An Evolutionary and Molecular Analysis of \textit{Bmp2} Expression* \\

Kevin L. Abrams‡, Junwang Xu§, Celine Nativelle-Serpentini§, Shabnam Dabirshahsahebi§, and Melissa B. Rogers§§

Published, JBC Papers in Press, February 2, 2004, DOI 10.1074/jbc.M313531200

The coding regions of many metazoan genes are highly similar. For example, homologs to the key developmental factor bone morphogenetic protein (BMP) 2 have been cloned by sequence identity from arthropods, mollusks, cnidarians, and nematodes. Wide conservation of protein sequences suggests that differential gene expression explains many of the vast morphological differences between species. To test the hypothesis that the regulatory mechanisms controlling this evolutionarily ancient and critical gene are conserved, we compared sequences flanking \textit{Bmp2} genes of several species. We identified numerous conserved noncoding sequences including some retained because the fish lineage separated 450 million years ago. We tested the function of some of these sequences in the F9 cell model system of \textit{Bmp2} expression. We demonstrated that both mouse and primate \textit{Bmp2} promoters drive a reporter gene in an expression pattern resembling that of the endogenous transcript in F9 cells. A conserved Sp1 site contributes to the retinoic acid responsiveness of the \textit{Bmp2} promoter, which lacks a classical retinoic acid response element. We have also discovered a sequence downstream of the stop codon whose conservation between humans, rodents, deer, chickens, frogs, and fish is striking. A fragment containing this region influences reporter gene expression in F9 cells. The conserved region contains elements that may mediate the half-life of the \textit{Bmp2} transcript. Together, our molecular and evolutionary analysis has identified new regulatory elements controlling \textit{Bmp2} expression.

Bone morphogenetic proteins (BMPs)\(^1\) were first identified because of their osteogenic properties (1–3). Subsequently, these developmentally critical proteins were found to be expressed widely in vertebrate embryonic structures (4). BMP2 and BMP4 are closely related proteins involved in key embryonic processes such as epithelio-mesenchymal interactions (5), apoptosis (6–9), and dorsal-ventral axis specification (10). Mice having null mutations in \textit{Bmp2} or \textit{Bmp4} (11, 12), BMP receptor genes (13–16), or the BMP signal transducing SMADs (17–21) die during early embryogenesis. The phenotypes of these mutants prove that BMP signaling is required for vital developmental processes.

The evolutionary conservation of the BMP2 and BMP4 proteins at both the functional and sequence levels is remarkable. Easily recognizable \textit{Bmp} genes have been cloned from distantly related phyla such as arthropods, mollusks, cnidarians, and nematodes. Except for the nematode BMPs, the amino acid sequences of nonvertebrate BMPs are 70–87% identical to human BMP2 (Ref. 22 and references therein). The \textit{Drosophila} protein DPP, which is 71% identical to human BMP2, is functionally interchangeable with mammalian BMPs in a mammalian bone induction assay (23). Conversely, the closely related mammalian BMP4 can rescue the dorsal-ventral axis defects of \textit{Drosophila} lacking DPP function (24). Indeed, entire signaling pathways involving BMPs are conserved. For example, BMPs and their antagonists appear to play analogous, although inverted, roles in dorsal-ventral axis formation in vertebrates and invertebrates (25).

Both BMP2 and dpp in \textit{Drosophila} are expressed in a highly tissue- and stage-specific pattern. Multiple promoters and alternative splicing produce a variety of \textit{dpp} transcripts in \textit{Drosophila} (26). Like BMP2, the \textit{dpp} mRNA has an unusually long 3’-untranslated region with portions highly conserved between flies (27, 28). Our work, and that of others, suggests that mammalian \textit{Bmp2} may be regulated in an equally complex manner (29–34).

The first indication that the vitamin A-derivative retinoic acid (RA) is involved in regulating the \textit{Bmp2} gene was our discovery of its strong induction in F9 embryonal carcinoma cells stimulated to differentiate with RA (35). F9 cells lacking retinoic acid receptor (RAR)\(^*\) fail to express \textit{Bmp2} in response to RA (36). RA also induces the \textit{Bmp2} gene in the developing chick limb (37). Many BMP2-expressing tissues develop abnormally in vitamin A-deficient embryos or after exposure to teratogenic levels of vitamin A or other retinoids. These include the heart and cardiovascular system, limbs, central nervous system, craniofacial structures, and vertebral column (4, 38).

F9 cells are widely used as a model of cellular differentiation and early embryonic development (39, 40). The morphology, pattern of gene expression, and cell cycle of undifferentiated embryonal carcinoma cells resemble that of the totipotent stem cells of the blastocyst inner cell mass. F9 cells differentiate rapidly and synchronously into primitive extraembryonic endoderm upon treatment with RA and into parietal extraembryonic endoderm upon treatment with RA and CAMP analogs (41). This model system has been used to identify genetic ele-
ments controlling the expression of many important developmental genes.

Bmp2 is expressed at three distinct levels in F9 cells (35, 42). Undetectable in undifferentiated stem cells, the Bmp2 transcript is readily detectable in RA-treated cells. The combination of elevated cAMP levels and RA increases the abundance of the Bmp2 transcript 5–6-fold more than RA alone. Cyclic AMP elevation alone induces neither differentiation nor expression of Bmp2. This indicates that maximally expressing Bmp2 requires a synergistic interaction between RA and cAMP.

Reasoning that the regulatory elements controlling this ancient and critical gene were conserved (43), we compared sequences flanking the Bmp2 genes from several species. We tested the function of several conserved noncoding sequences in F9 cells representing three embryonic cell types: stem cells, primitive endoderm, and parietal endoderm. The results suggest that, like other RA-induced genes lacking classical retinoid acid response elements (RAREs), a protein-protein interaction between retinoid receptors and Sp1 contributes to RA responsiveness. In addition, a portion of the 3′-untranslated region that has been conserved for 450 million years influences Bmp2 expression.

EXPERIMENTAL PROCEDURES
DNA Isolation
All plasmids were purified using QIAgen® plasmid purification kits. Genomic DNA was isolated from cells grown on 150-mm dishes, lysed with ligation-mediated PCR lysis buffer, incubated overnight at 37 °C, and purified by phenol/chloroform extraction (44). CV1 cells (African green monkey, Cercopithecus aethiops) were provided by Dr. Michele Glozak. HepG2 cells (human, Homo sapiens), A10 cells (rat, Rattus norvegicus), Madin-Darby canine kidney cells (dog, Canis familiaris), and CHL/JU cells (Chinese hamster, Cricetulus griseus) were from Dr. Richard Pollenz. Dr. Peter Medveczky provided genomic DNA from squirrel monkey (Saimiri sciureus), owl monkey (Aotus trivirgatus), common marmoset (Callithrix jacchus), and chicken (Gallus gallus). Dr. Richard Behringer provided DNA from bat (Carollia perspicillata) and tammar wallaby (Macropus eugenii). Purified calf thymus DNA (Bos taurus) was purchased from Sigma.

Genomic PCR
Genomic PCR conditions were adapted from the ligation-mediated PCR protocol described by Garrity and Wold (44) using these degenerate primers: 5′-ATTNNCCCTCTGTGGTTGTACTGCGC-3′ (nt 1217 to 1192) and 5′-NNNNGGGNNCCAGCTCATTGAAAGAG-3′ (nt 1094 to 1071). Unless otherwise indicated, all nucleotide positions are provided with respect to the murine Bmp2 distal promoter that is 2,201 nt upstream of the initiator codon (ATG). One of any four deoxynucleotides, indicated by “N,” was placed in positions that differed between the mouse and human sequences. Reactions contained 0.01 unit/μl of Vent DNA polymerase (New England Biolabs), 2.5–10.0 ng/μl of genomic DNA, 0.1 pmol/μl of primers, 0.8 mM dNTPs, 5% MgSO4, 20 mM NaCl, 20 mM MgSO4, 0.01% gelatin, 0.1% Triton X-100. Samples were covered with mineral oil and thermocycling was performed in a MasterCycler gradient thermocycler (Eppendorf). The cycling conditions were as follows: 2 cycles: denature at 97 °C for 3 min; anneal, 60 °C for 30 min; extend, 76 °C for 10 min followed by 29 cycles: 97 °C for 1 min, 60 °C for 30 min, and 76 °C for 3 min. The ramping time between denaturation at 97 °C and annealing at 60 °C was set to 1 °C/s. The reactions were separated on a 0.7% agarose gel and visualized by ethidium bromide staining.

Following the initial sequencing of the squirrel monkey and African green monkey Bmp2 sequences, the following degenerate primers were designed using ClustalW alignments of the rodent, primate, and deer Bmp2 genes: 5′-TCCAGCTGCACTACGCTGACCGC-3′; nt 1379 to 1280, 5′-ATGCTCGAGGAGGAGGACG-3′; nt 1459 to 1478, and 5′-CTGCTGAGGAGGAGGACG-3′; nt 1459 to 1478. PCR products were purified using QIAquick Gel Extraction Kit (QIAGEN) and cloned using the QIAeasy T-A Cloning Kit (QIAGEN). All plasmids were sequenced using T7 and Sp6 primers or the primers listed above. This PCR product also was digested with BamHI and XhoI and cloned into the luciferase reporter vector, pGL2Basic (Promega) cut with NheI and XhoI.

Southern Analysis
Following visualization of the genomic PCR products, Southern analysis was performed by standard methods (45). Probes for Southern blot analysis were made by primer extension using Bmp2-specific primers or random primers and [α-32P]dCTP (45).

Sequence Analysis
Sequence was performed by the Molecular Biology Core Facility at the H. Lee Moffitt Cancer Center (Tampa, FL) and the UMDNJ, New Jersey Medical School Molecular Resource Facility (Newark, NJ).

Cloning and Sequencing the Squirrel Monkey Bmp2 Promoter Region
A 2.3-kb squirrel monkey Bmp2 PCR product was isolated by gel electrophoresis and Sephaglas™ (Amersham Biosciences) purification. After adding an A to the ends with TaqDNA polymerase (Promega), the fragment was cloned into the pCR II TA cloning vector (Invitrogen) according to the manufacturer’s instructions. The resulting clone, pCRISMBmp2-3, Bmp2 was digested with NotI, BamHI, Smal or, NheI and SpeI, or SalI and Xhol to remove 1-, 1.2-, 1.1- and 1.3-kb fragments, respectively. The ends were religated to make pCRISMBmp2-3 Not, pCRISMBmp2-3 Bam, pCRISMBmp2-3 Smal, pCRISMBmp2-3 NheI, and pCRISMBmp2-3 Sal, respectively, pCRISMBmp2-5 Not was digested with SstII and the ends religated to create pCRISMBmp2-5 Not SstII.

Each plasmid was sequenced using T7 and Sp6 primers or the following squirrel monkey-specific primers: SMB2aF, 5′-CGTCCACACCCCTGCGCGCN-3′, SMB2bR, 5′-AGGTCGCTGGCGGACGCTGT. SMB2aR1, 5′-GACCCCACGCTAGCCACAAA, SMB2bF, 5′-TTACTTTTTCCTCC, SMB2aR2, 5′-GGTCGCTGGCGGACGCTGT.

Cloning and Sequencing the African Green Monkey Bmp2 Promoter Region
A 1696-nt fragment was amplified from African green monkey genomic DNA and sequencing was performed using the degenerate primers 5′-gggagggagtgtcaggtgag-3′ (mouse nt 1217 to 1192, lowercase letters indicate an added BamHI site) and 5′-GGG-CNACTCTTCAgGAGGCGAC-3′ (mouse nt 479 to 455, RevBmp2XhoI, lowercase letters indicate an added XhoI site) as described above. This PCR product also was digested with BamHI and XhoI and cloned into the BamHI and XhoI sites of pGL2Basic (Promega) cut with NheI and XhoI.

Luciferase Reporter Constructs
All nucleotide positions are provided with respect to the murine distal Bmp2 promoter that is 2,201 nt upstream of the initiator codon (ATG).

Mouse Plasmids
Bmp-LUC (nt 1237 to 471, pGL1.7XX)—The plasmid was constructed with XhoI and XbaI and subcloned into the luciferase reporter vector, pGL2Basic (Promega) cut with NheI and XhoI.

pGL2Basic-Bam—To remove the BamHI site from the 3′ cloning site, pGL2Basic (Promega Corp.) was digested with BamHI and the ends were made blunt using Klenow (New England Biolabs) in the presence of dNTPs and religated.

B716-LUC (nt 245 to 471, pGL716BXX)—pGL1.7XX was linearized with HindIII and partially digested with BamHI. A 724-nt fragment (716 nt of Bmp2) was ligated into pGL2Basic-Bam cut with BglII and HindIII.

BmpNot-LUC (nt 1237 to 193 and 160 to 471, pGL1.7XX Not)—The plasmid was constructed with XhoI and XbaI and subcloned into the luciferase reporter vector, pGL2Basic (Promega) cut with NheI and XhoI.

pGL2Basic-Bam—To remove the BamHI site from the 3′ cloning site, pGL2Basic (Promega Corp.) was digested with BamHI and the ends were made blunt using Klenow (New England Biolabs) in the presence of dNTPs and religated.

B716-LUC (nt 245 to 471, pGL716BXX)—pGL1.7XX was linearized with HindIII and partially digested with BamHI. A 724-nt fragment (716 nt of Bmp2) was ligated into pGL2Basic-Bam cut with BglII and HindIII.

2 curttools.curagen.com.
3 nbio.ncesu.edu/BioEdit/bioedit.html.
4 fruitfly.org seq_tools/proteor.
5 bimas.dcc.toh.gov/molbio/signal.
6 bio.cse.edu/pipmaker.
7 www-gsd.lbl.gov/vista.
Bmp2 Gene Regulation

pGL716BxXNA(nor)—To remove the 32-nt fragment between nt −192 and −181, plasmids pGL1.7XX and pGL716BxX were cut with NcoI and the ends religated.

B575-LUC (nt −104 to 471, pGL575pBX)—pGL716BxX was cut with BamHI and NheI. The ends were filled in with Klenow and religated.

BmpMPP2-F-LUC (pGL1.6Xdm1)—QuickChange™ site-directed mutagenesis (Stratagene) was used to make nucleotide substitutions in pGL1.7XX. Numbers are relative to the distal Bmp2 start site. At nt −1164 to −1163, TT was converted to CG and at nt −1157, C was converted to T. The sequences of the mutating oligomers were: nt −1180 to −1138, 5′-CGAACCTAGGTTATCCGAGCCTCGAGCT- GAAAGCTAGAC-3′ and nt −1138 to −1180, 5′-GTGCTAG- GTTCTTGCAAGGAGCTGAATCTGAGGTC-3′.

Primate Plasmids

AGM-LUC (Analogous to Mouse nt −104 to 465, pGLAGM599)—pBS-AGM599 was digested with BamHI, made blunt with Klenow, digested with XhoI, and cloned into pGL2BasicBam digested with SmaI and XhoI.

SqM-LUC (Analogous to Mouse nt −104 to 465, pGLSQM600)—A 1.6 kb fragment was PCR amplified from pCRHASmQ2.3KB with the primers described above. The PCR product was digested with BamHI, made blunt with Klenow, digested with XhoI, and then cloned into pGL2BasicBam digested with SmaI and XhoI.

F9 Cell Culture and Differentiation

F9 embryonal carcinoma cells were plated on dishes precoated with 1% gelatin and incubated at 37 °C with 10% CO2. The culture media consisted of Dulbecco’s modified Eagle’s medium supplemented with 10% heat-inactivated calf serum and 2 mM glutamine (Dulbecco’s modified Eagle’s medium). The cells were induced to differentiate into parietal endoderm by adding 1 μM all-trans-retinoic acid, 250 μM dibutyryl cAMP, and 500 μM theophylline (RACT). Undifferentiated control cells were treated with 250 μM dibutyryl cAMP and 500 μM theophylline (CT).

F9 Cell Transfection by Calcium Phosphate Precipitation

Transfections were performed essentially as described by Vasis et al. (46). Briefly, for 96-h drug treatments, F9 cells were plated at 1×105 or 0.5×105 (CT only) cells per 100-cm dish (Nunc) for 12 h, drugged for 48 h with CT or RACT, transfected by overnight calcium phosphate precipitation, and then cultured for an additional 48 h with drugs. Each 100-cm dish was cotransfected with 100 μg of reporter plasmid and 3 μg of pBACaZ (46) containing the β-galactosidase coding region driven by the constitutive β-actin promoter.

Luciferase Assays

Cells were extracted and luciferase activity was determined using the Promega Luciferase Assay system and a Monolight 2010 luminometer (Promega Corp.). 

Electrophoretic Mobility Shift Assays

Electrophoretic mobility shift assays were performed essentially as described in Gloczek et al. (47) except poly(dA-dT) was used to block nonspecific binding. Expression vectors encoding human RARβ and murine RXRγ (obtained from R. Evans) were linearized with BamHI and in vitro transcribed and translated using the TNT T7-coupled wheat germ protein system (Promega Corp.) according to the manufacturer’s instructions. Dr. F. Kashanchi provided human recombinant Sp1 from Promega Corp.

Chromatin Immunoprecipitation Assay (ChIP)

ChIP conditions were adapted from Kang et al. (48). F9 cells (−1×105) were fixed with 1% formaldehyde for 5 min at room temperature. 125 mM glycine was added to block cross-linking. Then, the cells were rinsed with phosphate-buffered saline, lysed (50 mM HEPES-KOH, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 140 mM NaCl, 10% glycerol, 0.5% Nonidet P-40, 0.25% Triton X-100, 1 mM phenylmethylsulfonyl fluoride, and 5 μg/ml each of leupeptin and aprotinin), and incubated for 10 min at 4 °C. Lysates were centrifuged, resuspended in wash buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 200 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, and 5 μg/ml each of leupeptin and aprotinin), and incubated for 10 min at 4 °C. The isolated nuclei were centrifuged and resuspended in 2 ml of RIPA buffer (10 mM Tris-HCl, pH 8.0, 1% Triton X-100, 0.1% SDS, 0.1% sodium deoxycholate, 1 mM EDTA, 0.5 mM EGTA, 140 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, and 5 μg/ml each of leupeptin, peptatin A, and aprotinin). 0.5 ml of suspended nuclei were sonicated to an average DNA length of 500–1200 base pairs using an ultrasonic processor with a microtip (8 × 10 s). 400 μg of sonicated DNA in 50 μl of RIPA buffer was centrifuged and the supernatant was incubated with 100 μl of protein A slurry (50% protein A-Sepharose in 15 mM Tris, pH 7.5, 1 mM EDTA, 150 mM NaCl, 0.5% Triton X-100, 1 mg/ml bovine serum albumin, and 400 μg/ml sonicated salmon sperm DNA) for 30 min. The precleared supernatant was immunoprecipitated with 2–4 μg of specific antibodies with rotation overnight at 4 °C. All antibodies were from Santa Cruz Biotechnology, Inc.: Sp1 (PEP2) SC-59, RARα (C-20) SC-551, RARβ (C-19) SC-552, RXRα (D-20) SC-553, and RXRγ (Y-20) SC-555. Immunocomplexes adsorbed onto 100 μl of protein A beads were washed sequentially with 1 ml of washing buffer 1 (20 mM Tris, pH 7.5, 2 mM EDTA, 500 mM NaCl, 0.5% Triton X-100), washing buffer 2 (20 mM Tris, pH 7.5, 2 mM EDTA, 500 mM NaCl, 0.5% Triton X-100), washing buffer 3 (20 mM Tris, pH 7.5, 1 mM EDTA, 250 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate), and washing buffer 4 (10 mM Tris, pH 7.5, 1 mM EDTA). Antibody-bound chromatin fragments were eluted from the beads with 500 μl of 1% SDS in 0.1 M NaHCO3. Formaldehyde-induced cross-links were reverted by heating at 65 °C overnight. DNA then was extracted using 80 μg/ml proteinase K digestion, phenol/chloroform extraction, and ethanol precipitation. The DNA was resuspended in 50 μl of sterile H2O and 5 μl was used as a template for PCR. After separation on 1% agarose gels, the PCR products were subjected to Southern blot analysis as described above.

RESULTS

Noncoding Sequences Upstream of the Bmp2 Gene Are Widely Conserved in Mammals and Chick—The Bmp2 coding sequences are highly conserved in distantly related animal phyla (3, 22). Bmp2 expression also is conserved in several vertebrate tissues during early development. For example, Bmp2 transcripts occur in the posterior regions of developing mouse and chick limb buds and zebrafish fin buds (4, 49). Like others, we observed that the regions upstream of the Bmp2 gene are highly conserved between humans and mice (32). Similarly, the untranslated regions (UTRs) of Bmp2 transcripts are conserved between many species. This suggests that critical regulatory elements mediating the expression of Bmp2 have been retained throughout evolution.

To test the hypothesis that Bmp2 regulatory elements are conserved, we developed a PCR strategy to amplify Bmp2 noncoding regions from other species. We identified conserved regions 1.2 kb upstream of the distal promoter (30) and at the 3′ end of exon 1. A degenerate upstream PCR primer was designed using a ClustalW alignment of the mouse and human Bmp2 genomic sequences. The degenerate exon primer was designed by comparing the human, mouse, rat, deer, and rabbit Bmp2 cDNA sequences. The locations of these primers and sequence accession numbers are shown in Fig. 1. Each primer sequence was compared with GenBank™ using BLAST™ and no significant similarity to other sequences was found.

The predicted 2.3-kb Bmp2 fragment was PCR amplified from murine (M. musculus, strain 129) genomic DNA using these degenerate primers. Similarly sized products (±200 nt) were also amplified from genomic DNA isolated from rat (R. norvegicus), hamster (C. griseus), human, squirrel monkey (Saimiri sciureus), African green monkey (C. aethiops), mar- moset (C. jacchus), dog (C. familiaris), and wallaby (M. eugenii) using these primers (not shown). We cloned and sequenced the amplified fragment from squirrel monkey (GenBank™ accession number MY491488). Subsequently, we designed degenerate internal primers flanking the distal promoter (nt −140 to 619, Fig. 1A). These primers amplified fragments from owl monkey (A. trivirgatus), cow (B. taurus), and chicken (G. gallus), as well as the species indicated previously (Fig. 1B).

*www.ncbi.nlm.nih.gov/BLAST.*
The ability of the primers to amplify fragments indicates that the primer sequence occurs in each genome. To confirm that the amplified fragments were the Bmp2 sequence, the PCR products were hybridized to a 1.7-kb Bmp2 fragment (-1237 to +471 nt relative to the distal promoter). This probe hybridized to the amplified fragments under normal stringency conditions (6× SSC, 50% formamide, 42 °C, Fig. 1B). These results demonstrated that the PCR products were indeed Bmp2 fragments.

**Sequences Downstream of the Bmp2 Translated Region Are Conserved between Mammals, Birds, Amphibians, and Fish**—

The genomes of several vertebrates have been completed or are nearing completion. Consequently, we could compare sequences within and flanking the Bmp2 gene of several different species that could not be analyzed easily using PCR. We used two different programs (PipMaker and VISTA) that were designed to provide global sequence alignments between extremely long sequences. Each program provides a useful visualization of the many possible alignments.

PipMaker and MultiPipMaker compare the percent identity of all alignments between gaps. The results are returned as a Percent Identity Plot (PIP) showing the percent identity of any gap-free segments within an alignment (50, 51). Unbroken lines indicate the alignment length. The height of the line indicates percent identity over that alignment. For example, the mature Bmp2 peptide sequences in mice and humans are nearly identical and only one-third base differences occur in the DNA sequence. Consequently, an unbroken line near 90% results from comparing the mouse and human coding sequences within exon 3 (Fig. 2A).

VISTA measures percent identity within a user-specified window beginning at each base pair of the entire sequence (52). The results are returned as a continuous curve representing identity level. For example, aligning the exon 3 coding regions from mouse and human produces a high plateau in VISTA (Fig. 2B). Whereas PipMaker data is anchored by comparing all sequences to one species, VISTA performs pairwise sequence alignments between 3 species (e.g. mouse-human, mouse-pufferfish, and human-pufferfish (53)). A statistical analysis of transcription factor binding sites identified by comparing bacterial genomes suggests that three species are sufficient for meaningful comparisons (54). Each program enables annotation of the results with functional information such as coding regions and CpG islands and is publicly available.

We have assumed that sequences identified in three or more species with both programs are likely to be functionally significant. Using these tools and this criterion, we identified non-coding sequences that are conserved between rodents, humans, deer, birds, frogs, and fish. Fig. 2, A and B, shows MultiPipMaker and VISTA analyses of the Bmp2 genes from mouse (M. musculus, strain C57Bl/6), human, and pufferfish (Fugu rubripes). Conservation upstream and surrounding the distal promoter region is readily apparent between the mouse and human sequences. Rodents and primates diverged ~91 million years ago (55).

An upstream conserved region corresponds to a high identity PIP bar and a VISTA peak (Fig. 2, A and B). This region includes the degenerate PCR primer site used to amplify Bmp2 sequences from diverse species (Fig. 1). Together, the computer and experimental analyses indicate that the sequence is conserved between 11 mammals and 1 bird. This evolutionary conservation suggests that elements controlling mammalian and perhaps avian Bmp2 expression reside in this region.

Both the distal (29–31, 56) and proximal (29, 57) promoter regions are conserved between rodents and primates. The distal promoter region exhibits slightly higher identity (compare the length and height of the PIP bars near each promoter). For example, seven mismatches occur in the 100 nt leading up to the mouse and human distal promoters, whereas the proximal promoters have 19 mismatches.

Conservation in the 5′ end of the gene was not evident in the amphibian or teleost sequences (the short overlapping PIP bars and VISTA peak upstream of the promoter are because of a simple sequence repeat found in mammals and fish). Thus, the sequences near the promoters may interact with mammalian and possibly avian-specific factors. Pufferfish diverged ~450 million years ago (55). Although mainly coding sequence conservation occurred over this period, functional conserved non-coding sequences have been identified. For example, many enhancers required for mammalian Hox gene expression, including retinoic acid response elements, are conserved between mammals and fish (Refs. 43 and 58, and references therein).

Both MultiPipMaker and VISTA analyses detected a highly conserved region at the beginning of the 3′-UTR of the Bmp2 transcript from mouse, human, and pufferfish (Fig. 2, A and B). This region is also conserved in rat, rabbit (Oryctolagus cuniculus), deer (Dama dama), chick, frogs (Xenopus laevis and X. tropicalis), and zebrafish (Bmp2b, Danio rerio). The Bmp2 conserved element is absent in the pufferfish Bmp4 gene (Fig. 2C), thus implying a function specifically in Bmp2 regulation. Conservation over a large phylogenetic distance suggests the sequence must interact with conserved cellular factors in a...
FIG. 2. MultiPipMaker and VISTA alignments of conserved sequences in vertebrate Bmp2 genes. A diagram of the Bmp2 gene structure is above each plot. The distal (D) and proximal (P) promoters at nt 1 and 735, respectively, are indicated by arrows (right square arrow). Untranslated regions in exons 1 and 3 are gray boxes. Translated regions in exons 2 and 3 are black. The upstream (upstr.) Conserved Noncoding Sequence (CNS), the CNS surrounding the distal promoter, and the 3'-UTR CNS are boxed. A, a PIP comparing all sequences for which Bmp2 3'-UTRs are available. 50–100% identity is shown. Multiple bars in one location indicate the many possible alignments between simple sequence repeats occurring in the paired sequences. All sequences are genomic except for rabbit, deer, and chick. Species and sequence accession numbers are: mouse, M. musculus strain C57Bl/6, NW_000178; rat, R. norvegicus, AABR02024700; rabbit, O. cuniculus, AF041421; human, H. sapiens, AF040249; deer, Dama dama, AF041421; chick, G. gallus, compiled from X75914 and BU423990; X. laevis, X. laevis, AJ315159; X. tropicalis,
Mouse and Monkey Bmp2 Sequences Drive Reporter Gene Expression Similarly in F9 Cells Treated with RA or RACT—

Interspecies sequence comparisons indicate that many sequences outside of the Bmp2 coding regions are conserved. These regions may contain regulatory elements subject to functional constraints. We have begun to experimentally test whether or not these conserved noncoding sequences are functionally relevant using differentiating F9 cells. Untreated or CT-treated F9 cells are undifferentiated and closely resemble cells of the blastocyst inner cell mass. RA or RACT treatment stimulates F9 cells to differentiate into primitive or parietal extraembryonic endoderm, respectively (41). Whereas undifferentiated cells express no detectable Bmp2, RA induces expression. Bmp2 expression is further elevated by RACT treatment (Fig. 3A) (35). We previously demonstrated that a 1.7-kb fragment containing the mouse Bmp2 distal promoter drove RA-dependent chloramphenicol acetyltransferase reporter gene expression in F9 cells (30). The 1.7-kb fragment contains 1237 nt of upstream sequence, the distal promoter, 471 nt of exon 1, and includes the conserved regions described above. As shown in Fig. 3A, like the endogenous Bmp2 transcript, luciferase gene expression driven by this region is induced in RA-treated cells and is elevated further in RACT-treated cells.

To test the hypothesis that the regulatory elements required to express Bmp2 during extraembryonic endoderm differentiation are conserved between mouse and primates, we cloned amplified Bmp2 DNA from squirrel monkey (S. sciureus) and African green monkey (C. aethiops). 2322 nt of squirrel monkey and 1046 nt of African green monkey sequences have been deposited into GenBank™ (accession numbers AY494188, AY494189, and AY494190). As shown in Fig. 3B, RACT induced transcription from monkey sequence-containing reporter genes by about 10-fold relative to undifferentiated cells (CT). The level of transactivation resembled that mediated by the mouse sequence. This indicates that the primate sequence contains a promoter and interacts appropriately with mouse regulatory factors during differentiation.

Deletion and Mutation Analysis of Conserved Bmp2 Upstream Sequences—To further delineate the elements involved in regulating Bmp2, various sequences were mutated or deleted from the full-length murine fragment (nt −1237 to +471). We tested each construct for its ability to drive reporter expression relative to the full-length fragment in F9 cells digested for 96 h.

MultiPipMaker and VISTA analyses detected a conserved sequence between −1237 and −1145 nt (Fig. 2). We also showed that primers designed from human and mouse sequences within this conserved sequence could PCR amplify fragments from many non-rodent and non-primate species (Fig. 1). This evolutionarily conservation suggests the region interacts with specific cellular factors. A perfect binding site (ATTTTA) for the P2F transcriptional repressor is located between nt −1146 to −1141 of the mouse sequence and at the homologous region of the rat, human, African green monkey, and squirrel monkey genes (Fig. 4A). P2F represses prolactin expression in some contexts (59). Mutating this site elevated reporter gene expression by 30% (Fig. 4C, p < 0.05). This suggests that P2F may function in F9 cells.

Deletion of the region between −1237 and −245 did not tightly constrained manner. A role in modulating Bmp2 gene expression is possible.

Mouse and Monkey Bmp2 Sequences Drive Reporter Gene Expression Similarly in F9 Cells Treated with RA or RACT—

Interspecies sequence comparisons indicate that many sequences outside of the Bmp2 coding regions are conserved. These regions may contain regulatory elements subject to functional constraints. We have begun to experimentally test whether or not these conserved noncoding sequences are functionally relevant using differentiating F9 cells. Untreated or CT-treated F9 cells are undifferentiated and closely resemble cells of the blastocyst inner cell mass. RA or RACT treatment stimulates F9 cells to differentiate into primitive or parietal extraembryonic endoderm, respectively (41). Whereas undifferentiated cells express no detectable Bmp2, RA induces expression. Bmp2 expression is further elevated by RACT treatment (Fig. 3A) (35). We previously demonstrated that a 1.7-kb fragment containing the mouse Bmp2 distal promoter drove RA-dependent chloramphenicol acetyltransferase reporter gene expression in F9 cells (30). The 1.7-kb fragment contains 1237 nt of upstream sequence, the distal promoter, 471 nt of exon 1, and includes the conserved regions described above. As shown in Fig. 3A, like the endogenous Bmp2 transcript, luciferase gene expression driven by this region is induced in RA-treated cells and is elevated further in RACT-treated cells.

To test the hypothesis that the regulatory elements required to express Bmp2 during extraembryonic endoderm differentiation are conserved between mouse and primates, we cloned amplified Bmp2 DNA from squirrel monkey (S. sciureus) and African green monkey (C. aethiops). 2322 nt of squirrel monkey and 1046 nt of African green monkey sequences have been deposited into GenBank™ (accession numbers AY494188, AY494189, and AY494190). As shown in Fig. 3B, RACT induced transcription from monkey sequence-containing reporter genes by about 10-fold relative to undifferentiated cells (CT). The level of transactivation resembled that mediated by the mouse sequence. This indicates that the primate sequence contains a promoter and interacts appropriately with mouse regulatory factors during differentiation.

Deletion and Mutation Analysis of Conserved Bmp2 Upstream Sequences—To further delineate the elements involved in regulating Bmp2, various sequences were mutated or deleted from the full-length murine fragment (nt −1237 to +471). We tested each construct for its ability to drive reporter expression relative to the full-length fragment in F9 cells digested for 96 h.

MultiPipMaker and VISTA analyses detected a conserved sequence between −1237 and −1145 nt (Fig. 2). We also showed that primers designed from human and mouse sequences within this conserved sequence could PCR amplify fragments from many non-rodent and non-primate species (Fig. 1). This evolutionarily conservation suggests the region interacts with specific cellular factors. A perfect binding site (ATTTTA) for the P2F transcriptional repressor is located between nt −1146 to −1141 of the mouse sequence and at the homologous region of the rat, human, African green monkey, and squirrel monkey genes (Fig. 4A). P2F represses prolactin expression in some contexts (59). Mutating this site elevated reporter gene expression by 30% (Fig. 4C, p < 0.05). This suggests that P2F may function in F9 cells.

Deletion of the region between −1237 and −245 did not tightly constrained manner. A role in modulating Bmp2 gene expression is possible.

Mouse and Monkey Bmp2 Sequences Drive Reporter Gene Expression Similarly in F9 Cells Treated with RA or RACT—

Interspecies sequence comparisons indicate that many sequences outside of the Bmp2 coding regions are conserved. These regions may contain regulatory elements subject to functional constraints. We have begun to experimentally test whether or not these conserved noncoding sequences are functionally relevant using differentiating F9 cells. Untreated or CT-treated F9 cells are undifferentiated and closely resemble cells of the blastocyst inner cell mass. RA or RACT treatment stimulates F9 cells to differentiate into primitive or parietal extraembryonic endoderm, respectively (41). Whereas undifferentiated cells express no detectable Bmp2, RA induces expression. Bmp2 expression is further elevated by RACT treatment (Fig. 3A) (35). We previously demonstrated that a 1.7-kb fragment containing the mouse Bmp2 distal promoter drove RA-dependent chloramphenicol acetyltransferase reporter gene expression in F9 cells (30). The 1.7-kb fragment contains 1237 nt of upstream sequence, the distal promoter, 471 nt of exon 1, and includes the conserved regions described above. As shown in Fig. 3A, like the endogenous Bmp2 transcript, luciferase gene expression driven by this region is induced in RA-treated cells and is elevated further in RACT-treated cells.

To test the hypothesis that the regulatory elements required to express Bmp2 during extraembryonic endoderm differentiation are conserved between mouse and primates, we cloned amplified Bmp2 DNA from squirrel monkey (S. sciureus) and African green monkey (C. aethiops). 2322 nt of squirrel monkey and 1046 nt of African green monkey sequences have been deposited into GenBank™ (accession numbers AY494188, AY494189, and AY494190). As shown in Fig. 3B, RACT induced transcription from monkey sequence-containing reporter genes by about 10-fold relative to undifferentiated cells (CT). The level of transactivation resembled that mediated by the mouse sequence. This indicates that the primate sequence contains a promoter and interacts appropriately with mouse regulatory factors during differentiation.

Deletion and Mutation Analysis of Conserved Bmp2 Upstream Sequences—To further delineate the elements involved in regulating Bmp2, various sequences were mutated or deleted from the full-length murine fragment (nt −1237 to +471). We tested each construct for its ability to drive reporter expression relative to the full-length fragment in F9 cells digested for 96 h.

MultiPipMaker and VISTA analyses detected a conserved sequence between −1237 and −1145 nt (Fig. 2). We also showed that primers designed from human and mouse sequences within this conserved sequence could PCR amplify fragments from many non-rodent and non-primate species (Fig. 1). This evolutionarily conservation suggests the region interacts with specific cellular factors. A perfect binding site (ATTTTA) for the P2F transcriptional repressor is located between nt −1146 to −1141 of the mouse sequence and at the homologous region of the rat, human, African green monkey, and squirrel monkey genes (Fig. 4A). P2F represses prolactin expression in some contexts (59). Mutating this site elevated reporter gene expression by 30% (Fig. 4C, p < 0.05). This suggests that P2F may function in F9 cells.

Deletion of the region between −1237 and −245 did not tightly constrained manner. A role in modulating Bmp2 gene expression is possible.
significantly affect expression in RACT-treated cells (Fig. 4C). However, removing the −245 to −104 segment (Fig. 4, B and C) reduced expression by 30% (\(p < 0.001\)). We also deleted a NotI fragment between −193 to −160 from the plasmids containing 1237 or 245 nt of sequence upstream of the promoter. Luciferase activities generated from the plasmids with the small internal deletion were indistinguishable from that of the plasmid controlled by only 104 nt of upstream sequence. The require-

---

**Fig. 4. Conserved noncoding sequences upstream of the distal promoter.** A, the mouse upstream sequence between −1191 and −1144 is aligned with rat, human, squirrel monkey, and African green monkey sequences. The conserved F2F site consensus sequence (TAAAAT) in the reverse orientation is in **bold face** and **shaded**. The F2F site nucleotides mutated in the reporter construct BmpMtF2F-LUC are indicated in lowercase below the sequence. Species and sequence accession numbers are: mouse, *M. musculus* strain C57Bl/6, NW_000178; rat, *R. norvegicus*, AABR02024700; human, *H. sapiens*, AF040249; squirrel monkey (Sq Monk), *S. sciureus*, (accession number AY494188); African green (Af Gr) monkey, *C. aethiops*, (accession numbers AY494189 and AY494190). Independently sequenced mouse strain 129 sequence compiled from L25602 and AF074942 was identical to the C57Bl/6 sequence. B, the BamHI fragment containing the GC-rich region with Sp1 binding sites (nt −54 to −98) is aligned with rat, human, and squirrel monkey sequences. Classical Sp1 sites (GGGCGG) in the reverse orientation are **shaded**. The two NotI restriction endonuclease sites are **bold face** and **underlined**. Species and sequence accession numbers are as in A. C, F9 cells were transfected, drug-treated, and analyzed for luciferase activity as described for Fig. 3 (\(n = 3–21\)). The diagram to the left of the bar graph illustrates the DNAs driving the reporter gene activity. The ends of the constructs are indicated below each bar. The black box and ∧ mark the deleted 32-base pair NotI fragment. A Student’s t test (two-sample assuming unequal variances) was used to compare the activity of each reporter construct to the full-length murine reporter (Bmp-LUC). The following constructs differed significantly: the full-length construct with the mutated F2F site, BmpMtF2F-LUC, \(p < 0.05\) (*), \(n = 5\); the −104 to 471 construct, B575-LUC, \(p < 0.001\) (**), \(n = 17\); the full-length construct lacking the NotI fragment between −193 and −160, BmpNot-LUC, \(p = 0.01\) (**), \(n = 3\); the −245 to 471 construct lacking the NotI fragment, B716Not-LUC, \(p < 0.01\) (**), \(n = 3\).
ment for this NotI fragment was also observed previously using a chloramphenicol transferase reporter gene (30). Thus, an element within or overlapping a 32-nt fragment upstream of the distal promoter accounts for one-third of the transactivation controlling the Bmp2 promoter.

Deleting the NotI fragment removes a consensus Sp1 site that binds Sp1 in vitro. Cold competitor oligomers with the same sequence or an unrelated sequence containing a consensus Sp1 binding site reduced binding (Fig. 5A). Oligomers containing the site within the NotI fragment bind to 500 pg of Sp1 protein and 1 μl of a 1:1 mixture of RARβ and RXRγ, and electrophoresed through nondenaturing gels. 400-fold molar excess of unlabeled Bmp2 or Sp1 consensus oligomers (ATTGATCGGGGCGGGCGGCCGC) were added as indicated. B, labeled Bmp2 oligomers were incubated with the indicated mass of Sp1 protein with 2 μl of a 1:1 mixture of RARβ and RXRγ (+) or with an equal volume of unprogrammed wheat germ extract (−). C, labeled Bmp2 oligomers were incubated with 500 pg of Sp1 with the indicated volumes of RARβ and RXRγ mixture. The volume and total protein concentration of each binding reaction was equalized with unprogrammed wheat germ extract.

Fig. 5. RAR and RXR potentiate Sp1 binding to Bmp2 sequences. A, Sp1 binds to annealed oligomers containing the Sp1 site within the NotI fragment (nt −193 to −143, Fig. 4B). Oligomers containing only the Sp1 site embedded in the downstream NotI site (nt −168 to −143) or the two sites downstream of the NotI fragment (nt −148 to −98) did not bind (data not shown). Radiolabeled oligomers were incubated with 500 pg of recombinant human Sp1 protein and 1 μl each of in vitro transcribed and translated RARβ and RXRγ, and electrophoresed through nondenaturing gels. 400-fold molar excess of unlabeled Bmp2 or Sp1 consensus oligomers (ATTGATCGGGGCGGGCGGCCGC) were added as indicated. B, labeled Bmp2 oligomers were incubated with the indicated mass of Sp1 protein with 2 μl of a 1:1 mixture of RARβ and RXRγ (+) or with an equal volume of unprogrammed wheat germ extract (−). C, labeled Bmp2 oligomers were incubated with 500 pg of Sp1 with the indicated volumes of RARβ and RXRγ mixture. The volume and total protein concentration of each binding reaction was equalized with unprogrammed wheat germ extract.

Sp1 and RARβ Interact Upstream of Bmp2—Electrophoretic mobility shift assays (Fig. 5) (30) showed that annealed, labeled oligomers containing the site within the NotI fragment bind to Sp1 protein in vitro. Cold competitor oligomers with the same sequence or an unrelated sequence containing a consensus Sp1 binding site reduced binding (Fig. 5A). Oligomers containing the three downstream Sp1 sites did not bind in vitro (data not shown). The sequence context surrounding Sp1 sites can influence binding. For example, five or more purines or pyrimidines on the 3′ side of an Sp1 consensus site can reduce its affinity for Sp1 relative to sites with mixed purines and pyrimidines (62). The 3′ upstream (relative to the Bmp2 promoter) side of the site that binds Sp1 in vitro contains more mixed purine and pyrimidine bases compared with the three non-binding sites. Thus, the context of the Sp1 site within the NotI fragment appears to promote Sp1 binding.

Fig. 6. RARβ and Sp1 act upstream of the distal promoter in CT-treated cells. A, PCR was used to amplify the Sp1 binding region upstream of the distal promoter (nt −286 to −24) from genomic DNA or DNA isolated from chromatin that was immunoprecipitated with Sp1 or RARβ antibodies (ChIP). The approximate locations of the forward (thick right arrow, −286 to −258) and reverse (thick left arrow, −24 to −44) PCR primers are indicated. B, chromatin was prepared from F9 cells treated with RA or dibutyryl cAMP and theophylline (CT) or both (RACT) for 3 days. DNA was extracted directly (Input) or from chromatin immunoprecipitated (IP) with Sp1 (α-Sp1) or RARβ (α-RARβ) antibodies. The PCR products were electrophoresed, Southern blotted, and hybridized to Bmp2-specific radiolabeled probes. The numbers below the ChIP blots indicate the relative signals in each lane after normalization to the signals in the Input lanes. As negative controls, PCR targeted to non-promoter regions such as the Bmp2 3′-UTR were performed on ChIP DNA. No signals were observed. C, F9 cells were CT-treated and transfected with the full-length murine Bmp2 reporter construct (nt −1237 to 471, Bmp-LUC) as described for Fig. 3 except that cells were exposed to mithramycin (MTM) at the time of transfection. The average luciferase activity ± S.E. (n = 2 to 4) was plotted against the MTM concentration.
Despite strong activation of the endogenous Bmp2 gene and Bmp2 promoter-driven reporter genes by RA in F9 cells, a classical RARE cannot be identified near the Bmp2 promoter. Classical RAREs contain two directly repeated "PuG(G/T)TCA" motifs separated by 5 nucleotides (63). Although the spacing and relative position of these motifs may vary, at least two usually occur (Ref. 47 and references therein). Only three, distantly spaced motifs occur within the full-length Bmp2 construct (nt /H11002 1236 to /H11001 471). None of these are between /H11002 245 to /H11002 104. The regulatory regions controlling several RA-inducible genes lacking classical RAREs have been shown to contain Sp1-binding sites whose activity is modulated by retinoid receptors (see Refs. 62 and 64, and references therein). Thus retinoid-activated receptors may control Bmp2 transcription by modulating Sp1 activity at this element.

We tested the hypothesis that retinoid receptors potentiate Sp1 binding at low Sp1 levels by mixing purified Sp1 protein with in vitro transcribed and translated RARα and RXRα. As shown in Fig. 5B, adding RARα and RXRα enhanced Sp1 binding to the oligomers containing the Sp1 site within the NotI fragment at low levels of Sp1. In contrast, we never observed binding to oligomers containing the three downstream Sp1 sites (not shown). The Sp1 to receptor ratio appeared to be critical because adding excess receptor protein reduced Sp1 binding (Fig. 5C).

Electrophoretic mobility shift assay can test whether or not a
protein can bind a sequence in vitro. However, cellular and molecular context greatly modifies protein activity. For example, depending on post-translational modification or concentration of Sp1, or on binding site context, Sp1 may activate or repress genes (65–70). Similarly, coactivator and corepressor levels alter nuclear receptor activity (71–73). Reporter genes permit mutational analysis of putative elements, but multiple copies of transiently transfected genes may disrupt the normal balance of regulatory factors. Therefore, we used ChIP to test whether or not Sp1 binds the conserved Sp1 sites in non-transfected cells.

We isolated chromatin from undifferentiated, CT-treated cells that express no Bmp2 mRNA, RA-treated cells that express some Bmp2 mRNA, and RACT-treated cells that express the highest level of Bmp2 mRNA (Fig. 3) (35, 42). This chromatin was immunoprecipitated with antibodies specific to Sp1, RARα and -β, and RXRβ and -γ. After extracting DNA, we attempted to amplify the region containing the conserved Sp1 sites (Fig. 6A). As Fig. 6B shows, a fragment was readily amplified from the CT-treated cell chromatin immunoprecipitated with anti-Sp1 and anti-RARβ antibodies. A faint band was also visible in the DNA from RA-treated cells. We could not amplify fragments from DNA from RACT-treated cells nor from chromatin precipitated with other antibodies (data not shown).

Likewise, we were unable to detect Sp1 or RARβ bound to other regions of the Bmp2 gene such as the 3'-UTR (data not shown). We normalized the intensity of the ChIP signals to the amount of total DNA (input) in each chromatin sample. The relative signal levels indicate that Sp1 and RARβ binding correlates inversely with the abundance of the Bmp2 transcript.

A pharmacological approach also can assess the activity of a regulatory protein binding the Bmp2 promoter. Mithramycin interferes with Sp1-mediated repression. Since mithramycin interferes with Sp1-mediated repression.

Conserved Noncoding Sequences in the Bmp2 3'-UTR—The sequence comparisons shown in Fig. 2 identified a sequence downstream of the Bmp2 stop codon that is conserved between mammals, birds, frogs, and fish. As Fig. 6 shows the MultiPipMaker sequence alignment of the highly conserved region. The 96 nt sequence overlaps the region shown in Fig. 8 is 96% identical between all mammals and is 93% identical between mouse and chick. With the exception of a gap between 9,712 and 9,728, even the two Xenopus sequences are 84% identical to the mouse sequence. Remarkably, large stretches within this region are identical between mouse and the two fish sequences (shaded gray). Indeed, identity over the 230 nt region conserved between mouse and pufferfish is about 69%.

A Reporter Gene Containing the Bmp2 Downstream Region Is Induced More Strongly by RA—To test the hypothesis that the region conserved between mammals and other vertebrates influences Bmp2 expression, we inserted the entire 3'-UTR downstream of luciferase in the 1.7-kb mouse reporter construct previously described (Bmp-LUC, Figs. 3 and 4). Fig. 8A shows the construct (Bmp-LUC-Bmp). The inserted fragment includes the Bmp2 stop codon, the conserved noncoding sequence, and four putative polyadenylation sites (57, 76, 77). The actual luciferase activity generated in RACT-treated cells from the reporter construct with the 3'-UTR was more than twice that observed for the 3'-UTR lacking construct (Fig. 8B). Furthermore, luciferase activity was not significantly elevated in CT-treated cells transfected with the 3'-UTR construct. Consequently, whereas RA induced Bmp-LUC expression by 9.5 ± 2.3-fold relative to non-RA-treated cells, inserting the 3'-UTR enabled RA to induce expression by 13.6 ± 1.5-fold. These results indicate that key regulatory elements occur within this region.

The Bmp2 3'-UTR Increases the Stability of the Luciferase mRNA—The overall AU content of the conserved region in mouse is about 60% (Fig. 7). AU-rich sequences are associated with post-transcriptional regulation. A specific AUUUA motif known as an AU-rich element (ARE) controls transcript stability (78). Seven AREs are conserved in the mammalian 3'-UTRs. The conserved AREs are embedded exclusively within the re-
scription shown in Fig. 7 and none occur in the remaining 1.6 kb of sequence included in the reporter gene (Fig. 8). Of these, four also occur in fish.

The presence of AREs suggests that the strong 3′-UTR element may control the half-life of the Bmp2 mRNA. To analyze the relative stability of luciferase transcripts containing or lacking the Bmp2 3′-UTR, transfected and RACT-treated F9 cells were exposed to the transcriptional inhibitors 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole or actinomycin D. The concentrations used, 25 μg/ml 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole or 2 μg/ml actinomycin D, have been shown to inhibit [5,6-3H]uridine incorporation by 90–95% (79).

We plotted the data as a percentage of the luciferase activity in cells not treated with inhibitor (time 0). Within 2 h of 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole addition, luciferase levels in cells transfected with the “promoter only” plasmid (Bmp-LUC) fell to 10% of those observed at time 0 (Fig. 8C). In contrast, luciferase levels in cells transfected with the Bmp2 3′-UTR-containing construct (Bmp-LUC-Bmp) were nearly 80% of time 0 levels. Similarly, the Bmp-LUC-Bmp message was more stable in cells treated with actinomycin D (not shown). Thus the Bmp2 3′-UTR contains an element that stabilizes a heterologous reporter message.

**DISCUSSION**

**Upstream Conserved Noncoding Sequences—**BMP2 is an indispensible growth factor required throughout vertebrate and invertebrate embryogenesis. Disruptions in BMP2 signaling influence human diseases, including cancers and pulmonary hypertension (80). Indeed, the human BMP2 gene has been linked to osteoporosis (81). Thus, understanding the mechanisms controlling BMP2 synthesis is relevant to congenital disorders and adult disease. BMP2 is an ancient and conserved member of the transforming growth factor β family. Related proteins from animals as distant as flies and humans are functionally interchangeable in some assays. Gene expression patterns also appear partially conserved. For example, Bmp2 mRNA is detectable in the homologous regions of the developing mouse and chick limb buds and the zebrafish fin buds (4, 49). We have identified and characterized conserved genomic regulatory elements controlling the induction of the Bmp2 gene in F9 embryonal carcinoma cells during parietal endoderm differentiation.

Genomic sequences containing the entire Bmp2 transcription unit are available from mice, rats, humans, two frogs (X. laevis and X. tropicalis), zebrafish, and pufferfish. MultiPipeMaker and VISTA alignment of the mouse promoter region to the frog and fish sequences failed to detect significant conservation. In contrast, the rodent and primate sequences contain obvious regions of high identity (Figs. 2 and 4). Reporter gene activity (Fig. 3) indicates that upstream regulatory elements required for inducing Bmp2 in F9 cells treated by RACT are functionally conserved between rodents and primates whose lineages separated about 91 million years ago (55). Rodents and primates are evolutionary sister groups (82); however, our ability to amplify sequences from other mammals and chick suggests that Bmp2 regulatory element conservation extends considerably further (Fig. 1). Because only limited chick sequence is available, directly comparing the putative chick promoter to mouse is not possible. However, the fragment amplified from chick DNA hybridized easily to the mouse promoter region under normal stringency conditions. This suggests that, despite ~310 million years of independent evolution (55), many aspects of Bmp2 promoter region function have been retained.

**Sp1 Regulation of Bmp2 Transcription—**A cluster of Sp1 consensus binding sites occurs in a GC-rich region upstream of the mouse, rat, human, and squirrel monkey promoters (Fig. 4). In addition to sequence identity, four distinct experimental approaches support the hypothesis that one or more conserved Sp1 sites regulate Bmp2 expression. First, deleting the Sp1 site between nt −192 to −161 significantly reduced reporter gene activity (Fig. 4C). Second, this sequence bound Sp1 in electrophoretic mobility shift assays (Fig. 5). Third, ChIP analyses showed that Sp1 bound this region in F9 cells (Fig. 6B). Fourth, mithramycin, an inhibitor of Sp1/chromatin interactions altered Bmp2 reporter gene activity (Fig. 6C). Thus, Sp1 may influence Bmp2 transcription.

We also found that the RARβ/RXRα heterodimer potentiates Sp1 binding to this site in vitro (Fig. 5, B and C). Furthermore, ChIP analyses indicate that RARβ binds the same region as Sp1 in F9 cells (Fig. 6B). Retinoid receptor and Sp1 interaction may help explain the absence of a classical RARE near the Bmp2 promoter. Indeed, several RA-inducible genes, including the Bmp2 relative TGF-β1, lack classical RAREs, but contain Sp1-binding sites required for induction by RA (62, 64). Like Bmp2, RARα potentiates Sp1 binding to these sites. Moreover, both RA and mithramycin influence the activity of reporter genes derived from these genes (62).

Several lines of evidence point to a direct physical interaction between Sp1 and retinoid receptors. These include coimmunoprecipitation of purified proteins, the fact that glutathione S-transferase-chimeric proteins pull-down these factors from cell extracts, and mammalian two-hybrid assays (62, 83, 84).

Unexpectedly, however, ChIP analyses of proteins binding the Bmp2 upstream region in F9 cells indicated an inverse correlation between RARβ and Sp1 binding and Bmp2 expression. RARβ and Sp1 strongly bound the region in undifferentiated, CT-treated cells that do not express Bmp2; binding was intermediate in RA-treated cells expressing an intermediate level of mRNA; and binding was undetectable in RACT-treated cells expressing the highest mRNA level (compare Figs. 3A and 6B). Additionally, mithramycin, which interferes with Sp1 binding to chromatin (62, 74, 75), induced a Bmp reporter gene specifically in undifferentiated cells (Fig. 6C). These results suggest that a mithramycin-sensitive RARβ-Sp1 complex may repress Bmp2 transcription in undifferentiated, CT-treated cells.

Many nuclear hormone receptors repress transcription in the absence of ligand by recruiting transcriptional corepressors (71, 72). However, the simplest hypothesis that RARβ directly represses Bmp2 is unlikely. First, no classical retinoic acid response elements (63) occur within this region. In addition, in vitro synthesized receptors cannot bind this region alone (Fig. 5B and data not shown). Thus, a cofactor probably facilitates binding to this region. Second, although RARα interacts with the corepressor SMRT (silencing mediator of retinoic acid and thyroid hormone receptor), RARβ does not (85, 86). We only detected RARβ binding upstream of Bmp2 (Fig. 6B). Our results are mirrored by the observation that RARβ binds an Sp1 element upstream of the folate receptor (FR) type β gene only in untreated cells (87). Loss of binding occurs after RA induces the FRβ gene. For these reasons, we favor the hypotheses that Sp1 recruits RARβ to Bmp2 and that an RARβ-Sp1 complex represses Bmp2.

Consistent with this hypothesis, RARβ mRNA abundance is greatly reduced in RACT-treated cells compared with CT-treated cells (35, 88, 89). In contrast, Sp1 is a ubiquitously expressed regulatory protein that recognizes GC-rich sequences upstream of many genes (70). Extracts from untreated F9 cells and F9 cells treated with RA and cAMP analogs have similar Sp1 binding activities (64). Sp1 itself may also act as either a repressor or activator depending on post-translational modification, cellular concentration, and binding site context.
Fish that diverged from mammals about 360 and 450 million years ago, respectively. Decapentaplegic (Dpp) is the Drosophila ortholog to Bmp2 and its close relative, Bmp4. Curiously, a 110-bp invariant region is also immediately downstream of the Dpp coding sequence from several fly species that diverged from each other 40 to 80 million years ago (27, 28, 90). Although the fly sequence does not easily align with the metazoan sequence, it is both AU-rich and positioned similarly in the 3′-UTR. This is consistent with the involvement of the region in a very ancient mechanism retained because of the invertebrate-vertebrate split.

The conserved region contains putative RNA stability elements that may mediate the half-life of the Bmp2 transcript. These “AREs” are found in the 3′-UTRs of many highly regulated genes involved in critical signaling processes, e.g., c-fos, c-myc, and various cytokines (78). AREs can influence RNA turnover, translation, and subcellular localization of the mRNA (78). The Bmp2 AREs are clustered after the translation termination codon in all species. In mouse, ARE motifs are absent in the 1.6-kb sequence downstream of the conserved region. This suggests that the elements were actively selected during evolution and that their location is constrained within the transcript.

AREs appear to play dual roles in RNA turnover, in that under different conditions they may stabilize or destabilize mRNAs. Our observations that the luciferase activity and the half-life of transcripts containing the conserved AREs is greater than for transcripts lacking the AREs is consistent with a role in RNA turnover. Undifferentiated, CT-treated cells expressed the reporter constructs containing or lacking the 3′-UTR equally. This suggests that mRNA turnover machinery in undifferentiated cells does not interact with the Bmp2 3′-UTR. In contrast, the 3′-UTR markedly stimulated expression and transcript half-life in differentiated, RACT-treated cells. Future experiments are directed toward determining whether a stabilizing protein binding the Bmp2 transcript is induced in differentiated cells.

In summary, genomic comparisons of transcription units and flanking DNAs from different species can reveal functional sequences that would be difficult to detect using purely experimental methods. Using both wet bench and computer-based methods, we have identified and tested the function of conserved sequences controlling Bmp2 expression. One conspicuously conserved region is downstream of the stop codon. The obvious conservation of this sequence in the Bmp2 3′-UTRs of mammals, birds, amphibians, and fish suggest high functional significance. The AU-rich nature of the element suggests a specific post-transcriptional role that may be empirically tested.
An Evolutionary and Molecular Analysis of Bmp2 Expression
Kevin L. Abrams, Junwang Xu, Celine Nativelle-Serpentini, Shabnam Dabirshahsahebi and Melissa B. Rogers

J. Biol. Chem. 2004, 279:15916-15928. doi: 10.1074/jbc.M313531200 originally published online February 2, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M313531200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 84 references, 48 of which can be accessed free at http://www.jbc.org/content/279/16/15916.full.html#ref-list-1