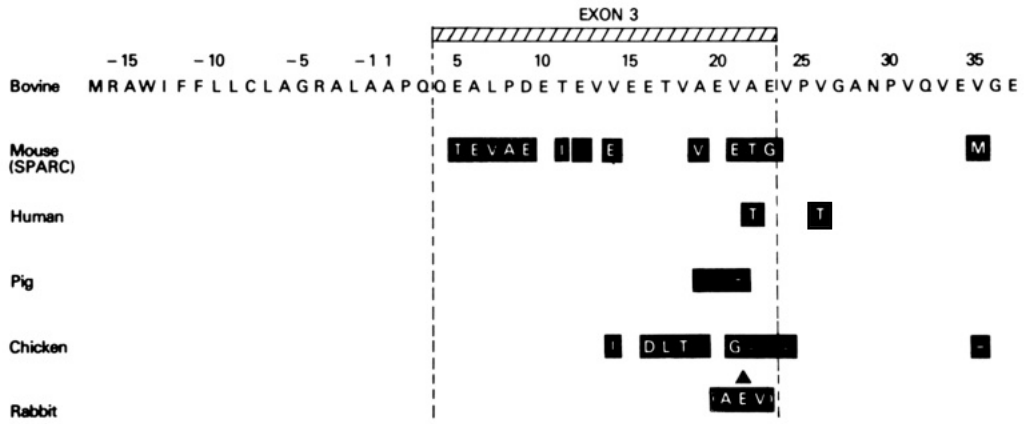

FIGURE 7: Amino acid sequence of osteonectin. Bovine and mouse sequences are as predicted from the cDNA sequence and include 17-residue leader sequences. Reverse type indicates residues different from bovine; the arrowhead above the rabbit sequence indicates an insertion of an AEV tripeptide. The region encoded by exon 3 of the bovine gene is indicated by the hatched line.

sequencing of poly(A+) mRNA purified from cultured bovine bone cells. Other examples of intervening sequences in the 5′-encoding sequence of genes have been found (Eckert & Green, 1986; Yang et al., 1986; Peralta et al., 1987). In the case of the c-myc gene, an intron in the 5′-noncoding sequence separates exon 1, which contains regulatory activity, from the coding portion of the gene (Yang et al., 1986). To determine whether the 5′-noncoding region of the osteonectin gene also contains regulatory activity, we are currently in the process of isolating both intron 1 and DNA encoding the start of osteonectin transcription.

Other features of the structural organization of the bovine osteonectin gene can be seen from Figure 5. For example, the coding region of the "pre" sequence of the molecule is entirely contained within exon 2. Since the first two amino acids of the mature secreted protein are absolutely conserved between all species so far examined and the fact that these two residues also fall within exon 2 suggest their importance for recognition of the signal peptide cleavage site. Figure 5 also indicates, in boxed areas, a pair of imperfect amino acid repeats identified from analysis of the cDNA. Again, these repeats are wholly contained within exons 3 and 4, respectively, located in both exons toward the 3′ terminus. It is possible that this repeat in the osteonectin molecule may have arisen by partial duplication of this exon during evolution, as has been suggested for the repeating structural units of type I collagen (Yamada et al., 1980). It has been proposed that exons of genes may code for functional domains of proteins (Gilbert, 1985). This concept has gained support recently with the cloning of the LDL (Sudhof et al., 1985), fibronectin (Odermatt et al., 1985), and pyruvate kinase (Lonberg & Gilbert, 1985) genes, where in each case analysis of the exon–intron structure shows that identifiable domains of the protein are encoded within discrete exons. Studies are currently in progress to define a three-dimensional model of the osteonectin molecule in order to study the properties of various domains of the molecule. These studies should enable us to define whether the small exons we find encoding osteonectin (40–50 amino acid units) do indeed represent discrete functional or structural units of the protein.

It is interesting to compare the primary structure of the N-termini of osteonectin from various species. It is clear that the first 37 amino acids are homologous between most species. For the mouse protein, SPARC (Mason et al., 1986a), and bovine osteonectin (J. D. Termine et al., personal communication), the complete amino acid sequence is known by prediction from the cDNA sequence. In this case, there is total sequence homology between the "pre" sequence of both species and extensive homology (>90%) throughout the remainder of the molecule. It is notable, then, that of the 22 amino acid differences between bovine and mouse, 12 occur in the first 23 amino acids of the mature protein. In the chicken molecule, too, there is a cluster of amino acid changes in the same region. The rabbit and pig molecules are identical with the bovine molecule through the N-terminus except for an insertion and deletion, respectively, of the amino acid triplet Ala-Glu-Val. These changes are again in the first 23 amino acids with respect to the bovine molecule. It appears, therefore, that osteonectin is highly conserved between species, changes being noted primarily at the N-terminus of the molecule. It is interesting that this relatively variable region appears to be encoded entirely by exon 3, as defined in the bovine gene. Confirmation of this speculation awaits full sequence data for the osteonectin molecule from the other species and determination of the structural organization of the corresponding osteonectin genes.

Added in Proof

Since this paper was submitted, the domain structure of the mouse analogue has appeared in the literature (Engel et al., 1987), indicating multiple functional properties of the protein.

Acknowledgments

We thank Matthew K. Williamson, UCSD, for expert amino acid sequencing work and Marjorie Sullivan for electron micrograph analysis.

References

Domenicucci, C., Wasi, S., Sodek, J., & Goldberg, H. A. (1987) *J. Dent. Res. 66*, 75.

Dziadek, M., Paulsson, M., Aumailley, M., & Timpl, R. (1986) *Eur. J. Biochem. 161*, 455–464.

Eckert, R. L., & Green, H. (1986) *Cell (Cambridge, Mass.) 46*, 583–589.

Engel, J., Taylor, W., Paulsson, M., Sage, H., & Hogan, B. (1987) *Biochemistry 26*, 6958–6965.

Fisher, L. W., Hawkins, G. R., Tuross, N., & Termine, J. D. (1987) *J. Biol. Chem. 262*, 9702–9708.

Gehron Robey, P., & Termine, J. D. (1985) *Calcif. Tissue Int. 37*, 453–460.

Gehron Robey, P., & Fisher, L. W. (1987) in *Calcium Regulation and Bone Metabolism: Basic and Clinical Aspects* (Cohn, D. V., Martin, T. J., & Meunier, P. J., Eds.) Vol. 9, pp 438–443, Elsevier Science, New York.

Gehron Robey, R., Findlay, D. M., Young, M. F., Beresford, J. N., Stubbs, J., Fisher, L. W., & Termine, J. D. (1986) *J. Bone Miner. Res. 1*, 288.

Gehron Robey, P., Fisher, L. W., Stubbs, J. T., & Termine, J. D. (1987) *UCLA Symp. Mol. Cell. Biol., New Ser. 46*, 115–126.
Geliebter, J., Zeff, R. A., Melvoid, R. W., & Nathenson, S. G. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3371–3375.
Gilbert, W. (1985) *Science (Washington, D.C.) 235*, 823–824.
Hewick, R. M., Hunkapiller, M. W., Hood, L. E., & Dreyer, W. J. (1981) *J. Biol. Chem. 256*, 7990–7997.
Hunkapiller, M. W., & Hood, L. E. (1983) *Methods Enzymol. 91*, 486–493.
Lonberg, N., & Gilbert, W. (1985) *Cell (Cambridge, Mass.) 40*, 81–90.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Mason, I. J., Taylor, A., Williams, J. G., Sage, H., & Hogan, B. L. M. (1986a) *EMBO J. 5*, 1465–1472.
Mason, I. J., Murphy, D., Munke, M., Franke, U., Elliot, R. W., & Hogan, B. L. M. (1986b) *EMBO J. 5*, 1831–1837.
Miyada, C. G., Klofelt, C., Reyes, A. A., McLaughlin-Taylor, E., & Wallace, R. B. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2890–2894.
Odermatt, E., Tamkun, J. W., & Hynes, R. O. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6571–6575.
Peralta, E. G., Winslow, J. W., Peterson, G. L., Smith, P. H., Ashkenazi, A., Ramachandran, J., Shimerlik, M. I., & Capon, D. J. (1987) *Science (Washington, D.C.) 236*, 600–605.
Romberg, R. W., Werness, P. G., Lollar, P., Riggs, B. L., & Mann, K. G. (1985) *J. Biol. Chem. 260*, 2728–2736.
Romberg, R. W., Werness, P. G., Riggs, B. L., & Mann, K. G. (1986) *Biochemistry 25*, 1176–1180.
Sage, H., Tupper, J., & Bramson, R. (1986) *J. Cell. Physiol. 127*, 373–387.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Smith, B. D., Martin, G. R., Miller, E. J., Dorfman, A., & Swarm, R. (1975) *Arch. Biochem. Biophys. 166*, 181–186.
Stenner, D. D., Tracy, R. P., Riggs, B. L., & Mann, K. G. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6892–6896.
Sudhof, T. C., Goldstein, J. L., Brown, M. S., & Russell, D. W. (1985) *Science (Washington, D.C.) 228*, 815–822.
Termine, J. D., Belcourt, A. B., Conn, K. M., & Kleinman, H. K. (1981a) *J. Biol. Chem. 256*, 10403–10408.
Termine, J. D., Kleinman, H. K., Whitson, S. W., Conn, K. M., McGarvey, M. L., & Martin, G. R. (1981b) *Cell (Cambridge, Mass.) 26*, 99–105.
Tilgham, S. M., Curtis, P. J., Tiemeier, D. C., Leder, P., & Weissman, C. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 1309–1313.
Wieringa, B., Meyer, F., Reisser, J., & Weissman, C. (1983) *Nature (London) 301*, 38–43.
Yamada, Y., Avvedimento, V., Mudryji, M., Ohkubo, H., Vogeli, G., Irani, M., Pastan, I., & deCrombrugghe, B. (1980) *Cell (Cambridge, Mass.) 22*, 887–892.
Yang, J.-Q., Remmers, E. F., & Marcu, K. B. (1986) *EMBO J. 5*, 3553–3562.
Young, M. F., Bolander, M. E., Day, A. A., Ramis, C. I., Gehron Robey, P., Yamada, Y., & Termine, J. D. (1986) *Nucleic Acids Res. 14*, 4483–4497.

Effects of Eucaryotic Initiation Factor 3 on Eucaryotic Ribosomal Subunit Equilibrium and Kinetics†

Dixie J. Goss,*,‡ Donna Rounds,‡ Teresa Harrigan,‡ Charles L. Woodley,§ and Albert J. Wahba§
Department of Chemistry, Hunter College of the City University of New York, New York, New York 10021, and Department of Biochemistry, The University of Mississippi Medical Center, Jackson, Mississippi 39216
Received April 22, 1987; Revised Manuscript Received October 8, 1987

ABSTRACT: In order to understand the possible role of eucaryotic initiator factor 3 (eIF-3) in maintaining a pool of eucaryotic subunits, we have measured the effects of eIF-3 on the equilibria and kinetics of ribosomal subunit association and dissociation. The ribosomal subunit interactions have been studied by laser light scattering, which does not perturb the system. We find that eIF-3 reduces the apparent association rate of reticulocyte, wheat germ, and *Artemia* ribosomes. The kinetics of the reassociation for a shift in [Mg^{2+}] from 0.5 to 6 mM are best explained by a model where eIF-3 dissociates from the 40S subunits prior to association of the 40S and 60S subunits. Static titrations indicate there is some binding of eIF-3 to 80S ribosomes at lower [Mg^{2+}].

The Mg^{2+}-dependent association and dissociation of ribosomal subunits play a key role in the initiation of protein synthesis in both eucaryotes and procaryotes. This process is regulated by initiation factor 3 in procaryotic systems [see Maitra et al. (1982) for a review]. The procaryotic initiation factor 3 is a single polypeptide of about M_r 23 000.

Eucaryotic initiation factor 3 (eIF-3) has been isolated by several groups from different sources (Schreier & Staehelin, 1973; Safer et al., 1976; Benne & Hershey, 1976; Trachsel et al., 1977; Thompson et al., 1977; Brown-Luedi et al., 1982; Nygard & Westermann, 1983; Spremulli et al., 1979; Seal et al., 1983). In all cases, eIF-3 was shown to be a large multicomponent protein complex composed of at least nine different polypeptides. Benne and Hershey (1976) reported that rabbit reticulocyte eIF-3 bound to the 40S and not the 60S ribosomal subunits. Trachsel and Staehlin (1979), using glutaraldehyde fixation, found eIF-3 bound to the 40S and not the 60S ribosomal subunit. Behlke et al. (1986) have demonstrated by electron microscopy the binding of eIF-3 to

† Supported in part by grants from Research Corp., a PSC-CUNY Faculty Research award, an American Heart Association-NYC Affiliate Investigator award (D.J.G.), and NSF DMB 86-007070 and NIH GM 24451.
* Address correspondence to this author.
‡ Hunter College of the City University of New York.
§ The University of Mississippi Medical Center.

0006-2960/88/0427-1489$01.50/0