Novel Regulators of Bone Formation: Molecular Clones and Activities

JOHN M. WOZNEY, VICKI ROSEN, ANTHONY J. CELESTE, LISA M. MITSOCK, MATTHEW J. WHITTERS, RONALD W. KRIZ, RODNEY M. HEWICK, ELIZABETH A. WANG

Protein extracts derived from bone can initiate the process that begins with cartilage formation and ends in de novo bone formation. The critical components of this extract, termed bone morphogenetic protein (BMP), that direct cartilage and bone formation as well as the constitutive elements supplied by the animal during this process have long remained unclear. Amino acid sequence has been derived from a highly purified preparation of BMP from bovine bone. Now, human complementary DNA clones corresponding to three polypeptides present in this BMP preparation have been isolated, and expression of the recombinant human proteins have been obtained. Each of the three (BMP-1, BMP-2A, and BMP-3) appears to be independently capable of inducing the formation of cartilage in vivo. Two of the encoded proteins (BMP-2A and BMP-3) are new members of the TGF-β supergene family, while the third, BMP-1, appears to be a novel regulatory molecule.

THE STRUCTURAL MASS, STRENGTH, AND SHAPE OF BONE tissue are maintained through a continuous balance between its formation and destruction. The balance is shifted toward bone formation by mechanical stress such as exercise and toward destruction by disuse or by chronic disease states such as osteoporosis. The regenerative capacity of bone is substantial, enabling it to recover from severe breaks and cracks by repeating the developmental process characteristic of embryonic skeletal formation.

One model for analyzing bone growth and repair is the induction of bone formation at an extraskeletal ectopic site. Living transitional epithelium and various epithelial cells cause bone formation when transplanted into the connective tissue of an animal (*1*). Whether this phenomenon is due to a diffusible factor, extracellular matrix components, cell-cell interactions, or a combination of these has not been determined. Urist in his pioneering work (*2*) demonstrated that extracts from demineralized bone could induce new bone formation if implanted into ectopic sites in rodents. Extensive characterization of this process indicates that it mirrors the normal process of in vivo cartilage and bone formation (*3*). The implant is invaded by migrating mesenchyme which differentiates over a period of approximately 7 days into cartilage forming cells (chondroblasts and chondrocytes). Bone cells (osteoblasts and osteoclasts) then appear in the area, and are responsible for the gradual removal of cartilage and the deposition of new bone in its place. The end result is the replacement of the originally implanted material with an ossicle of new bone complete with functional marrow. The active component in the bone extract was identified as being proteinaceous and named bone morphogenetic protein (BMP) by Urist and collaborators many years ago, although it was unclear whether BMP activity could be ascribed to any single molecule (*4*).

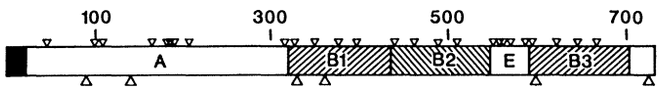
Purification and characterization of BMP remained until recently an elusive goal, partially because of the cumbersome in vivo assay that defines it. Several proteins thought to have BMP activity have been isolated from bovine and human bone (*5, 6*) and from a murine osteosarcoma (*7*). The molecular sizes and isoelectric points indicate that these are all different proteins, and none of these has been sequenced or cloned. Many well-known growth factors and cytokines have been implicated in bone growth and repair (*8*). Factors such as fibroblast growth factor (FGF), transforming growth factor β (TGF-β), and platelet-derived growth factor (PDGF) can be found in bone, have various effects on cartilage and bone cells in vitro, and have been proposed to contribute to BMP activity. However, none has been shown to induce cartilage or bone growth in vivo.

The physical properties and specific activity of our highly purified bovine bone protein preparation with BMP activity (*9*) distinguish it from any previously described factor. In this purification procedure, the final step was SDS–polyacrylamide gel electrophoresis under nonreduced conditions; the activity was recovered in a fraction corresponding to a size of 30 kD. After reduction, which destroys biological activity, the fraction yielded three major individual polypeptides with sizes of 30, 18, and 16 kD. This mixture of polypeptides represents a purification of approximately 300,000-fold with respect to the initial bone extract, and only 40 μg could be obtained from 40-kg quantities of bovine bone powder. In that further biochemical purification to determine which of the three components was necessary for BMP activity would require additional bovine material and time, we cloned each of the polypeptides in the mixture and then determined the biological activities of each of the recombinant molecules. The mixture of polypeptides was digested with trypsin and the amino acid sequences of several individual tryptic peptides were obtained. We now describe the isolation of full-length complementary DNA's (cDNA's) encoding the human equivalents of three polypeptides originally purified from bovine bone and present data on the biochemical and biological characterization of these proteins, which we have designated BMP-1, BMP-2A, and BMP-3. While BMP-1 appears to be unrelated to other known growth factors, the derived amino acid sequences of BMP-2A and BMP-3 indicate that they are closely related to proteins

The authors are with the Tissue Growth and Repair Program at Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140.

involved in embryonic morphogenesis in other animal systems, and we regard them as new members of the growth factor family that includes TGF-β and the inhibins. We have also identified by cross-hybridization with a BMP-2A probe another cDNA encoding a protein called BMP-2B, which is still another member of the same growth factor family. Our findings indicate that the BMP activity in the original preparation is due to a mixture of regulatory molecules and that the complex development process of cartilage and bone formation is most likely controlled, at least in part, by the interactions of these molecules.

BMP-1 cDNA's. At the outset it was unclear which tryptic fragments were derived from which of the three polypeptides present in the BMP preparations. We therefore obtained recombinant DNA clones corresponding to each tryptic sequence individually. In order to obtain human cDNA's, even though the amino acid sequences were obtained from bovine protein, we first isolated a bovine genomic clone, and then used this recombinant as a probe to determine a suitable human messenger RNA (mRNA) source and thus isolate a human cDNA encoding the entire protein.




```
  1  MPGVARLPLLLGLLLLPRPGRPLDLADYTYDLAEEDDSEPLNYKDPCKAA  50
 51  AFLGDIALDEEDLRAFQVQQAVDLRRHTARKSSIKAAVPGNTSTPSCQST 100
101  NGQPQRGACGRWRGRSRSRRAATSRPERVWPDGVIPFVIGGNFTGSQRAV 150
151  FRQAMRHWEKHTCVTFLERTDEDSYIVFTYRPCGCCSYVGRRGGGPQAIS 200
201  IGKNCDKFGIVVHELGHVVGFWHEHTRPDRDRHVSIVRENIQPGQEYNFL 250
251  KMEPQEVESLGETYDFDSIMHYARNTFSRGIFLDTIVPKYEVNGVKPPIG 300
301  QRTRLSKGDIAQARKLYKCPACGETLQDSTGNFSSPEYPNGYSAHMHCVW 350
351  RISVTPGEKIILNFTSLDLYRSRLCWYDYVEVRDGFWRKAPLRGRFCGSK 400
401  LPEPIVSTDSRLWVEFRSSSNWVGKGFFAVYEAICGGDVKKDYGHIQSPN 450
451  YPDDYRPSKVCIWRIQVSEGFHVGLTFQSFEIERHDSCAYDYLEVRDGHS 500
501  ESSTLIGRYCGYEKPDDIKSTSSRLWLKFVSDGSINKAGFAVNFFKEVDE 550
551  CSRPNRGGCEQRCLNTLGSYKCSCDPGYELAPDKRRCEAACGGFLTKLNG 600
601  SITSPGWPKEYPPNKNCIWQLVAPTQYRISLQFDFFETEGNDVCKYDFVE 650
651  VRSGLTADSKLHGKFCGSEKPEVITSQYNNMRVEFKSDNTVSKKGFKAHF 700
701  FSEKRPALQPPRGRPHQLKFRVQKRNRTPQ                      730
```

Fig. 1. Amino acid sequence of human BMP-1 derived from a cDNA clone (λU20S1-1). The human homologs of the two tryptic peptides derived from bovine BMP that were used in the isolation of the clone are underlined. The diagram above the sequences illustrates the putative domain structure of BMP-1; the triangles above mark the position of the 27 cysteine residues, while the triangles below indicate the six potential sites for N-linked glycosylation. A bovine BMP-1 clone was first isolated from a genomic library constructed from a Sau 3A partial digest of bovine liver DNA inserted into the Bam HI vector EMBL3. The library was screened with two oligonucleotide probes (*33*). Probe 1 was TCCTCATCCAGGGCAATGTCGCCCAGGAAGGC; probe 2 was a mixture of ARRTCYTCYTCRTCYAA and ARRTCYTCYTCRTCNAG (192-fold degenerate). One recombinant, which hybridized to both probes, was plaque purified and the hybridizing region was localized to a 0.8-kb Eco RI fragment. This fragment was isolated, labeled by nick-translation, and used to isolate a clone from a human genomic library (*34*). The sequence of the human exon was synthesized and cloned into an M13 vector. With this as a probe, polyadenylated RNA's (3 μg) derived from various human cell lines were examined for hBMP-1 mRNA expression (*35*). An oligo(dT) primed cDNA library in λgt10 was synthesized from U-2 OS polyadenylated RNA (*34*). Recombinants (1×10^6) of this library were screened in duplicate with the synthetic human BMP-1 exon sequence. One duplicate positive (λU2OS1-1) was obtained. Two other cDNA clones were obtained by rescreening of 500,000 recombinants of the same library with the 5′ Eco RI fragment (1.2-kb) of λU2OS1-1. Ambiguities in the nucleotide sequences are: R is A or G; Y is C or T; N is any nucleotide; and D is A, G, or T.

The tryptic fragment containing the amino acid sequence AAFLGDIALDEEDLG (*10*) was used to design two different types of oligodeoxyribonucleotide probes: one a relatively long oligonucleotide; the other a pool of 17-nucleotide (nt) probes. A recombinant from a bovine genomic library was obtained which hybridized to both of these probes. The DNA sequence analysis of this recombinant indicated that it encoded the desired tryptic sequence. In addition, the amino acid sequence of another tryptic fragment of the bovine protein fraction, AFQVQQAADL, is also contained within the same coding region immediately following the original tryptic sequence. We have named the protein encoded by this gene BMP-1.

RNA hybridization analysis indicated that the human osteosarcoma cell line U-2 OS synthesized two low-abundance BMP-1 mRNA's of approximately 4.0 and 2.8 kb. Subsequent screening of a U-2 OS cDNA library resulted in the isolation of three recombinant clones. The complete nucleotide sequence of the 2.5-kb insert of one of these contains an open reading frame of 2190 bp (*11*), predicting a primary translation product of 730 amino acids (Fig. 1). The first in-frame methionine codon is followed by a string of hydrophobic amino acid residues characteristic of a leader sequence of a secreted protein.

The primary structure of BMP-1 indicates that the protein can be divided into distinct domains (Fig. 1). Domain A contains the sequences corresponding to the two tryptic peptides found in bovine BMP. A search of the NBRF (National Biomedical Research Foundation) protein database (*12*) with the BMP-1 sequence shows that domain A also contains a region of sequence similarity (amino acid residues 121 to 321) with a protease isolated from crayfish (*13*), which has an unusual cleavage specificity and does not appear to be related to any known protease family. Whether BMP-1 has protease activity is unknown. BMP-1 also contains a small region (50 amino acids) of sequence similarity to epidermal growth factor (EGF) and the EGF-like domains of various proteins (domain E). These EGF-like regions have been found in a number of different proteins including clotting factors, structural proteins such as a fibroblast proteoglycan (*14*) and cartilage matrix protein (*15*), and several neurogenic gene products (*16*). Although the role of this region in BMP-1 is unclear, various functions have been postulated for EGF-like domains, including calcium ion binding and protein-protein interactions. A comparison of the BMP-1 amino acid sequence with itself indicates that the other domains (B1, B2, B3) are regions of internal sequence similarity. Each block of sequence is 113 amino acid residues long; the first two are tandemly repeated (residues 322 to 434 and 435 to 547), the third (591 to 703) follows the EGF-like domain. An alignment of these three sequences shows that in each block there are four absolutely conserved Cys residues, suggesting a common secondary conformation for these domains.

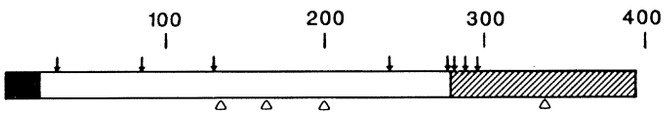
BMP-2A and BMP-2B cDNA's. In a similar manner, a bovine genomic clone encoding the tryptic sequence NYQDMVVEG (*10*) was isolated through the use of two (pooled) 17-nt probes. A portion of this recombinant was used as a probe to detect a corresponding human mRNA from U-2 OS. The U-2 OS cDNA library was hybridized with the bovine probe, and two classes of human cDNA clones were evident, as indicated by strong or weak hybridization signals under stringent washing conditions. The strongly hybridizing clones have extensive homology to the probe and clearly represent the human homolog; the protein encoded by these cDNA's is designated BMP-2A. Sequence analysis of the more weakly hybridizing clones indicates that they are quite similar to the bovine BMP-2 gene at the 3′ end, but less so in their 5′ portions. Because of its high homology in the carboxyl-terminal portion, we have termed the encoded protein BMP-2B. Analysis of a 1.6-kb BMP-2A and a 1.9-kb BMP-2B clone indicates that they contain

open reading frames of 1188 bp and 1224 bp, sizes from which we predict the hBMP-2A and hBMP-2B precursor proteins to be 396 and 408 amino acid residues long, respectively (Fig. 2). The predicted amino acid sequences of these polypeptides both start with hydrophobic strings of amino acids characteristic of secreted proteins. BMP-2A and BMP-2B each contain four potential NH_2-linked glycosylation sites, three of them at corresponding positions (Fig. 2).

BMP-3 cDNA's. We were able to obtain sequence information from one bovine BMP preparation significantly enriched in the 16-kD reduced protein, and therefore assign several of the tryptic peptide sequences of the 30-kD nonreduced bovine BMP material to this component. The tryptic peptide sequences SLKPSNHATIQSIV, SFDAYYCS, and VYPNMTVESCA (*10*) were used to design probes consisting of three different 17-nt pools. One bovine genomic recombinant was obtained that hybridized to all three probes; DNA sequence analysis showed that it encoded all the desired tryptic sequences as well as another (VDFADI), which was also found in the 16-kD reduced polypeptide. We have named the protein encoded by this gene BMP-3.

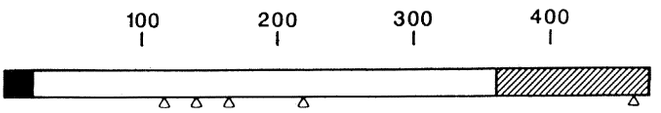
Clones for the human BMP-3 gene were obtained with portions of the bovine BMP-3 gene as probes. These clones were used to determine a mRNA source for hBMP-3. Expression of BMP-3 mRNA was detected in the human lung small cell carcinoma cell line H128, and cDNA clones encoding the entire hBMP-3 (human) protein were obtained from a primer-extended H128 cDNA library. The DNA sequence of these recombinants indicates that they encode a 472–amino acid residue hBMP-3 protein (Fig. 3). The COOH-terminal portion of BMP-3 has about 49 percent sequence similarity to BMP-2A and BMP-2B (see below).

Expression of BMP cDNA's. To determine the in vivo activities of the BMP polypeptides, the cDNA inserts of the BMP-1, BMP-2A, and BMP-3 clones were transferred into derivatives of the mammalian cell expression vector pMT2 (*17*) and transfected into monkey (COS-1) cells. Characterization of the recombinant BMP polypeptides synthesized by the COS cells through pulse-labeling with [^{35}S]methionine showed expression of the expected full-length




```
  1 MVAGTRCLLALLLPQVLLGGA..AGLVPELGRRKFA..AASSGRPSSQPS 46
  1 MIPGNRMLMVVLLCQVLLGGASHASLIPETGKKKVAEIQGHAGGRRSGQS 50

 47 DEVLSEFELRLLSMFGLKQRPTPSRDAVVPPYMLDLYRRHSG....QPGS 92
 51 HELLRDFEATLLQMFGLRRRPQPSKSAVIPDYMRDLYRLQSGEEEEEQIH 100

 93 PAPDHRLERAASRANTVRSFHHEESLEELPETSGKTTRRFFFNLSSIPTE 142
101 STGLEYPERPASRANTVRSFHHEEHLENIPGTSENSAFRFLFNLSSIPEN 150

143 EFITSAELQVFREQMQDALGNNSSFHHRINIYEIIKPATANSKFPVTRLL 192
151 EVISSAELRLFREQVDQGPDWERGFHRINIYEVMKPPAEVVPGHLITRLL 200

193 DTRLVNQNASRWESFDVTPAVMRWTAQGHANHGFVVEVAHLEEKQGVSKR 242
201 DTRLVHHNVTRWETFDVSPAVLRWTREKQPNYGLAIEVTHLHQTRTHQGQ 250

243 HVRISRSLHQDEHSWSQIRPLLVTFGHDGKGHPLHKREKRQAKHK....Q 288
251 HVRISRSLPQGSGNWAQLRPLLVTFGHDGRGHALTRRRRAKRSPKHHSQR 300

289 RKRLKSSCKRHPLYVDFSDVGWNDWIVAPPGYHAFYCHGECPFPLADHLN 338
301 ARKKNKNCRRHSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCPFPLADHLN 350

339 STNHAIVQTLVNSVNSKIPKACCVPTELSAISMLYLDENEKVVLKNYQDM 388
351 STNHAIVQTLVNSVNSSIPKACCVPTELSAISMLYLDEYDKVVLKNYQEM 400

389 VVEGCGCR 396
401 VVEGCGCR 408
```

Fig. 2. Alignment of BMP-2A (top row) and BMP-2B (bottom row) amino acid sequences derived from cDNA clones. Potential N-linked glycosylation sites are indicated with open triangles. The human homolog of the tryptic peptide derived from bovine BMP is underlined. The schematic diagram above indicates the putative leader sequence (solid black) and mature region (shaded) of BMP-2A; arrows above mark the positions of dibasic residues in the primary amino acid sequence and the triangles below indicate the potential N-linked glycosylation sites. A bovine BMP-2A clone was isolated from a bovine genomic library constructed from a Sau 3A partial digest of bovine liver DNA inserted into the Bam HI vector λJ1. The library was screened with two oligonucleotide probes: probe 1 was ACNACCATRTCYTGRAT (32-fold degenerate); probe 2 was CARGAYATGGTNGTNGA (64-fold degenerate) with the TMAC hybridization procedure (*33*). One recombinant that hybridized to both probes was plaque purified and the oligonucleotide hybridizing region was localized to a Sac I fragment; its DNA sequence indicated it encoded the desired fragment. The 5′ end of the Hind III–Sac I fragment (0.6 kb) of this bovine genomic clone was subcloned into M13. RNA blot analysis with a single-stranded M13 probe from this construct indicated the presence of a 3.8 kb mRNA made by U-2 OS. The oligo(dT) primed U-2 OS cDNA library was screened with the nick-translated Hind III–Sac I bovine BMP-2 fragment under stringent and nonstringent hybridization conditions to yield hBMP-2A and hBMP-2B cDNA clones (*36*).




```
  1 MAGASRLLFLWLGCFCVSLAQGERPKPPFPELRKAVPGDRTAGGGPDSEL  50
 51 QPQDKVSEHMLRLYDRYSTVQAARTPGSLEGGSQPWRPRLLREGNTVRSF 100
101 RAAAAETLERKGLYIFNLTSLTKSENILSATLYFCIGELGNISLSCPVSG 150
151 GCSHHAQRKHIQIDLSAWTLKFSRNQSQLLGHLSVDMAKSHRDIMSWLSK 200
201 DITQFLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYILVYANDAAIS 250
251 EPESVVSSLQGHRNFPTGTVPKWDSHIRAALSIERRKKRSTGVLLPLQNN 300
301 ELPGAEYQYKKDEVWEERKPYKTLQAQAPEKSKNKKKQRKGPHRKSQTLQ 350
351 FDEQTLKKARRKQWIEPRNCARRYLKVDFADIGWSEWIISPKSFDAYYCS 400
401 GACQFPMPKSLKPSNHATIQSIVRAVGVVPGIPEPCCVPEKMSSLSILFF 450
451 DENKNVVLKVYPNMTVESCACR                               472
```

Fig. 3. Amino acid sequence of human BMP-3 derived from a cDNA clone. The human homologs of the four tryptic peptides derived from bovine BMP are underlined. The diagram above indicates the putative leader sequence (solid black) and the mature 16-kD region (shaded) of BMP-3; triangles indicate the potential N-linked glycosylation sites. In order to isolate a bovine BMP-3 genomic clone, three 17-nt probes were synthesized on the basis of the tryptic peptide sequences VYPNMTVESCA (probe 1: ACNGTCATRTTNGGARTA, 64-fold degenerate); SFDAYYCS (probe 2: CARTARTANGCRTCRAA, 64-fold degenerate); and SLKPSNHATIQSIV (probe 3: TGDATNGTNGCRTGRTT, 192-fold degenerate). About 400,000 recombinants of a bovine genomic library were screened with probe 1 in TMAC (*33*). All positives were replated, three nitrocellulose replicas were made and screened with probes 1, 2, and 3 again in TMAC. One triplicate positive was plaque purified and shown to encode the three tryptic peptides. Human BMP-3 genomic clones were obtained by screening a human genomic library (*34*) with the oligonucleotides designed from the bovine sequence (*37*). Human BMP-3 mRNA was detected in the cell line H128 (ATCC HTB 120) by ribonuclease protection. A primer-extended H128 cDNA library in λgt10 was made with the AATGATTGAATTAAGCAATTC as a primer. This oligonucleotide was synthesized on the basis of the DNA sequence of the 3′ untranslated region of the human BMP-3 gene. A 350-bp portion of the bovine genomic BMP-3 clone was subcloned into pSP65 and amplified; the insert was excised and labeled with ^{32}P by nick-translation. Of 375,000 recombinants screened with this probe, 17 positives were obtained. One of these contains all the coding information for hBMP-3.

as well as smaller molecular size forms of each protein. The level of expression was dependent on the individual BMP being expressed, but in the best case was less than 1 percent of the total protein produced. Direct implantation of crude COS-conditioned medium into rats resulted in no reproducible cartilage formation, due to the low level of expression coupled with the inflammatory response generated in the animals because of the presence of 0.5 percent fetal calf serum in the medium. All of the BMP's bound to heparin-Sepharose as determined by experiments with pulse-labeled [^{35}S]methionine and this fractionation resulted in an approximately tenfold purification. Activity of the fraction bound by heparin-Sepharose was therefore examined. Rats were implanted with each BMP protein corresponding to amounts ranging from 0.2 to 10.0 ml of crude COS cell–conditioned medium (approximately 5 to 250 μg of protein). Each of the partially purified BMP-1, BMP-2A, and BMP-3 proteins directed de novo cartilage formation in a dose-related manner (Table 1). No activity was seen when purified medium from mock-transfected COS cells was implanted. The dose response generated by the COS cell–expressed material was limited by its purity, and further studies on the activities of the recombinant BMP's were done with CHO (Chinese hamster ovary) cells or *Escherichia coli* expression systems where larger quantities of the proteins are synthesized.

After amplification of the mammalian cell expression constructs, we find that CHO cells synthesize different molecular sizes of each of the BMP proteins. The protein species and amounts of each protein observed are dependent on the individual cell line. When bound to heparin-Sepharose, serum-free conditioned medium from CHO cells making BMP-1, BMP-2A, or BMP-3 induces the formation of new cartilage in the rat implant system after 7 days (Fig. 4). The cartilage formed shows the developmental progression of chondroblasts from resting cells through the hypertrophic stage. The cells are organized in columns typical of the orderly array seen in the cartilage growth plate of developing long bone. The ratio of hypertrophied to resting chondroblasts varies from one cartilage nodule to another (compare Fig. 4, A and B). No cartilage formation is observed when material that is not bound by heparin is implanted, and only the acellular carrier matrix particles are present (Fig. 4D).

Cartilage formation is also induced by recombinant BMP-1 expressed in *E. coli*. A 50-kD, NH_2-terminal fragment of BMP-1 was expressed in *E. coli*, solubilized from washed inclusion bodies, and purified by heparin-Sepharose affinity chromatography to approximately 50 percent purity, although the amount of BMP-1 in an active conformation is not known. When 300 or 100 ng of this material was implanted in vivo, comparatively large areas of cartilage were formed (Fig. 5, A and B). Implantation of only 10 ng resulted in significantly less cartilage formation (Fig. 5C), while a similarly prepared negative control of the 30-kD COOH-terminal BMP-1 fragment had no activity (Fig. 5D). No activity was seen in similar experiments with various constructions of BMP-2A or BMP-3.

The TGF-β family. BMP-2A, BMP-2B, and BMP-3 are new members of the TGF-β family of growth and differentiation factors. An alignment of the COOH-terminal portions of the BMP's with eight other members of the family (Fig. 6A) shows that there are

Table 1. In vivo activities of BMP proteins expressed by COS cells. Each pMT2CX plasmid (7 μg) containing BMP-1, BMP-2A, or BMP-3 cDNA's was transfected per one 100-mm plate of COS cells (*17*). Conditioned medium with 0.5 percent fetal calf serum (4 ml per plate) from each transfection was concentrated, dialyzed against 20 m*M* tris, *p*H 7.4, 0.15*M* NaCl, and applied to a heparin-Sepharose column. Bound proteins were eluted with 20 m*M* tris, 2.0*M* NaCl and assayed for cartilage induction in rats (*9*). Results are summarized from several different experiments. Activity scores are as described (*9*): a score of +2 represents 20 to 30 percent of the section to be cartilage, a score of +1 represents 10 to 20 percent to be cartilage, and so forth. BMP-1 and BMP-2A were implanted for 10 days; BMP-3 for 6 days; purified conditioned medium from mock transfections showed no activity at 6 or 10 days. The amount of protein implanted is given in milliliter equivalents of conditioned medium from the COS cell transfections.

Protein	Conditioned medium (ml)	Activity	Protein	Conditioned medium (ml)	Activity
BMP-1	11.7	+2	BMP-3	4.5	0
	4.0	+2		1.5	0
	2.6	+1		0.5	0
	0.7	0		0.2	+2
				0.17	+1
BMP-2A	4.0	0	Mock	10.4	0
	0.9	+2		3.4	0
	0.8	+1		1.1	0
	0.16	0		0.67	0
				0.33	0
				0.17	0

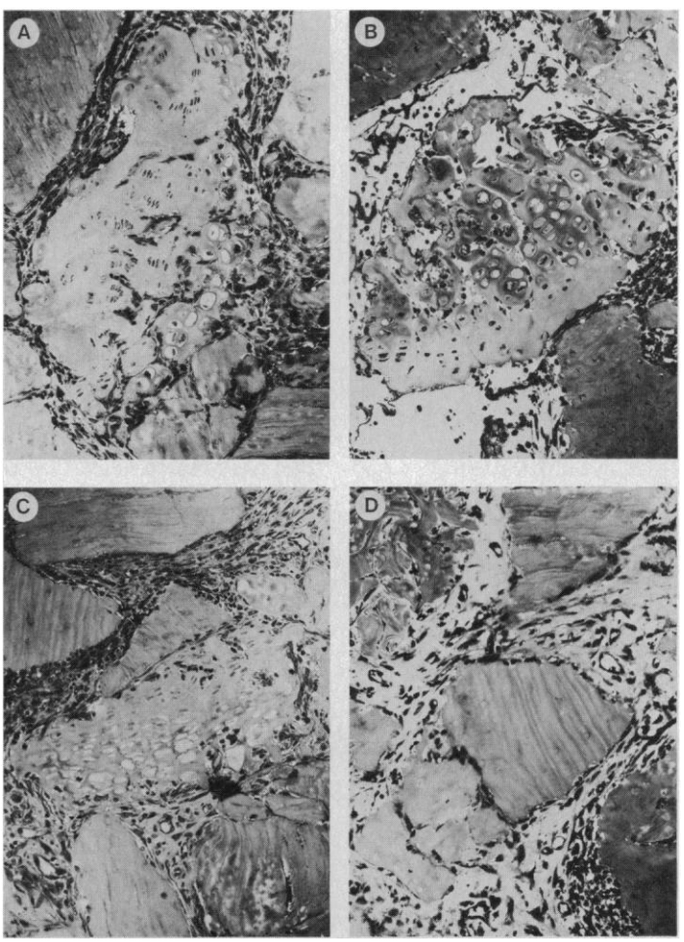


Fig. 4. Cartilage-inducing activity of human recombinant BMP proteins. Photographs (×75) of toluidine blue–stained sections of 7-day implants of BMP proteins partially purified from the conditioned media from CHO cells transfected with the BMP cDNA's. (**A**) BMP-1. Approximately 20 ng of total protein was implanted. The scoring for this sample was C+2. (**B**) BMP-2A, 40 ng of total protein, C+2. (**C**) BMP-3, 40 ng total protein, C+1. (**D**) Implantation of heparin nonbound material results in no cartilage formation. Serum-free conditioned media was collected at 24 hours, concentrated, loaded directly on a heparin-Sepharose column at 1 to 5 mg of protein per milliliter of gel, fractionated, and assayed (*9*). The purity of each material was not determined. The quantity of material was estimated by absorbance at 280 nm.

three general regions of sequence conservation: residues 1 to 36; the residues immediately surrounding the cysteine pair at residues 61,62; and the COOH-terminal ten residues. Seven Cys residues are absolutely conserved amongst all 11 proteins. A pairwise comparison of these COOH-terminal domains shows that BMP-2A and BMP-2B are very similar in this region (92 percent), indicating that they are even more closely related than TGF-β1 and TGF-β2 (Fig. 6B); BMP-2A is clearly the human homolog of the protein present in bovine bone. The identical tryptic peptide sequence (NYQDMVVEG) found in the bovine protein is present in hBMP-2A; the corresponding sequence in hBMP-2B has one amino acid difference (NYQEMVVEG). It is not yet clear whether the bovine homolog of BMP-2B is present in bovine bone or any other tissue.

The three BMP's are most homologous to the inhibin β subunits and to two factors implicated in developmental processes in lower animals, decapentaplegic (dpp) and Vgl (see below). BMP-2A and -2B each show about 75 percent sequence identity with the *Drosophila* dpp protein, which is involved in dorsal-ventral specification during embryogenesis, and later during development in correct formation of the imaginal disks (*18*). Considering the evolutionary distance between arthropods and chordates, we suggest that one of these BMP's may be the human homolog of the dpp protein. BMP-2A and -2B also have considerable sequence similarity (about 57 percent) with Vgl, as does BMP-3 (49 percent). Vgl mRNA is localized to the vegetal hemisphere of *Xenopus* oocytes, and has been proposed to be a signal used by the endoderm cells to commit ectodermal cells to become the embryonic mesoderm (*19*). This is supported by the finding that TGF-β2 (*20*) or a TGF-β-like factor (*21*), or both, in synergy with FGF will induce mesoderm formation in *Xenopus* embryo explants.

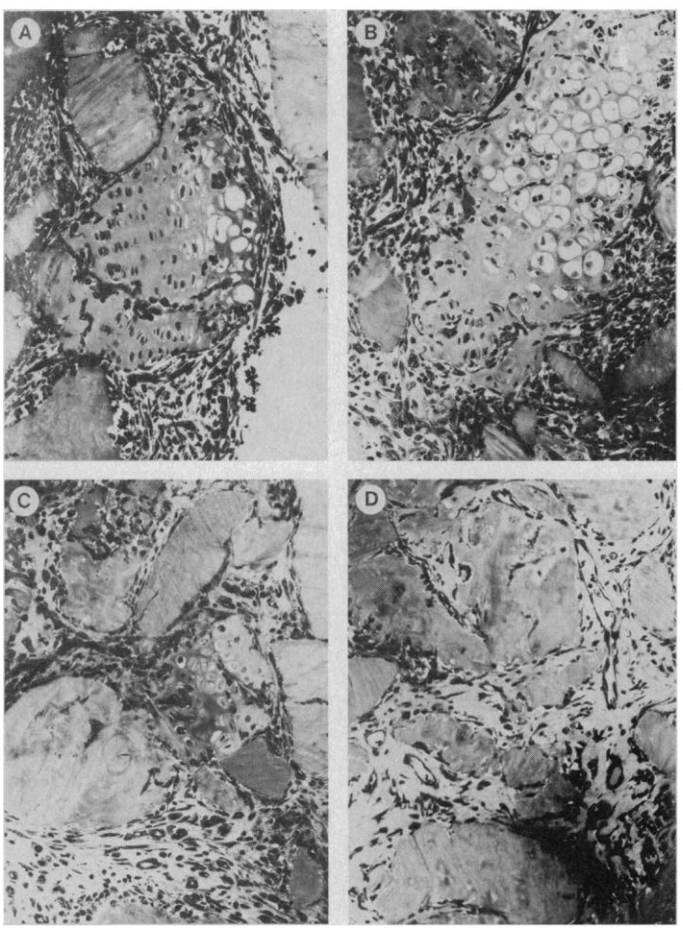


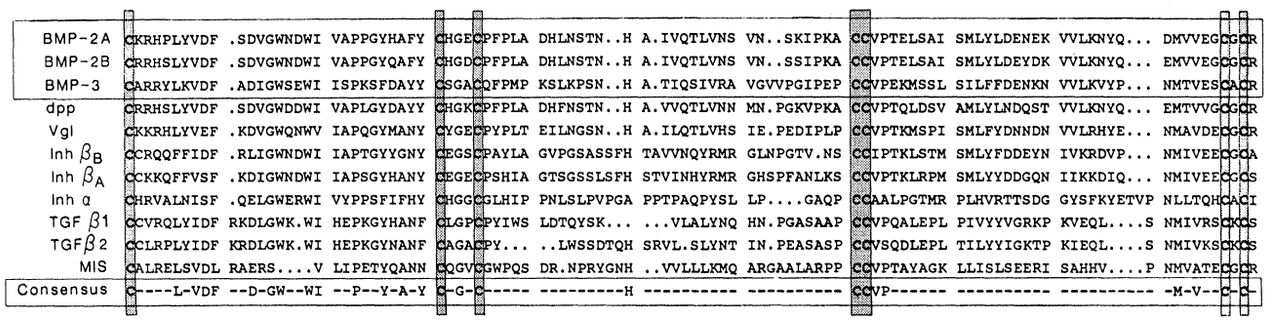
Fig. 5. Cartilage-inducing activity of human recombinant BMP-1 produced by *E. coli*. Photographs (×75) of toluidine-stained sections of 7-day implants of partially purified 50 kD NH_2-terminal fragment of BMP-1; (**A**) 300 ng, C+2; (**B**) 75 ng, C+3; (**C**) 10 ng, C+1. (**D**) is a negative control of the 30-kD COOH-terminal fragment of BMP-1 expressed in *E. coli*. C score was 0. *Escherichia coli* containing a plasmid encoding either residues 26-464 (NH_2-terminal 50 kD) or 543-730 (COOH-terminal 30 kD) of BMP-1 under control of the λp_L promoter (*38*) was grown and inclusion bodies were prepared (*39*). The BMP-1 protein was solubilized from the inclusion bodies in 8*M* urea, 50 m*M* tris, *p*H 8.5, and 5 m*M* dithiothreitol, loaded on a heparin-Sepharose column at 5 mg of protein per milliliter of gel, and fractionated (*9*).

The amino acid sequences of the COOH-terminal portions of the BMP's are next most closely related to the β subunits of the inhibins (about 44 percent for the BMP-2's, 38 percent for BMP-3). The inhibin α, β_A, and β_B polypeptides combine to form a set of molecules with complex, often opposite, activities. The αβ heterodimer inhibits follicle-stimulating hormone (FSH) secretion (*22*), while the β homodimer stimulates FSH secretion (*23*); β_A homodimers induce differentiation of erythroleukemia cells, while the αβ heterodimer inhibits the differentiation induced by the β homodimer (*24*). By analogy with inhibin, the ability of BMP-2A and BMP-3 to form homodimers and heterodimers could create multiple signals that could be used to modulate the bone formation process.

We would expect the processing and subunit structure of BMP-2A and BMP-3 to be similar to other members of the TGF-β family. Information contained in tryptic and NH_2-terminal sequences (*9*) indicates that the 16-kD polypeptide in our purified bovine BMP is encoded by the COOH-terminal portion of the BMP-3 gene. The precursor is cleaved at paired dibasic residues resulting in a mature protein with a predicted size of about 12 to 13 kD. Given the sequence similarity between BMP-2A and BMP-3, we would expect correct processing of the BMP-2A precursor to give a COOH-terminal fragment of 13 to 14 kD. We assume that the difference between the sizes predicted from the amino acid sequence of BMP-3 and BMP-2A and the 16-kD and 18-kD bovine proteins is due to glycosylation. Each of the predicted amino acid sequences contains a potential NH_2-linked glycosylation site, and both purified bovine proteins are sensitive to digestion with endoglycosidase F (*9*). As both proteins migrate at approximately 30 kD under nonreducing conditions in SDS gels, both must be present as dimers; however, whether as homodimers or heterodimers is unknown.

In vivo activity. Each protein that we have identified independently induces cartilage formation in the in vivo assay system we use, a result that has not been seen with any other known growth factor (*6, 25*). This observation is of particular interest as TGF-β1 and TGF-β2, although related to BMP-2A and BMP-3, do not direct new cartilage or bone formation when tested in vivo (*25, 26*) in spite of the fact that they do increase the formation of cartilage specific macromolecules by fibroblasts in vitro (*27*) as well as modulate the expression of other extracellular matrix components by a variety of other cell types (*28*).

It seems quite likely that the BMP activity derived from bovine bone matrix represents the combined action of multiple factors acting at specific points during bone development. Once significant amounts of purified and active protein from each of our recombinant factors are available, a full series of dose response and time course studies can be performed to determine whether the islands of cartilage formation observed in our assay can develop into larger areas of cartilage and bone. While simple mixing of the three BMP proteins may result in this goal, it is more likely that detailed characterization of various subunit compositions and analysis of their modes of action will be necessary to define optimal BMP activity. Thus the configuration and the ratio of these new growth factors may modulate the specificity of the in vivo cartilage or bone-



BMP-2A	29	38	36	23	45	44	58	74	49	92
BMP-2B	29	36	35	22	43	44	56	76	48	
BMP-3	32	34	34	30	37	38	49	43		
dpp	27	37	37	23	40	44	48			
Vgl	32	36	38	23	47	39				
inh β_B	27	38	38	26	64					
inh β_A	25	46	42	27						
inh α	27	28	26							
TGF-β2	24	74								
TGF-β1	31									

Fig. 6. Comparison of BMP-2A, BMP-2B, and BMP-3 to other members of the TGF-β family. (**A**) Alignment of the COOH-terminal regions of these proteins beginning with the first conserved cysteine residue. (**B**) Pairwise comparison between these same domains. Numbers given in (B) are the percentage of identical residues derived with the use of GAP [University of Wisconsin Genetics Computer Group (*40*)] with the GAP weight parameter being set to 3.0 and the GAP length weight parameter being set to 0.2 on the portions of proteins given in (A). Abbreviations: dpp, the protein encoded by the decapentaplegic transcript of *Drosophila* (*18*); Vgl, product of the *Xenopus* transcript present in the vegetal pole of oocytes (*19*); inh β and α, human inhibin β and α subunits (*22*); and MIS, human Mullerian inhibiting substance (*41*).

forming activity. It is also possible that additional factors, derived from bone but without inherent bone-forming activity, are necessary to enhance the activity of the factors we have described. Although we did not expect three distinct proteins to have similar activities in the in vivo assay, it is not unusual for different growth factors or cytokines to exhibit similar activities in particular assay systems (*29*). In addition, as many growth factors have activities distinct from those for which they were originally identified, it will also be of interest to assess the role of the BMP proteins in tissues other than bone.

The unrelated molecule BMP-1, which we isolated along with the two proteins in the TGF-β family, BMP-2A and BMP-3, contains an EGF-like domain and the A domain, which has sequence similarity to a known protease (*13*). The TGF-β1 binding protein also contains EGF-like domains; it has been suggested that it could be a protease involved in the activation of TGF-β1 (*30*). Similarly, nerve growth factor (NGF) has been isolated, complexed to polypeptides with protease activity and with sequence similarity to proteases. Again, these associated proteins have been implicated in the processing of NGF to its activated form (*31*). BMP-1 may also be involved in binding or activation of the other BMP molecules.

Relatively impure preparations of nonrecombinant BMP have already been used to repair nonunion fractures and skull defects in sheep, monkeys, and humans with varying degrees of success (*32*). Obvious limitations of these treatments include the small quantities of active bone-inducing material in the crude preparations, the possible presence of inhibitors of bone formation in the crude mixture, and implantation of BMP across species barriers. Use of local bone induction to treat large bony defects caused by trauma, surgical resection, or periodontal disease might require all the recombinant human factors necessary for classic BMP activity formulated in an appropriate carrier. For treatment of other types of defects, such as damaged cartilage, a subset of these factors locally delivered may be sufficient. Finally, we can now begin to examine what roles these BMP proteins play in systemic skeletal defects such as osteoporosis.

REFERENCES AND NOTES

1. C. B. Huggins, *Arch. Surg.* **22**, 377 (1931); K. Wlodarski, *Exp. Cell Res.* **57**, 446 (1969).
2. M. R. Urist, *Science* **150**, 893 (1965); ——— *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3511 (1973).
3. A. H. Reddi and C. Huggins, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1601 (1972); A. H. Reddi, *Collagen Rel. Res.* **1**, 209 (1981).
4. M. R. Urist, A. Mikulski, A. Lietze, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1828 (1979).
5. M. R. Urist *et al.*, *ibid.* **81**, 371 (1984); M. R. Urist *et al.*, *Methods Enzymol.* **146**, 294 (1987).
6. T. K. Sampath *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7108 (1987).
7. K. Takaoka *et al.*, *Biomed. Res.* **2**, 466 (1981).
8. E. Simpson, *Trends Biochem. Sci.* **9**, 527 (1984); P. V. Hauschka *et al.*, *J. Biol. Chem.* **261**, 12665 (1986); J. T. Triffitt, *Acta Orthop. Scand.* **58**, 673 (1987).
9. E. A. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
10. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
11. The DNA sequences of the cDNA clones described in this article have been submitted to GenBank, Los Alamos National Laboratory, Los Alamos, NM.
12. National Biomedical Research Foundation, Georgetown University Medical Center, Washington, DC; Release 15.0.
13. K. Titani *et al.*, *Biochemistry* **26**, 222 (1987).
14. T. Krusius, K. R. Gehlsen, E. Ruoslahti, *J. Biol. Chem.* **262**, 13120 (1987).
15. W. S. Argraves *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 464 (1987). *U.S.A.* **84**, 464 (1987).
16. E. Knust *et al.*, *EMBO J.* **6**, 761 (1987); M. R. Kelley, S. Kidd, W. A. Deutsch, M. W. Young, *Cell* **51**, 539 (1987).
17. D. D. Pittman and R. J. Kaufman, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2429 (1988).
18. R. W. Padgett, R. D. St. Johnson, W. M. Gelbart, *Nature* **325**, 81 (1987).
19. D. L. Weeks and D. A. Melton, *Cell* **51**, 861 (1987).
20. D. Kimelman and M. Kirschner, *ibid.*, p. 869.
21. F. Rosa *et al.*, *Science* **239**, 783 (1988).
22. A. J. Mason *et al.*, *Nature* **318**, 659 (1985).
23. W. Vale *et al.*, *ibid.* **321**, 776 (1986); J. Ling *et al.*, *ibid.* **321**, 779 (1986).
24. Y. Eto *et al.*, *Biochem. Biophys. Res. Commun.* **142**, 1095 (1987); J. Yu *et al.*, *Nature* **330**, 765 (1987).

25. V. Rosen, J. M. Wozney, E. A. Wang, unpublished results.
26. H. Bentz, R. Armstrong, S. M. Seyedin, in *Development and Diseases of Cartilage and Bone Matrix*, A. Sen and T. Thornhill, Eds. (Liss, New York, 1987), p. 137.
27. S. M. Seyedin *et al.*, *J. Biol. Chem.* **261**, 5693 (1986).
28. R. A. Ignotz and J. Massague, *ibid.*, p. 4337; M. B. Sporn, A. B. Roberts, L. M. Wakefield, R. K. Assoian, *Science* **233**, 532 (1986).
29. S. C. Clark and R. Kamen, *ibid.* **236**, 1229 (1987); E. Rozengurt, *ibid.* **234**, 161 (1986); J. Folkman and M. Klagsbrun, *ibid.* **235**, 442 (1987).
30. K. Miyazono, U. Hellman, C. Wernstedt, C.-H. Heldin, *J. Biol. Chem.* **263**, 6407 (1988).
31. P. J. Isackson, J. C. Dunbar, R. A. Bradshaw, A. Ullrich, *Int. J. Neuroscience* **26**, 95 (1985).
32. D. Ferguson, W. L. Davis, M. R. Urist, W. C. Hurt, E. P. Allen, *Clin. Orthoped. Relat. Res.* **219**, 251 (1987); T. C. Lindholm, T. S. Lindholm, I. Alitalo, M. R. Urist, *ibid.* **227**, 265 (1988); E. E. Johnson, M. R. Urist, G. A. M. Finerman, *ibid.* **230**, 257 (1988); J. Glowacki, L. B. Kaban, J. E. Murray, J. Folkman, J. B. Mulliken, *Lancet* **i**, 959 (1981).
33. Probe I was hybridized in 5× SSC, 0.1 percent SDS, 5× Denhardt's, containing salmon sperm DNA (100 μg/ml) (standard hybridization buffer, SHB) at 45°C and washed in 5× SSC, 0.1 percent SDS at 45°C. Probe 2 was hybridized in 3*M* tetramethylammonium chloride (TMAC), 0.1*M* sodium phosphate, *p*H 6.5, 1m*M* EDTA, 5× Denhardt's, 0.6 percent SDS, salmon sperm DNA (100 μg/ml) at 48°C, and washed in 3*M* TMAC, 50m*M* tris, *p*H 8.0 at 50°C for 1 hour.
34. J. J. Toole *et al.*, *Nature* **312**, 342 (1984).
35. A. J. Celeste *et al.*, *EMBO J.* **5**, 1885 (1986).
36. Duplicate nitrocellulose replicas of 1,000,000 recombinants of the oligo(dT) primed U-2 OS cDNA library were hybridized to the nick-translated Hind III–Sac I bovine BMP-2 fragment in SHB at 65°C, and washed with 1× SSC, 0.1 percent SDS. Twelve duplicate positives were replated for secondaries, duplicate replicas were made, and both sets were hybridized as above. One set of filters was washed with 1× SSC, 0.1 percent SDS, and the other with 0.1× SSC, 0.1 percent SDS; both at 65°C. Four recombinants hybridized strongly (hBMP-2A), and seven hybridized more weakly (hBMP-2B), under the more stringent washing conditions. Full-length BMP-2A and BMP-2B clones were obtained by rescreening of the library with appropriate cDNA's.
37. The human genomic library was screened with the oligonucleotides AATTCCGGGGTTCAATCCATTGCTTTCTTCTTGCCTTCTTCAGGGTCTGT and TTCGCTCCAGCCAATATCTGCGAAGTCCACTTTAAGGTACCGTCTGGCAC in SHB at 50°C and washed in 2× SSC, 0.1 percent SDS at 50°C.
38. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
39. D. G. Kleid *et al.*, *Science* **214**, 1125 (1981).
40. J. Devereux, P. Haeberli, O. Smithies, *Nucleic Acids Res.* **12**, 387 (1984).
41. R. L. Cate *et al.*, *Cell* **45**, 685 (1986).
42. We thank K. Kelleher, L. Haines, and R. Adalpe for their efforts in all the DNA sequence analysis; D. Pittman and B. Murray for the COS cell expression experiments; R. Kaufman, D. Israel, K. Kerns, and J. Nove for CHO expression experiments; J. McCoy for *E. coli* expression experiments; M. Byrne and P. LaPan for RNA protection experiments; J. Judd, T. O'Toole, and D. Vanstone for oligonucleotide synthesis; M. Ryan for help with the protein sequencing; T. Harada, M. J. Kriz, D. Luxenberg, and D. McQuaid for in vivo assays; P. Cordes for histology and photography; D. Merberg and S. Aram for help with computer analysis; R. Kay for promoting the project; and R. Kamen, S. Clark, and E. Fritsch for comments on the manuscript.

11 August 1988; accepted 14 November 1988

You want a laser-scanning microscope and you still believe in Santa Claus?