JOURNAL OF BONE AND MINERAL RESEARCH
Volume 16, Number 6, 2001
© 2001 American Society for Bone and Mineral Research

Characterization of Mouse *Atp6i* Gene, the Gene Promoter, and the Gene Expression*

WENJIE DENG,[1] PHILIP STASHENKO,[1,2] WEI CHEN,[1] YUQIONG LIANG,[1] KEN SHIMIZU,[1] and Y.-P. LI[1,2]

ABSTRACT

Solubilization of bone mineral by osteoclasts depends on the formation of an acidic extracellular compartment through the action of a V-type ATPase. We previously cloned a gene encoding a putative osteoclast-specific proton pump subunit, termed *OC-116 kDa,* approved mouse *Atp6i* (ATPase, H^+ transporting, [vacuolar proton pump] member I). The function of *Atp6i* as osteoclast-specific proton pump subunit was confirmed in our mouse knockout study. However, the transcription regulation of *Atp6i* remains largely unknown. In this study, the gene encoding mouse *Atp6i* and the promoter have been isolated and completely sequenced. In addition, the temporal and spatial expressions of *Atp6i* have been characterized. Intrachromosomal mapping studies revealed that the gene contains 20 exons and 19 introns spanning ~11 kilobases (kb) of genomic DNA. Alignment of the mouse *Atp6i* gene exon sequence and predicted amino acid sequence to that of the human reveals a strong homology at both the nucleotide (82%) and the amino acid (80%) levels. Primer extension assay indicates that there is one transcription start site at 48 base pairs (bp) upstream of the initiator *Met* codon. Analysis of 4 kb of the putative promoter region indicates that this gene lacks canonical TATA and CAAT boxes and contains multiple putative transcription regulatory elements. Northern blot analysis of RNAs from a number of mouse tissues reveals that *Atp6i* is expressed predominantly in osteoclasts, and this predominant expression was confirmed by reverse-transcription polymerase chain reaction (RT-PCR) assay and immunohistochemical analysis. Whole-mount in situ hybridization shows that *Atp6i* expression is detected initially in the headfold region and posterior region in the somite stage of mouse embryonic development (E8.5) and becomes progressively restricted to anterior regions and the limb bud by E9.5. The expression level of *Atp6i* is largely reduced after E10.5. This is the first report of the characterizations of *Atp6i* gene, its promoter, and its gene expression patterns during mouse development. This study may provide valuable insights into the function of *Atp6i,* its osteoclast-selective expression, regulation, and the molecular mechanisms responsible for osteoclast activation. (J Bone Miner Res 2001;16:1136–1146)

Key words: osteoclast gene, characterization of *Atp6i* gene promoter, gene expression pattern, gene organization

INTRODUCTION

OSTEOCLASTS DEGRADE both the inorganic and the organic components of bone.(1) Dissolution of the hydroxyapatite mineral phase is dependent on acidification of the osteoclastic resorption lacuna, via the action of a V-type proton pump.(2–5) Given this important extracellular acidification function, considerable interest has been generated in the structure and function of the osteoclast proton pump as well as its diversification. The unique feature of the osteoclast proton pump prompted us to hypothesize that there is an osteoclast-specific proton pump. Previously, we isolated a novel human gene encoding a putative 116-kDa subunit of the osteoclast proton pump, termed *OC-116 kDa*

*Presented in part at the annual meeting of the International and American Associations for Dental Research, Washington, DC, USA, 2000.

[1]Department of Cytokine Biology, The Forsyth Institute, Boston, Massachusetts, USA.
[2]Harvard-Forsyth Department of Oral Biology, Forsyth Dental Center and Harvard School of Dental Medicine, Boston, Massachusetts, USA.

(approved human *ATP6I*). We investigated the role of *Atp6i* in osteoclastic bone resorption by generating knockout mice with a null mutation in the *Atp6i* gene and showed that the severe osteopetrosis phenotype of *Atp6i*$^{-/-}$ mice results from a defect of osteoclast-mediated extracellular acidification.[6] This result indicates that osteoclast-specific V-ATPases (i.e., V-type proton pump) are structurally and functionally different from other V-ATPases.

However, the genomic organization and sequence of *Atp6i* have not been characterized, and the *Atp6i* gene expression regulation remains largely unknown. To approach an understanding of the factors modulating *Atp6i* gene expression and to better define the function of *Atp6i*, we characterized the mouse *Atp6i* gene together with its regulatory region and the temporal and spatial expression of the *Atp6i* gene.

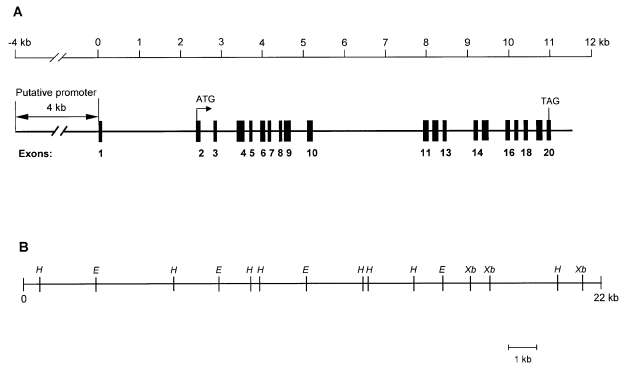

FIG. 1. Genomic organization and restriction map of mouse *Atp6i*. (A) The mouse *Atp6i* gene is composed of 20 exons and 19 introns distributed over approximately 11 kb. Solid bars represent exons. The locations of the initiator *Met* code and stop code TAG are indicated. (B) Restriction map of *λMAtp6i1* containing mouse *Atp6i* gene. *E*, *Eco*RI; *H*, *Hin*dIII; *Xb*, *Xba*I.

MATERIALS AND METHODS

Isolation of mouse Atp6i *gene*

A mouse 129/SVJ genomic library, constructed on lambda FIX II (Stratagene, La Jolla, CA, USA), was used for cloning the mouse *Atp6i* gene using a 2600-base pair (bp) human *ATP6I*/*OC-116 kDa*[7] complementary DNA (cDNA) [^{32}P]-labeled by random primer labeling kit. About 1×10^6 plaques were screened with a [^{32}P]-labeled probe derived from the human full-length *ATP6I* cDNA. The positive clones for mouse *Atp6i* were isolated and subcloned into pBC SK (+) vector (Stratagene). The structure of the isolated genes was determined by restriction endonuclease mapping and Southern blot analysis. Restriction endonuclease–digested DNA was fractionated on 0.8% agarose gel and transferred to nylon membranes. The DNA fragments containing portions of the *Atp6i* gene exons were identified by hybridization with nick-translated DNA probes from different portions of the human *ATP6I* cDNA. Standard methods were used for library screening, purification of phage DNA inserts, mapping of restriction enzyme cut sites, and Southern blotting.[8]

DNA sequencing

A 22-kilobase (kb) fragment containing the 5′-end of the mouse *Atp6i* gene and all 20 exons was isolated from a bacteriophage clone and subcloned into pBC SK (+) vector (Stratagene). The subclone was sequenced as double-stranded templates with Sequenase in combination with dideoxynucleotide chain terminators according to the supplier's recommendation (U.S. Biochemical Corp., Cleveland, OH, USA) or with an ABI 377 sequencer. Initial DNA sequencing was performed with T7 and T3 primers. Subsequent reverse orientation sequencing was performed using insert-specific primers. The DNA sequence was assembled and analyzed for the location of exons and introns by comparison to the sequence of human *ATP6I* cDNA using the DNAsis II program (Hitachi Software, San Bruno, CA, USA). The putative promoter region of the gene was searched for transcription factor binding sites using the DNAsis II program.

Northern blotting

Total RNA from mouse tissues including brain, liver, thymus, lung, kidney, spleen, skeletal muscle, and tartrate-resistant acid phosphatase–positive (TRAP$^+$) multinuclear osteoclast-like cells (OCLs) derived from mouse osteoclast precursor-5 (MOCP-5) were prepared as described.[9] Northern blot analysis of mouse *Atp6i* was performed as described.[10] Fifteen micrograms of total RNA was fractionated by electrophoresis through 1% agarose/formaldehyde gel followed by Northern blot transfer to nitrocellulose membrane presoaked in 20× SSC and then covalently bound by baking for 1 h at 80°C. Visualization by ethidium bromide was used to evaluate the integrity of messenger RNA (mRNA) on equal loading per lane. The nylon membrane was hybridized with radioactively labeled human *ATP6I* cDNA as a probe. After washing, the filter was exposed to X-ray film (Fuji, Tokyo, Japan) at −80°C for autoradiography. Analysis of the films was carried out using densitometric scanning (Molecular Dynamics, Sunnyvale, CA, USA).

Reverse-transcription polymerase chain reaction

Reverse-transcription polymerase chain reaction (RT-PCR) was performed using 0.5 μg of total RNA from mouse multiple tissues and OCLs by the Access RT-PCR System (Promega, Madison, WI, USA). The conditions for RT-PCR were 48°C for 45 minutes followed by 25 cycles of 94°C for 45 s, 64°C for 30 s, and 72°C for 1 minute. The PCR products were analyzed by separation on 2% agarose gels. The level of mRNA expression in each tissue was normalized to the expression of mouse GAPDH. The following primers were used for RT-PCR: *Atp6i*, 5′-ATGTTCCGGAGTGAAGAGGTG-3′ and 5′-TCCTGGATGCGCAGCAG-

TABLE 1. EXON-INTRON JUNCTION OF THE MOUSE *ATP6I* GENE

Exon number and size (nt)	*cDNA position of exon*	*5′ Splicing site*	*Intron number and size (kb)*	*3′ Splicing site*
1 (44)	1–44	CCCATTCCG gtatttcga	1 (2.383)	gcccttaag GATCATGGG
2 (121)	45–165	TTCAGAGAC gtaagtggg	2 (0.282)	ccatcacag CTCAACGAA
3 (79)	166–244	AGACGTTTA gtaagttgg	3 (0.493)	tgtgtctag CCTTCTTGC
4 (222)	245–466	AGCCCGCCG gtcagcccc	4 (0.087)	aagtggcag CTGAACATA
5 (79)	467–545	GAAGGTCAA gtgagtcgg	5 (0.167)	tgttggcag TTTTGTGGC
6 (127)	546–672	CCAGTGACG gtgaatttg	6 (0.079)	ctttcacag GGTGAGCCT
7 (83)	673–755	CACAGACTG gtgagccac	7 (0.158)	cttcctcag CTTCCACTG
8 (94)	756–849	CTTCAGGAA gtgagtgcc	8 (0.074)	tgtccccag GTCCTGGGG
9 (213)	850–1062	AGCGGCTCA gtgagcatc	9 (0.308)	gtaccccag AGTGAGGAA
10 (145)	1063–1207	TAAACCCTG gtaagatcc	10 (2.772)	tgcctgcag CTCCCTACA
11 (140)	1208–1347	CAAAACGAG gtacgaacc	11 (0.104)	gacccccag ATCTGGCAG
12 (158)	1348–1505	AGGCTGGAG gtatgacct	12 (0.079)	ttactgcag TGACGAGTA
13 (91)	1506–1596	ATTGACCCG gtaagtggg	13 (0.667)	ctaccccag ATCTGGAGC
14 (119)	1597–1715	CAACCACGT gtgagctaa	14 (0.076)	ttcccacag GCACTTTGG
15 (214)	1716–1929	CATGGGCAG gtaggccag	15 (0.298)	tgcccacag GAAGTGGTA
16 (135)	1930–2064	GGCCAGCAG gtagacgca	16 (0.135)	tcttctcag GATGAGGAC
17 (108)	2065–2172	GAAACTGAG gtgggcaga	17 (0.073)	cttaccgtg CAGTTTGTC
18 (118)	2173–2290	CCCATGCCC gtgagtgac	18 (0.158)	tgtgtgcag AGCTGTCTG
19 (178)	2291–2468	GTTGCACTG gtgagcacc	19 (0.077)	cacacccag GGTGGAGTT
20 (141)	2469–2609			

GTC-3′; glyceraldehyde-3-phosphate dehydrogenase (GAPDH), 5′-GCCATCAACGACCCCTTCATTG-3′ and 5′-ACTCCACGACATACTCAGCACC-3′.

Identification of the transcription start site of mouse Atp6i *gene in osteoclast*

The initiation site for the transcription of mouse *Atp6i* gene in osteoclasts was analyzed by primer extension as described.[11] Two synthetic 21 base-oligomers were used. They are located at −6 to −26 (5′-GGAATGGGTG-TGGCCTGTGAC-3′) and −29 to −49 (5′-CTGCCCAC-TGCCCGCCCTGTC-3′) upstream of the initiator *Met* codon (as +1) of the mouse *Atp6i* cDNA predicted from the mouse exons compared with the human cDNA. The primers were end-labeled with [γ-^{32}P]adenosine triphosphate (ATP) using T4 polynucleotide kinase (Life Technologies, Inc., Rockville, MD, USA). The [^{32}P]-labeled primers were hybridized with 10 μg of total RNA from mouse OCLs prepared as described[9] and extended using avian myeloblastosis virus (AMV) reverse transcriptase. The extended cDNA product was analyzed on a sequencing gel with DNA sequence reactions as size markers.

Atp6i *immunostaining and histological analysis*

For histological analysis, newborn mice were fixed in 4% paraformaldehyde, dehydrated in ethanol, and embedded in paraffin. Serial sections were taken at 5 μm and stained with hematoxylin and eosin according to standard procedures. The expression of *Atp6i* in osteoclasts was investigated by immunostaining various tissues in newborn mouse sections with an *ATP6I*-specific polyclonal antibody[12] and developed with peroxidase-coupled secondary antibodies according to the manufacturer's instructions (Rabit Elite ABC Kit; Vector Laboratories, Burlingame, CA, USA).

Whole-mount in situ hybridization

Whole-mount RNA in situ hybridization was done as described[13] using a 2.6-kb cDNA fragment containing the entire *ATP6I* coding region. Digoxigenin-labeled RNA probe for *ATP6I* was used. In brief, embryos were dissected free of extraembryonic membranes, and any cavities (e.g., the neural tube and amniotic cavity) were dissected open to avoid high backgrounds caused by the subsequent trapping of reagents. Embryos were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) overnight and washed with PBS and 0.1% Tween 20 (PBT) twice for 5 minutes. The embryos were then dehydrated with a methanol series in PBT, rehydrated through the methanol/PBT series in reverse, and washed twice with PBT Embryosine (Sigma, St. Louis, MO, USA). Embryos were incubated in 6% (vol/vol) hydrogen peroxide in PBT for 1 h, washed with PBT, and treated with 10 μg/ml proteinase K in PBT for 15 minutes before washing twice for 5 minutes each with freshly prepared 2 mg/ml of glycine in PBT and twice with PBT. Embryos were then refixed with fresh 0.2% (vol/vol) glutaraldehyde/4% (vol/vol) paraformaldehyde in PBT for 20 minutes. After washing twice with PBT and adding prehybridization solution (50% formamide, 5× SSC, pH 5, 50 μg/ml yeast RNA, 1% sodium dodecyl sulfate [SDS], and 50 μg/ml heparin), embryos were rocked at 70°C for 2–3 h. The 0.4 ml of hybridization mix containing approximately 1 μg/ml

FIG. 2. DNA sequence of the mouse *Atp6i* gene. Coding sequences are shown together with the translated amino acid sequence. The transcription initiation site is marked as +1. The consensus polyadenylation signal sequences (AATAAA) are underlined. The 3′ flanking sequence beyond the polyadenylation site is shown below. Exons are indicated by capital letters. The 5′ flanking sequence, introns, and 3′ flanking sequence are indicated by small letters.

```
Mouse Atp6i            1    MGSMFRSEEVALVQLLLPTGSANNCVSQLGELGFVEFRDLNESVSAFQRRFVVDVRRCEE
Human ATP6I/OC-116kDa  1    MGSMFRSEEVALVQLFLPTAAAYTCVSRLGELGLVEFRDLNASVSAFQRRFVVDVWRCEE
Consensus                    MGSMFRSEEVALVQL-LPT--A--CVS-LGELG-VEFRDLN-SVSAFQRRFVVDV-RCEE

Mouse Atp6i            61   LEKTFTFLREELQRAGLTLAPPEGTLPAPPPRDLLRIQEETDRLAQELRDVRGNQQALRA
Human ATP6I/OC-116kDa  61   LEKTFTFLQEEVRRAGLVLPPPKGRLPAPPPRDLLRIQEETERLAQELRDVRGNQQALRA
Consensus                    LEKTFTFL-EE--RAGL-L-PP-G-LPAPPPRDLLRIQEET-RLAQELRDVRGNQQALRA

Mouse Atp6i            121  QLHQLRLHSAVL--GQKPQPAAEHTEDPSLRQHPASRHRGPHSDLKVNFVAGAVEPYKAA
Human ATP6I/OC-116kDa  121  QLHQLQLHAAVLRQGHEPQLAAAHTDGASERTPLLQAPGGPHQDLRVNFVAGAVEPHKAP
Consensus                    QLHQL-LH-AVL--G--PQ-AA-HT---S-R--------GPH-DL-VNFVAGAVEP-KA-

Mouse Atp6i            179  ALERLLWRACRGFLIASFRETEGQLEDPVTGEPATWMTLFISYWGEQIGQKIRKITDCFH
Human ATP6I/OC-116kDa  181  ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEQIGQKIRKITDCFH
Consensus                    ALERLLWRACRGFLIASFRE-E--LE-PVTGEPATWMT--ISYWGEQIGQKIRKITDCFH

Mouse Atp6i            239  CHVFPYLEQEEARFRTLQQLQQQQSQELQEVLGETDRFLSQVLGRVQQLLPPWQVQIHKMK
Human ATP6I/OC-116kDa  241  CHVFPFLQQEEARLGALQQLQQQQSQELQEVLGETERFLSQVLGRVLQLLPPGQVQVHKMK
Consensus                    CHVFP-L-QEEAR---LQQLQQQQSQELQEVLGET-RFLSQVLGRV-QLLPP-QVQ-HKMK

Mouse Atp6i            299  AVYLTLNQCSVNTTHKCLIAEVWCAARDLPTVQQALQSGSSEEGVSAVAHRIPCQDMPPT
Human ATP6I/OC-116kDa  301  AVYLALNQCSVSTTHKCLIAEAWCSVRDLPALQEALRDSSMEEGVSAVAHRIPCRDMPPT
Consensus                    AVYL-LNQCSV-TTHKCLIAE-WC--RDLP--Q-AL---S-EEGVSAVAHRIPC-DMPPT

Mouse Atp6i            359  LIRTNRFTSSLQGIVDAYGVARYREVNPAPYTIITFPFLFAVMFGDVGHGLLMFLFALAM
Human ATP6I/OC-116kDa  361  LIRTNRFTASFQGIVDRYGVGRYQEVNPAPYTIITFPFLFAVMFGDVGHGLLMFLFALAM
Consensus                    LIRTNRFT-S-QGIVD-YGV-RY-EVNPAPYTIITFPFLFAVMFGDVGHGLLMFLFALAM

Mouse Atp6i            419  VLTENRPAVKAAQNEIWQTFFGGRYLLLLMGLFSVYTGFIYNECFSRATTIFPSSWSVAA
Human ATP6I/OC-116kDa  421  VLAENRPAVKAAQNEIWQTFFRGRYLLLLMGLFSIYTGFIYNECFSRATSIFPSGWSVAA
Consensus                    VL-ENRPAVKAAQNEIWQTFF-GRYLLLLMGLFS-YTGFIYNECFSRAT-IFPS-WSVAA

Mouse Atp6i            479  MANQSGWSDEYLSQHSMLTLNPNITGVFLGPYPFGIDPIWSLATNHLSFLNSFKMKMSVI
Human ATP6I/OC-116kDa  481  MANQSGWSDAFLAQHTMLTLDPNVTGVFLGPYPFGIDPIWSLAANHLSFLNSFKMKMSVI
Consensus                    MANQSGWSD--L-QH-MLTL-PN-TGVFLGPYPFGIDPIWSLA-NHLSFLNSFKMKMSVI

Mouse Atp6i            539  LGVTHMAFGVFLSIFNHVHFGQAHRLLLETLPELIFLLGLFGYLVFLIVYKWVNVSAASA
Human ATP6I/OC-116kDa  541  LGVVHMAFGVVLGVFNHVHFGQRHRLLLETLPELTFLLGLFGYLVFLVIYKWLCVWAARA
Consensus                    LGV-HMAFGV-L--FNHVHFGQ-HRLLLETLPEL-FLLGLFGYLVFL--YKWl-V-AA-A

Mouse Atp6i            599  SSAPSILIHFINMFLFSQNPTNHLLFHGQEVVQYVLVVLGFGYRSYPVAGHTLVPAAPST
Human ATP6I/OC-116kDa  601  AS-PSILIHFINMFLFSHSPSNRLLYPRQEVVQATLVVLALAMVPILLLG---TPLHLLH
Consensus                    -S-PSILIHFINMFLFS--P-N-LL---QEVVQ--LVVL----------G----P-----

Mouse Atp6i            659  ATEETLREGPAGQQDEDTDKLLASPDASTLENSWSPDEEKAGSPGDEE-TEFVPSEIFMH
Human ATP6I/OC-116kDa  657  RHRRRLRRRPADRQEENKAGLLDLPDASV--NGWSSDEEKAGGLDDEEEAELVPSEVLMH
Consensus                    -----LR--PA--Q-E----LL--PDAS---N--S-DEEKAG---DEE--ElVPSE--MH

Mouse Atp6i            718  QAIHTIEFCMGCVSNTATYLRLWALSLAHAQLSEVLWAMVMRIGLGMGREIGVAAVVLVP
Human ATP6I/OC-116kDa  715  QAIHTIEFCLGCVSNTASYLRLWALSLAHAQLSEVLWAMVMRIGLGLGREVGVAAVVLVP
Consensus                    QAIHTIEFC-GCVSNTA-YLRLWALSLAHAQLSEVLWAMVMRIGLG-GRE-GVAAVVLVP

Mouse Atp6i            778  VFAAFAVLTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGTGYKLSPFTFTVDSD
Human ATP6I/OC-116kDa  775  IFAAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGTGYKLSPFTFAATDD
Consensus                    -FAAFAV-TVAILLVMEGLSAFLHALRLHWVEFQNKFYSGTGYKLSPFTF----D
```

FIG. 3. Alignment of mouse *Atp6i* amino acid sequence derived from the genomic exons with that of human *ATP6I/OC-116 kDa* predicted from the cDNA.

of digoxigenin-labeled RNA probe was added and incubated overnight at 70°C with rocking. After hybridization, the embryos were washed, treated with ribonuclease, and high-stringently washed to remove unannealed probe. This was followed by incubation with antidigoxigenin antibody conjugated to alkaline phosphatase in the dark. The reaction was monitored at intervals under a dissecting microscope until the color had developed to the desired extent.

RESULTS

Isolation of the mouse Atp6i *genomic clone*

Eight λ-bacteriophage clones were identified and isolated from the Sv129 strain mouse genomic library using a full-length human *ATP6I* cDNA probe. The DNA from eight λ-clones was hybridized with different portions of human cDNA by Southern blot analysis. This resulted in the isolation of one bacteriophage clone *λMAtp6i1* with an insert spanning approximately 22 kb of genomic DNA. The approximate location of the 5′-end of the mouse gene and exons within the cloned region was determined by Southern blot analysis using different portions of the human *ATP6I* cDNA as probes.

Gene structure and sequence

The *Eco*RI genomic DNA fragments from *λMAtp6i1* were subcloned into pBC SK (+) (Stratagene) for sequencing. A primer walking strategy was used to obtain 15-kb contiguous sequences covering all exons, introns, the 4-kb putative promoter region of the gene to the transcription start site, and 0.5-kb 3′ sequences to the polyadenylation signal (Fig. 1A). The genomic sequence is divided into 20 exons ranging from 44 to 222 nucleotides (nt; Table 1). Exon 1 contains the 5′ untranslated region. Exon 2 contains the translation initiation site. Exon 20 contains the stop codon and the 3′ untranslated region. There is a putative consensus polyadenylation signal (AATAAA) located 23 bp

upstream from the poly(A) tract. The 19 introns ranged in size from 73 bp to 2.8 kb, with the entire gene approximately 11 kb (Fig. 2). Sequences of the genomic clone were used to construct a partial restriction map of the entire *Atp6i* gene (Fig. 1B). The restriction map was confirmed by Southern blot using different portions of human *ATP6I* cDNA as the probes.

We have determined the exon-intron boundaries of mouse *Atp6i*. As shown in Table 1, the exon-intron boundary sequences conformed to the GT/AG rule with a single exception of GT/TG in intron 17. The boundary sequences generally were consistent with the 5′ and 3′ splice consensus sequences. Variations from the consensus sequences were all permissible for primate exon-intron boundary sequences.

Alignment of the mouse *Atp6i* gene exon sequences to the human cDNA revealed that there is strong homology at both the nucleotide (82%) and the amino acid (80%) levels (Fig. 3). In contrast, the sequence homology with the human gene at the 5′ and 3′ untranslated regions is only 60% and 56%, respectively.

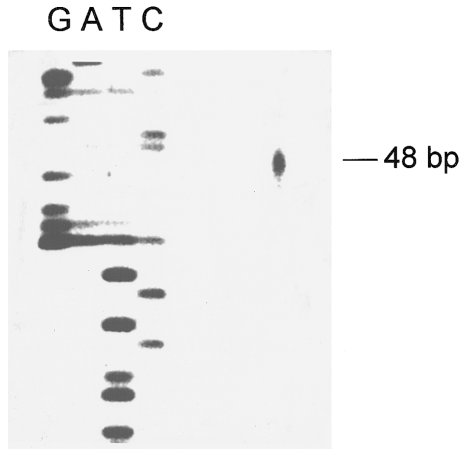

FIG. 4. Identification of the mouse *Atp6i* transcription initiation site by primer extension. The primer extension product was determined on denaturing polyacrylamide gel electrophoresis using the adjacent DNA sequencing reaction (GATC) as a size marker. The major primer extension products were 48 bp long, from nt −1 to the putative transcription start site.

Identification of the transcription start site of the Atp6i *gene in osteoclasts by primer extension*

To determine the mouse *Atp6i* transcription initiation site, primer extension analysis was performed using mouse *Atp6i*–specific antisense oligonucleotide primers. The 5′ untranslated region (UTR) of the mouse *Atp6i* gene was predicted by homology comparison with that of the human cDNA. Based on this potential exon 1 sequence, we made the antisense primer for the primer extension experiment. Two end-labeled antisense primers, −6 to −26 (5′-GGAATGGGTGTGGCCTGTGAC-3′) and −29 to −49 (5′-CTGCCCACTGCCCGCCCTGTC-3′), relative to the ATG initiation codon, were annealed to total RNA isolated from OCLs. The primers were extended with AMV reverse transcriptase and analyzed by polyacrylamide-urea gel electrophoresis. The primer extension assay indicates that there is a transcription start site at 48 bp upstream of the initiator *Met* codon (Fig. 4). No product was observed with the upstream primer (−29 to −49; data not shown). Therefore, the transcript initiation at 48 bp is the major transcription initiation site in mouse OCLs.

Structural analysis of the putative promoter region of the mouse Atp6i *gene*

To identify potential regulatory elements, a 4000-bp sequence of the *Atp6i* gene promoter was analyzed by database search. There is no TATA or CAAT box located 5′ of the transcriptional start site.

Although the important functional role of transcription factors c-*fos*, nuclear factor κB (NF-κB), and PU.1 in osteoclast differentiation has been evaluated in vivo by mouse knockout,[14–16] the factors that regulate osteoclast gene expression have not been delineated. Because *Atp6i* is expressed specifically in osteoclasts, putative AP-1, NF-κB, and PU.1 response elements were specifically sought in this genomic sequence. Several AP-1, NF-κB, and one PU.1 response elements were found. The potential GATA and bHLH binding sites located in the 5′-flanking regulatory region of the mouse *Atp6i* gene (Fig. 5, Table 2) indicate that members of bHLH and GATA transcription factor families also may play an important role in osteoclast development. Other putative regulatory elements found included AP-3, Ets1, H-APF-1, SP1, PEA, and C/EBP (Fig. 5; Table 2).

Tissue distribution of Atp6i *mRNA and protein*

The tissue and cellular distribution of the human *ATP6I* has been characterized.[7] Mouse tissue distribution of *Atp6i* mRNA was investigated by Northern blot analysis using the human full-length cDNA fragment as a hybridization probe. The results indicated that the mouse *Atp6i* was specifically and highly expressed in OCLs and was undetectable in other tissues including brain, liver, thymus, lung, kidney, spleen, skeletal muscle, etc. (Fig. 6A). This result is consistent with our human data.[7] However, with long exposure, a 3.8-kb transcript was detected in several mouse tissues, including brain, heart, liver, lung, spleen, kidney, and thymus, in about 100-fold lower level when compared with that in osteoclasts (Figs. 6B and 6D). To further investigate the tissue distribution of *Atp6i*, we performed RT-PCR using total RNA from various mouse tissues (Fig. 7). The results show that at 25 cycles of RT-PCR, the mouse *Atp6i* was predominantly and highly expressed in OCLs. A very weak expression was detected in mouse liver and heart (Fig. 7). The mouse tissue distribution of *Atp6i* protein was investigated by immunostaining newborn mouse sections with anti-*ATP6I* polyclonal antibodies. Sections of newborn mouse embryo were incubated with the polyclonal antibodies and then developed with peroxidase-coupled secondary antibodies. Antibodies did not react with extraskeletal tis-

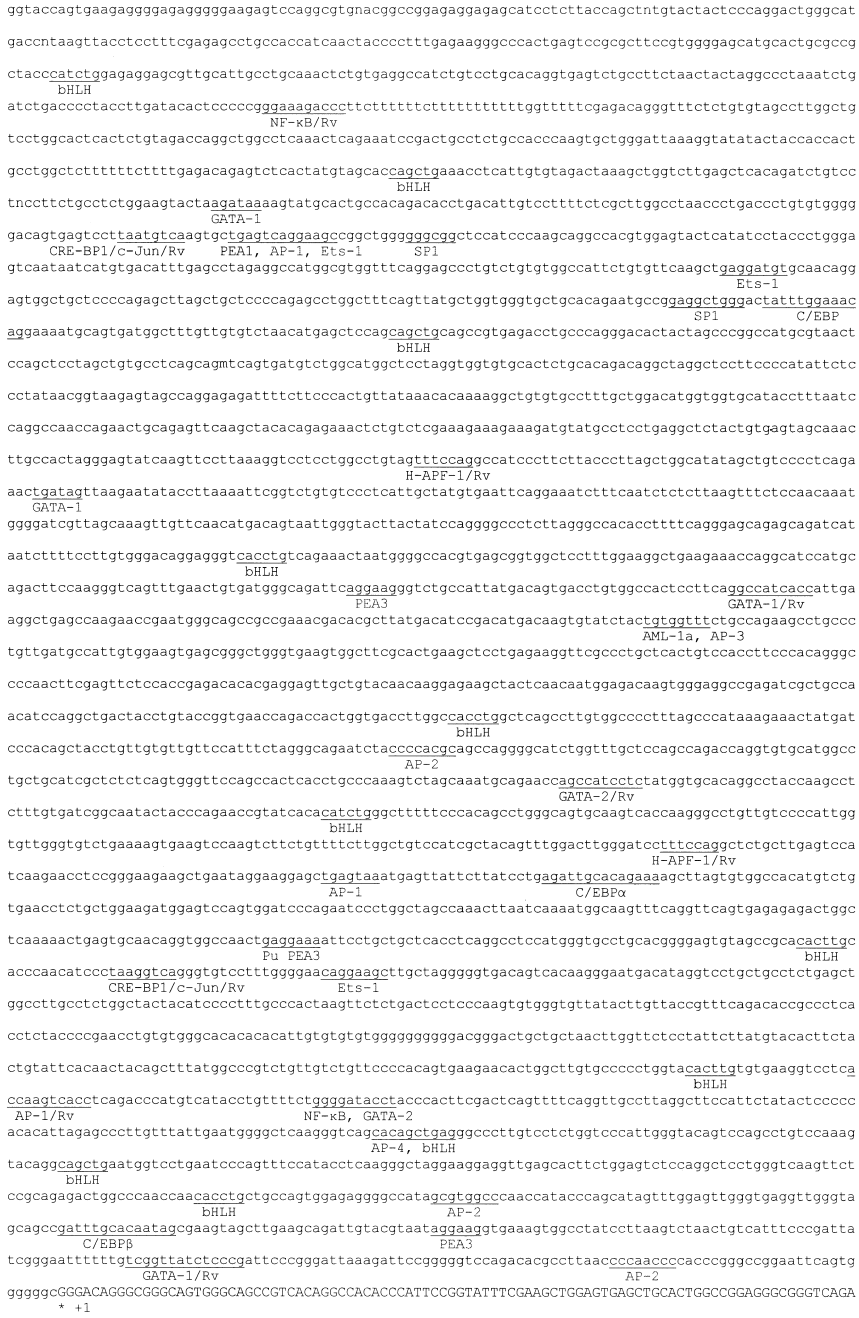

FIG. 5. Nucleotide sequence of the putative promoter region of mouse *Atp6i* gene. Numbers on the left are the nucleotide positions relative to the transcriptional initiation site. Putative regulatory elements are underlined.

sues or cells including heart, lung, liver, muscle, kidney, and brain (Figs. 8B–8D). In contrast, osteoclasts in the lumbar vertebrae showed intense *Atp6i* expression (Fig. 8A, arrows).

Expression of Atp6i *at somite stage and organogenesis stage revealed by whole-mount RNA in situ hybridization*

We found that the expression of *Atp6i* during embryo development is dynamic. No expression was detected before and at E7.5 (Fig. 9A). *Atp6i* expression was detected initially in the headfold region and the posterior region in somite stage (E8.5; Fig. 9B, arrows). By E9.5, expression of *Atp6i* became progressively restricted in regions of midbrain, forebrain, and limb bud (Fig. 9C, arrows). The expression level of *Atp6i* was reduced largely on E10.5 and became undetectable on E12.5 (Figs. 9D and 9E). In our previous studies, we showed that *Atp6i*-deficient mice exhibit severe osteopetrosis because of loss of osteoclast-mediated extracellular acidification. The expression of *Atp6i* at the early stage of mouse development led us to

TABLE 2. THE POTENTIAL RECOGNITION SEQUENCES FOR TRANSCRIPTION FACTORS IN THE MOUSE *ATP6I* REGULATION REGION

Name of cis element	*Position*	*Sequence*
AP-1	−3280	TGAGTCA
	−1269	TGAGTAA
AP-2	−256	GCGTGGCC
	−1661	CCCCACGC
	−29	CCCAACCC
AP-3	−2031	TGTGGTTT
bHLH	−1754, −284	CACCTG
	−1469	CATCTG
	−1013, −626	CACTTG
	−461	CAGCTG
Ets1	−3275, −962	CAGGAAGC
	−3122	GAGGATGT
GATA-1	−3382	AGATAA
	−2503	TGATAG
PEA1	−3280	TGAGTCA
PEA3	−2166, −156	AGGAAG
	−1075	AGGAAA
SP1	−3259	GGGCGG
	−3028	GAGGCTGGG
NF-κB	−570	GGGGATACCT
NF-κB/Rv	−3676	GGAAAGACCC
AML-1a	−2031	TGTGGT
H-APF-1/Rv	−2558, −1329	TTTCCAG
C/EBP	−3017	TATTTGGAAACAG
C/EBPα	−1243	AGATTGCACAGAAA
C/EBPβ	−200	GATTTGCACAATAG

examine the developmental normality of *Atp6i*$^{-/-}$ embryo. No developmental defect of *Atp6i*$^{-/-}$ embryo was found from the examination (data not shown). The significance of the expression of *Atp6i* at the early stage of mouse development remains to be explored.

DISCUSSION

Atp6i has been identified as a critical osteoclast proton pump subunit in bone degradation and remodeling.[6] Xie and Stone et al.[17] were the first to characterize the 116-kDa subunit in vacuolar ATPase from a variety of sources. We determined *ATP6I* as a possible osteoclast-specific subunit by comparing the amino acid sequence of *ATP6I* (a3) with that of the 116-kDa subunit (a1) that was cloned by Xie and Stone.[17] The amino acid of a1 showed 47% homology with that of a3.[7] In this study we aimed to characterize the mouse *Atp6i* gene with its regulatory region and analyze the temporal and spatial *Atp6i* expression as a starting point for the investigation of the regulation of the gene. The mouse *Atp6i* gene, including approximately 4 kb of the putative promoter region, was sequenced and characterized. We found that the *Atp6i* gene has 20 exons. The sequence

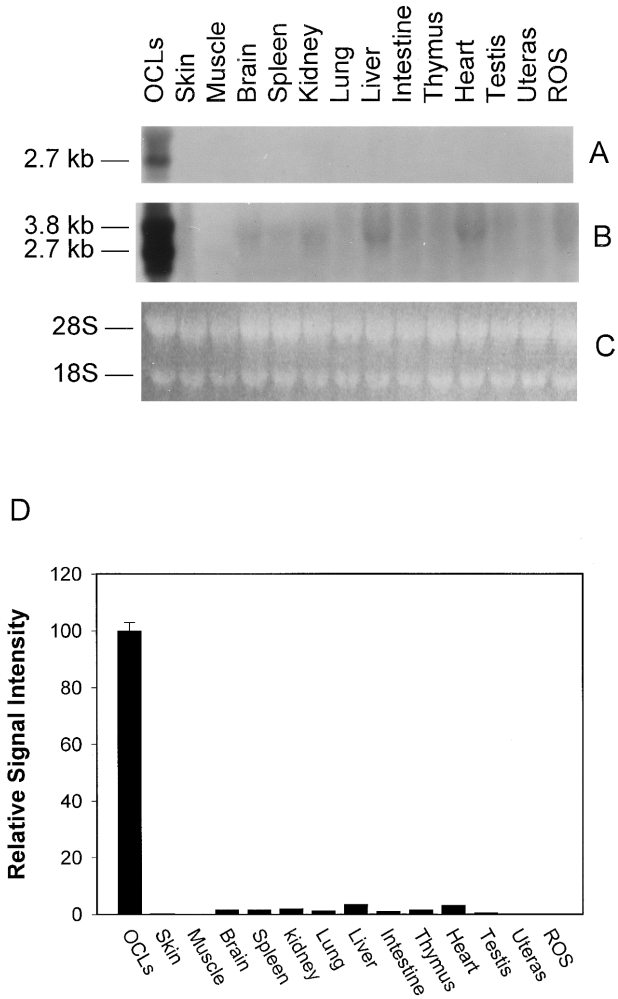

FIG. 6. Northern hybridization of human *ATP6I* cDNA to total cell RNA from cells and mouse tissues. Fifteen micrograms of total RNA from cells and various mouse tissues were blotted onto a nylon filter. (A) Two days exposure of autoradiography after hybridization with human *ATP6I* cDNA probe. (B) Eighteen days exposure of autoradiography after hybridization with human *ATP6I* cDNA probe. (C) Visualization by ethidium bromide of total RNA showing equivalent RNA loading. (D) The quantitative levels of *Atp6i* RNA on Northern blot was evaluated from the value of densitometric scanning and normalized to the level of 18S and 28S.

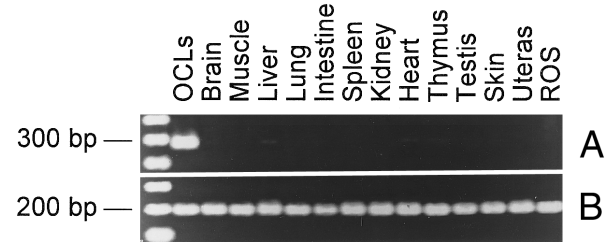

FIG. 7. Tissue-specific expression of mouse *Atp6i* by RT-PCR. RT-PCR was performed using total RNA isolated from cells and various mouse tissues as described in the Materials and Methods section. (A) The expression pattern of *Atp6i* in different cells and mouse tissues at 25 cycles of RT-PCR. *Atp6i* is expressed in osteoclast at a high level and very low level in mouse liver and heart. (B) Mouse GAPDH was used to normalize the mRNA expression levels.

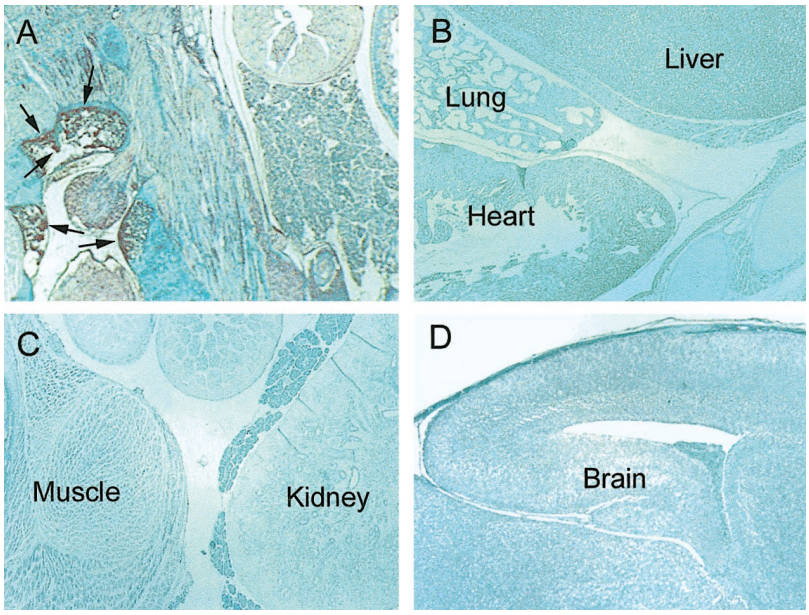

FIG. 8. Characterization of the *Atp6i* gene expression by immunostaining with section of newborn mouse. Newborn mouse sections were incubated with anti-*ATP6I* rabbit polyclonal antibodies and then developed with peroxidase-coupled secondary antibodies. (A) Osteoclasts in the vertebrae showed intense *Atp6i* expression (arrows). (B–D) Antibodies did not react with extraskeletal tissues or cells including heart, lung, liver, muscle, kidney, and brain.

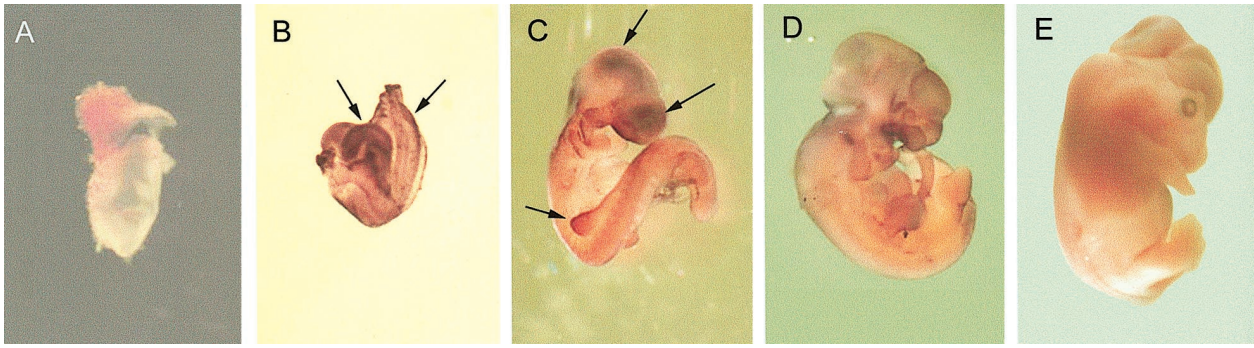

FIG. 9. Characterization of the *Atp6i* gene expression by whole-mount in situ hybridization during embryonic development. Whole-mount in situ hybridization showed that *Atp6i* expression (A) was undetectable at E7.5 and (B, arrows) was initially detected in headfold region and posterior region in somite stage (E8.5) and (C, arrows) became progressively restricted to anterior regions and limb bud by E9.5. The expression level of *Atp6i* (D) was largely reduced on E10.5 and (E) was undetectable on E12.5.

derived from these exons is highly conserved with the human *ATP6I* cDNA. Northern analysis of various mouse tissues indicated that *Atp6i* was expressed selectively and predominantly in osteoclasts, a finding consistent with a previous observation of the human *ATP6I* gene.[7] This is the first report describing the genomic sequence and the structure of mouse *Atp6i* gene. The structure and organization of human *ATP6I* gene have yet to be characterized. Osteopetrosis phenotype of *Atp6i* knockout mice present many similarities with the fatal infantile malignant form of human osteopetrosis that is lethal within the first decade in the absence of bone marrow transplantation.[18,19] Without knowledge of the genetic defect explaining the osteoclastic dysfunction, it has been difficult to delineate the mechanism of the malignant form of osteopetrosis. Our study may facilitate a characterization of human osteopetrosis disease genes.

Three isoforms (a1, a2, and a3) of the 116-kDa subunit encoded by different genes were reported.[20–22] The subunit that Stone et al.[17] have purified from clathrin-coated vesicles of brain as well as from osteoclasts is the a1 isoform that is different from *ATP6I* (*OC-116 kDa*).[7] The a1 isoform is highly expressed in brain and osteoclast.[17] In contrast, *Atp6i*, also termed a3, is expressed predominately in osteoclasts. Based on our original results,[7] we proposed that *ATP6I* is expressed specifically in osteoclasts. However, when mRNA instead of total RNA was used in Northern blot, other investigators[20] detected expression of *Atp6i* (a3) in other tissue. To verify this information we performed an additional Northern blot analysis on osteoclasts and other tissues with longer exposure time. Interestingly, when the blot was exposed for 2 days, a strong signal was detected in osteoclasts. However, expression of *Atp6i* in other tissues was undetectable. In contrast, after 18 days exposure, a very

low expression of *Atp6i* in high-molecular spliced variant was detected in liver, heart, brain, etc. (Fig. 6B). The expression level of *Atp6i* in osteoclasts is 40-, 45-, and 80-fold higher than that in liver, heart, and brain, respectively (Figs. 6B and 6D). We also have carried out immunohistochemical analysis on different tissues. At protein expression level, *Atp6i* was detected only in osteoclasts (Fig. 8A), but not in other tissues (Figs. 8B–8D). Moreover, we used a more sensitive assay RT-PCR to confirm the very low expression level of *Atp6i* in other tissues. Based on our data in this study, we state that *Atp6i* is expressed predominately in osteoclasts. Using whole-mount in situ hybridization, we found that *Atp6i* is expressed at somite and organogenesis stages of mouse embryonic development. However, the histological analysis of E12.5 *Atp6i* null mutation did not reveal any developmental defect (data not shown). A redundancy function of the other proton pump subunit may be substituting for the *Atp6i* function. The function of *Atp6i* at the early stage of mouse embryonic development remains to be explored.

The signals that regulate the expression of *Atp6i* are unknown. However, recently, genetically engineered osteopetrotic mutant mice have yielded important insights into the regulation of osteoclast differentiation. Disruption of the c-*fos* proto-oncogene, which is a component of the AP-1 complex, blocks osteoclast development at the point of divergence from the common CFU-GM precursor, before the expression of osteoclast genes (e.g., *Atp6i* and cathepsin K) and leads to an osteopetrotic phenotype.[14] PU.1 knockout mice are osteopetrotic and are devoid of both osteoclasts and macrophages.[16] Iotsova et al. reported that mice lacking NF-κB1 and NF-κB2 (double-knockout mice) developed osteopetrosis because of a defect in osteoclast differentiation.[15] Because *Atp6i* is expressed specifically in osteoclasts, the osteopetrotic phenotype in these knockout mice suggests that AP-1, PU.1, and NF-κB sequences may exist in the promoter region of *Atp6i* and may be critical in regulating expression of the *Atp6i* gene. Therefore, AP-1, NF-κB, and PU.1 binding site sequences were sought specifically by computer search, and several consensus sequences for AP-1 and NF-κB binding site were found. However, only one PU.1 binding site was found in the 4-kb promoter region. In addition, several putative regulatory *cis* elements for Ets1, C/EBP, AP-2, H-APF-1, AP-3, PEA1, and PEA3 were observed. Recently, a clustering of PEA3/Ets and AP-1 sequences in the 92-kDa type IV collagenase was found to be essential for transcriptional activation on induction of *ras* in ovarian tumor–derived OVCAR 3 cells.[23] The 92-kDa type IV collagenase is an enzyme highly expressed in osteoclasts.[24] The involvement of interleukin-6 (IL-6) in stimulation of osteoclast activity has been described in numerous reports.[25,26] Consensus H-APF-1 sequences for IL-6 is located at −1329.

Recently, we have characterized the mouse cathepsin K gene that is expressed abundantly and selectively in osteoclasts. Like the putative promoter region for mouse cathepsin K gene,[12] the mouse *Atp6i* putative promoter region also shows no conventional TATA or CAAT boxes typical of most RNA polymerase II transcriptional units. The multiple putative transcription regulation elements mentioned previously in mouse *Atp6i* promoter region also are present in mouse cathepsin K promoter region, indicating that they may be important in regulating the expression of *Atp6i* gene. Interestingly, the NF-κB binding site in mouse *Atp6i* promoter region (−570) is located much more proximally than that in mouse cathepsin K promoter region (−8886) relative to the transcription start site. Further studies are required to characterize and examine the function of these putative regulatory elements in controlling *Atp6i* gene expression.

In summary, we have isolated and sequenced the complete mouse *Atp6i* gene, including the 4-kb putative promoter region. Multiple regulatory elements in the *Atp6i* gene were identified. The temporal and spatial expression of *Atp6i* has been determined. These findings will provide direction toward future studies to define the mechanism of *Atp6i* transcriptional control.

ACKNOWLEDGMENTS

We thank Dr. Kalu Ogbureke and Ms. Susan Orlando for the critical reading of this article. This work was supported by the National Institutes of Health (NIH) grants DE-07378 (P.S.) and AR44741 (Y.-P.L.).

REFERENCES

1. Blair HC, Kahn AJ, Crouch EC, Jeffrey JJ, Teitelbaum SL 1986 Isolated osteoclasts resorb the organic and inorganic components of bone. J Cell Biol **102:**1164–1172.
2. Sundquist K, Lakkakorpi P, Wallmark B, Vaananen K 1990 Inhibition of osteoclast proton transport by bafilomycin A1 abolishes bone resorption. Biochem Biophys Res Commun **168:**309–313.
3. Vaes G 1968 On the mechanisms of bone resorption. The action of parathyroid hormone on the excretion and synthesis of lysosomal enzymes and on the extracellular release of acid by bone cells. J Cell Biol **39:**676–697.
4. Baron R, Neff L, Louvard D, Courtoy PJ 1985 Cell-mediated extracellular acidification and bone resorption: Evidence for a low pH in resorbing lacunae and localization of a 100-kD lysosomal membrane protein at the osteoclast ruffled border. J Cell Biol **101:**2210–2222.
5. Blair HC, Schlesinger PH 1992 The mechanism of osteoclast acidification. In: Rifkin BR, Gay CV (eds.) Biology and Physiology of the Osteoclast. CRC Press, Boca Raton, FL, USA, pp. 259–287.
6. Li YP, Chen W, Liang Y, Li E, Stashenko P 1999 Atp6i-deficient mice exhibit severe osteopetrosis due to loss of osteoclast-mediated extracellular acidification. Nat Genet **23:** 447–451.
7. Li YP, Chen W, Stashenko P 1996 Molecular cloning and characterization of a putative novel human osteoclast-specific 116-kDa vacuolar proton pump subunit. Biochem Biophys Res Commun **218:**813–821.
8. Sambrook J, Fritsch EF, Maniatis T 1989 Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York, NY, USA.
9. Chen W, Li YP 1998 Generation of mouse osteoclastogenic cell lines immortalized with SV40 large T antigen. J Bone Miner Res **13:**1112–1123.

10. Li YP, Chen W, Stashenko P 1995 Characterization of a silencer element in the first exon of the human osteocalcin gene. Nucleic Acids Res **23:**5064–5072.
11. Feng ZM, Li YP, Chen CL 1989 Analysis of the 5′-flanking regions of rat inhibin alpha- and beta-B-subunit genes suggests two different regulatory mechanisms. Mol Endocrinol **3:**1914–1925.
12. Li YP, Chen W 1999 Characterization of mouse cathepsin K gene, the gene promoter, and the gene expression. J Bone Miner Res **14:**487–499.
13. Wilkinson DG 1993 In situ hybridization. In: Stern CD, Holland PWH (eds.) Essential Developmental Biology: A Practical Approach. IRL Press, Oxford, UK, pp. 257–274.
14. Grigoriadis AE, Wang ZQ, Cecchini MG, Hofstetter W, Felix R, Fleisch HA, Wagner EF 1994 c-Fos: A key regulator of osteoclast-macrophage lineage determination and bone remodeling. Science **266:**443–448.
15. Iotsova V, Caamano J, Loy J, Yang Y, Lewin A, Bravo R 1997 Osteopetrosis in mice lacking NF-kappaB1 and NF-kappaB2. Nat Med **3:**1285–1289.
16. Tondravi MM, McKercher SR, Anderson K, Erdmann JM, Quiroz M, Maki R, Teitelbaum SL 1997 Osteopetrosis in mice lacking haematopoietic transcription factor PU1. Nature **386:** 81–84.
17. Mattsson JP, Schlesinger PH, Keeling DJ, Teitelbaum SL, Stone DK, Xie XS 1994 Isolation and reconstitution of a vacuolar-type proton pump of osteoclast membranes. J Biol Chem **269:**24979–24982.
18. Fischer A, Griscelli C, Friedrich W, Kubanek B, Levinsky R, Morgan G, Vossen J, Wagemaker G, Landais P 1986 Bone-marrow transplantation for immunodeficiencies and osteopetrosis: European survey, 1968–1985. Lancet **2:**1080–1084.
19. Gerritsen EJ, Vossen JM, van Loo IH, Hermans J, Helfrich MH, Griscelli C, Fischer A 1994 Autosomal recessive osteopetrosis: Variability of findings at diagnosis and during the natural course. Pediatrics **93:**247–253.
20. Toyomura T, Oka T, Yamaguchi C, Wada Y, Futai M 2000 Three subunit a isoforms of mouse vacuolar H(+)-ATPase. Preferential expression of the a3 isoform during osteoclast differentiation. J Biol Chem **275:**8760–8765.
21. Nishi T, Forgac M 2000 Molecular cloning and expression of three isoforms of the 100-kDa subunit of the mouse vacuolar proton-translocating ATPase. J Biol Chem **275:**6824–6830.
22. Mattsson JP, Li X, Peng SB, Nilsson F, Andersen P, Lundberg LG, Stone DK, Keeling DJ 2000 Properties of three isoforms of the 116-kDa subunit of vacuolar H+-ATPase from a single vertebrate species. Cloning, gene expression and protein characterization of functionally distinct isoforms in *Gallus gallus*. Eur J Biochem **267:**4115–4126.
23. Gum R, Lengyel E, Juarez J, Chen JH, Sato H, Seiki M, Boyd D 1996 Stimulation of 92-kDa gelatinase B promoter activity by ras is mitogen-activated protein kinase kinase 1-independent and requires multiple transcription factor binding sites including closely spaced PEA3/ets and AP-1 sequences. J Biol Chem **271:**10672–10680.
24. Wucherpfennig AL, Li YP, Stetler-Stevenson WG, Rosenberg AE, Stashenko P 1994 Expression of 92 kD type IV collagenase/gelatinase B in human osteoclasts. J Bone Miner Res **9:**549–556.
25. Manolagas SC, Bellido T, Jilka RL 1995 New insights into the cellular, biochemical, and molecular basis of postmenopausal and senile osteoporosis: Roles of IL-6 and gp130. Int J Immunopharmacol **17:**109–116.
26. Udagawa N, Takahashi N, Katagiri T, Tamura T, Wada S, Findlay DM, Martin TJ, Hirota H, Taga T, Kishimoto T 1995 Interleukin (IL)-6 induction of osteoclast differentiation depends on IL-6 receptors expressed on osteoblastic cells but not on osteoclast progenitors. J Exp Med **182:**1461–1468.

Address reprint requests to:
Yi-Ping Li
The Forsyth Institute
140 Fenway
Boston, MA 02115, USA

Received in original form March 29, 2000; in revised form October 13, 2000; accepted December 11, 2000.