Molecular cloning of the gene encoding the mouse parathyroid hormone/parathyroid hormone-related peptide receptor

[G protein-coupled receptors/(G+C)-rich promoters/polyadenylylation signals/growth hormone releasing factor receptor]

KIMBERLY A. MCCUAIG, JOHN C. CLARKE, AND JOHN H. WHITE*

Department of Physiology, McGill University, McIntyre Medical Sciences Building, 3655 Drummond Street, Montreal, P.Q., Canada H3G 1Y6

Communicated by Charles C. Richardson, February 14, 1994 (received for review November 9, 1993)

ABSTRACT The parathyroid hormone/parathyroid hormone-related peptide receptor (PTHR) is a G-protein-coupled receptor containing seven predicted transmembrane domains. We have isolated and characterized recombinant bacteriophage *AEMBL3* genomic clones containing the mouse PTHR gene, including 10 kilobases of the promoter region. The gene spans >32 kilobases and is divided into 15 exons, 8 of which contain the transmembrane domains. The PTHR exons containing the predicted membrane-spanning domains are heterogeneous in length and three of the exon-intron boundaries fall within putative transmembrane sequences, suggesting that the exons did not arise from duplication events. This arrangement is closely related to that of the growth hormone releasing factor receptor gene, particularly in the transmembrane region, providing strong evidence that the two genes evolved from a common precursor. Transcription is initiated principally at a series of sites over a 15-base-pair region. The proximal promoter region is highly (G+C)-rich and lacks an apparent TATA box or initiator element homologies but does contain CCGCCC motifs. The presumptive amino acid sequence of the encoded receptor is 99%, 91%, and 76% identical to those of the rat, human, and opossum receptors, respectively. There is no consensus polyadenylylation signal in the 3' untranslated region. The poly(A) tail of the PTHR transcript begins 32 bases downstream of a 35-base-long A-rich sequence, suggesting that this region directs polyadenylylation.

The parathyroid hormone/parathyroid hormone-related peptide receptor (PTHR) is bound specifically by a conserved 34-amino acid region present in both parathyroid hormone (PTH) and PTH-related peptide (PTHrP). PTH regulates calcium and phosphate metabolism by binding to receptors expressed in kidney and bone (1-5). PTHrP was first identified as a major cause of malignancy-associated hypercalcemia (6, 7); however, its normal physiological role remains largely unknown. Whereas PTH expression is limited to the parathyroid, PTHrP is expressed in a wide variety of normal and malignant tissues and appears to act mainly in a para- or autocrine manner (8-11). The PTHR is a G-protein-coupled receptor containing seven predicted transmembrane domains (refs. 1-3 and references therein). Binding of ligand to the PTHR stimulates cAMP production, raises intracellular calcium, and increases levels of inositol 1,4,5-trisphosphate (2).

The G-protein-coupled family of receptors is vast and includes receptors for peptide hormones, >100 odorants, neurotransmitters, and a number of other regulatory factors (12). Based on similarities between ligands and receptors (13, 14), the PTHR belongs to a subfamily that includes receptors for growth hormone releasing factor, vasoactive intestinal peptide, calcitonin, secretin, glucagon-like peptide, and glucagon. Genes for several mammalian adrenergic and serotonin receptors have been cloned and are intronless (15-19). Although the luteinizing hormone receptor contains 11 exons, the transmembrane and cytoplasmic regions of the protein are encoded by a single exon (20). Here, we have cloned the entire PTHR gene and show that it contains multiple exons, 8 of which encode the transmembrane domains. The exon-intron boundaries are very similar to those of the mouse growth hormone releasing factor receptor (GHFR) gene (13). The proximal promoter is (G+C)-rich and contains several putative binding sites for the transcription factor SpI. Interestingly, polyadenylylation is initiated downstream of an unusual A-rich sequence in a region that lacks a consensus polyadenylylation signal.[†]

MATERIALS AND METHODS

Library Screening. A λ EMBL3 genomic library (Clontech), from adult male BALB/c liver DNA, was screened using nick-translated probes corresponding to the entire rat PTHR cDNA or to 115 bp of the 5' untranslated sequence and signal sequence. Filters (S&S Nytran) were screened in 5× SSPE [1× SSPE = 10 mM sodium phosphate, pH 7.7/180 mM NaCl/1 mM EDTA], 5× Denhardt's solution, 40% deionized formamide, 1% SDS, 10% dextran sulfate, and 100 μ g of denatured salmon sperm DNA per ml at 42°C for 18 hr. The membrane was washed to a final stringency in 0.1% SSC [20× SSC = 0.3 M sodium citrate, pH 7.0/3 M NaCl] and 0.1% SDS at 55°C for 30 min. Positive clones were purified by three rounds of screening with the same probe.

DNA Sequencing. Phage DNA was prepared by polyethylene glycol precipitation and purification from a cesium chloride gradient (21). Fragments containing exons, determined by Southern blotting using Hybond-N membranes (Amersham) under conditions described above, were subcloned into pBluescript SK+ (Stratagene) and sequenced by the dideoxy chain-termination method using primers corresponding to T3 or T7 promoters or to rat or mouse DNA sequences.

S1 Nuclease Assays. Probe was prepared by insertion of a 560-bp Xho I-Apa I fragment (see Fig. 3) in Bluescript SK+ (Stratagene). The recombinant plasmid (0.5 μ g) was digested with Xho I, purified, and incubated in 40 mM Tris·HCl, pH 8.0/10 mM dithiothreitol/4 mM spermidine/10 mM NaCl/50 μ g of bovine serum albumin per ml/10 mM MgCl₂/0.5 mM (each) ATP, GTP, and UTP/0.01 mM CTP/50 μ Ci of [α -³²P]CTP (1 Ci = 37 GBq)/20 units of RNasin (Promega)/30 units of T7 RNA polymerase (Pharmacia) at 37°C for 60 min. DNase I (10 units, GIBCO) was then added to digest the DNA template. Following phenol extraction and ethanol precipitation, 50,000 cpm of probe was hybridized to 10 μ g of total

*To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: GHFR, growth hormone releasing factor receptor; IL-6, interleukin 6; NFIL-6, nuclear factor IL-6; PTH, parathyroid hormone; PTHrP, PTH-related peptide; PTHR, PTH/PTHrP receptor; RT-PCR, reverse transcriptase-polymerase chain reaction.

[†]The sequence reported in this paper has been deposited in the GenBank data base (accession no. L28108).

mouse kidney RNA, incubated overnight at 55°C in 30 μ l of 40 mM Pipes, pH 6.4/1 mM EDTA/0.4 M NaCl/80% formamide, and then diluted in 300 μ l of 50 mM sodium acetate, pH 5.0/4.5 mM ZnSO4/20 μ g of salmon sperm DNA per ml, and S1 nuclease (Pharmacia) was added as indicated. After 60 min at 37°C, the reaction was terminated by adding 80 μ l of 4 mM ammonium acetate/50 mM EDTA/50 μ g of tRNA per ml and ethanol precipitated. Products were heated in 50% formamide at 90°C for 3 min prior to loading on a 6% polyacrylamide sequencing gel.

Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR). RT-PCR was performed essentially as described (22) with total RNA from mouse kidney using the primers 5'-

Proc. Natl. Acad. Sci. USA 91 (1994)

GACTCGAGTCGACGGTACCT₁₇-3' and 5'-AACCACTG-GCGTTGACTTC-3', which recognize poly(A) and cytoplasmic domain sequences, respectively. Amplified products (30 cycles: 94°C, 1 min; 46°C, 90 sec; 72°C, 1 min) were digested with *Kpn* I and with *Pvu*II, which recognizes a sequence in the 3' untranslated region, and inserted into BlueScript SK+ for sequencing.

RESULTS AND DISCUSSION

Isolation and Sequencing of Genomic Clones Encoding the Mouse PTHR Gene. One million plaques of a BALB/c mouse λ EMBL3 genomic library were screened with a nick-translated probe containing the entire rat PTHR cDNA, and two clones,



FIG. 1. Structure of the mouse PTHR. (A) The λ EMBL3 clones $\lambda 1$, $\lambda 3R$, and $\lambda 5R$ containing PTHR exonic sequence and 10 kb of promoter region are shown above the PTHR gene structure. Exons U and SS containing untranslated sequence and the putative signal sequence, respectively, are represented by stippled bars. The four exons containing the extracellular sequence of the receptor (E1–E4) are in white. Exons containing the transmembrane region (T1–T7b) are in black and the exon containing the C-terminal cytoplasmic domain and 3' untranslated sequence (C) is represented by the striped bar. The *Bam*HI (B) and *Sac* I (S) restriction sites are also indicated. The predicted cDNA (below) shows the position of exon-intron boundaries within the coding sequence and flanking regions. The exons are represented as above. Positions of the transmembrane domains are indicated below by the horizontal white bars (see also Fig. 4). The length of each exon is shown on the right. (B) Positions of the splice donor and acceptor sites for each intron along with its length (or estimate) are listed on the right. (C) Alignment of the 3' untranslated sequences of the mouse and human (above) and mouse and rat (below). Translational stop codons (TGA) are indicated in bold with fine over- or underlines, and the T residue is assigned the position +1. The A-rich sequence, which is strongly conserved between the mouse and human and weakly conserved between the mouse and rat, is indicated by bold over- and underlines. The position of the start of the poly(A) tail in the mouse is indicated by the arrowhead. (D) Northern blot analysis was performed as described (22) using a rat cDNA probe and 30 μ g of total RNA isolated from the rat osteoblast-like osteosarcoma cell line ROS17/2.8 (RAT) and mouse kidney (MOUSE).







FIG. 2. Sequence of the mouse PTHR gene promoter. (A) Determination of transcriptional initiation sites by S1 nuclease analysis. The 5' end of the sequencing primer is 42 bp downstream of the 5' end of the S1 probe. Sequences shown are those corresponding to the start sites. (B) Sequence of 1698 bp of the mouse PTHR promoter region. Transcriptional start sites determined by S1 nuclease analysis are indicated by overhead arrows, with the thickest line corresponding to the most frequently used site (+1). A minor site detected by S1 nuclease analysis is indicated by the asterisk. The start sites identified by primer extension analysis (23) are indicated by the arrowheads underneath. Selected restriction sites are lightly underlined and indicated. The Apa I (+179) and Xho I (-485) sites used to generate the probe for S1 nuclease are shown. Potential binding sites for transcriptional regulators are heavily underlined and the names are indicated. NFIL-6, nuclear factor interleukin 6.

B

 λ 3R and λ 5R (Fig. 1A), were isolated and characterized. Southern blotting analysis of several restriction digests indicates that these clones represent contiguous fragments of genomic DNA (data not shown). The λ 5R clone contains six exons, of which five contain sequences encoding the putative fourth to seventh transmembrane domains (T4b-T7b; Fig. 1A). A sixth exon (C) encodes the cytoplasmic domain and contains 3' untranslated sequence. The λ 3R clone contains seven exons, of which three (T1-T3/4a) contain sequences encoding the first three transmembrane domains and four (E1-E4) encode most of the extracellular portion of PTHR but not the signal sequence. The library was then screened with a nick-translated 115-bp BamHI-HaeII fragment homologous to the 5' end of the rat cDNA (2). The clone $\lambda 1$ was isolated (Fig. 1A), and two exons (U and SS) were mapped. These exons contain 5' untranslated sequence and putative signal sequence, respectively, which are similar to the 5' end of the rat cDNA (data not shown, but see Fig. 4). The signal sequence exon is separated from the E1 exon by an intron of at least 15 kb as indicated by Southern blotting analysis of mouse genomic DNA (data not shown). In total, the exons of the mouse PTHR gene span at least 32 kb (see Fig. 1).

Determination of Transcriptional Initiation Sites and Analysis of the PTHR Promoter. Transcriptional initiation sites were determined by S1 nuclease mapping of mouse kidney RNA using a continuously labeled RNA probe (see *Materials* and Methods). The probe is homologous to the PTHR gene from an Apa I site centered 45 bp upstream of the 3' boundary of first exon (U) to the Xho I site located 560 bp upstream of the Apa I site (see Fig. 2B). A series of start sites were detected, clustered over a 15-bp stretch in a region that is highly (G+C)-rich (Fig. 2). A second much weaker site was reproducibly detected ≈ 80 bp upstream (Fig. 2A and indicated by the asterisk in Fig. 2B). We also performed primer extension in this region using two different primers, one centered over the *Bam*HI site (see Fig. 2B) and another hybridizing to sequences 109 bp downstream (data not shown). In both cases, the major extension products stop within 1-3 nucleotides of the principal site determined by S1 nuclease analysis (Fig. 2B; data not shown). There are no TATA homologies or initiator elements (24) in this region; however, there are three sites that conform to the extended homology recognized by the SpI transcription factor (Fig. 2B) along with several other potential SpI binding sites containing a single nonconsensus nucleotide.

The proximal promoter region also contains two sites (Fig. 2A), centered 160 and 270 bp upstream of the principal transcriptional initiation site, which correspond to the T(G/T)NNGNAA(G/T) motif recognized by the activator NFIL-6 (25). NFIL-6, which is a member of the C/EBP family of transcription factors, was found to induce expression of the interleukin 6 (IL-6) gene in response to interleukin 1 (26). Interestingly, IL-6 is secreted by stromal cell precursors of osteoblasts and mature osteoblasts and has been shown to be an activator of bone resorption by stimulating osteoclast formation (27), raising the possibility that NFIL-6 may regulate several pathways that lead to stimulation of bone resorption. There are also several potential binding sites for members of the ets family of transcriptional regulators (Fig. 2B) that recognize sequence motifs with C/AGGAA cores (28, 29). The ets family contains several members that are expressed in a wide variety of tissues, including kidney (29, 30). Phorbol ester and factors that increase intracellular

	λ	
PTHR	MGTARIAPS.LALL.LCCPVLSSAYALVDADDVFTKEEQIFLLHRAQAQCDK	50
GRFR	MDGLMMATRILCLLSLC.GV.	19
	1	
PTHR	LLKEVLHTAAN IMESOKGNTPASTSGEPREKAPGKFYPESKEWKDVPTG	100
GRFR	TLGHLHLECOFT TOLEDD . ELACLOAA . EGTNETSIG	54
PTHR	SRRRGRPCLPENDNIVCHPLGAPGEVVAVPCPDYIYDTNHK, GHAYBRCDR	150
GRFR	CPGTIDGLICIPPTGSGONVSLPCPEFFSHFGSDTGFVKBDCTT	98
	· · · · · · · · · · · · · · · · · · ·	
PTHR	NGSWEVV. PGHNRTWANYSECLKFMTNETREEVEDRLGMIYTVOYEMELA	200
GRFR	TG.WSNPFPPYPVACPVPLELLTHERSYFSTVKIITTCHSISIV	142
	_	
PTHR	SLTVAVLILAYFTCLECTREVIENHMFLSTMLRAAS IFVEDAVLYSGFTL	250
GRFR	ALCVAIAILVAL RELECTION THE TOLEAT FILEASAVELED AAIFOODST	192
	·	
PTHR	DEAERLTEELHIIAQVPPPPAAAAVGYAGCKVAVTFFLYFLATNYYWIL	300
GRFR	DHCSMST	221
PTHR	VEGLYLHSLIFMAFFSERKYLNGFTIFCNGLPAVFVAVNVGVRATLANTG	350
GRFR	ABAVYLSCILLASTSPRSIPAFIWLVLAGINGLPVLCTGTWVGCELAFED	271
PTHR	CHOLS. BGHKKWIIQVFILADVVLWILFIMIIRVLATKLRETNAGRCDTR	400
GRPR	CHULDNUSP CHWIIKGPIVLEWGVERGLEILICILLRELL. EP AQUGLHTR	321
	1 77 1	
0700		453
CDFD		101
GUL V		300
	IIV	
PTHR	OFFVALLYCHCHCOARTRESUSPICIALDER PARSOSSYSYCEWUSH	502
GRFR	OF IVAVINCELICONTREES BRENGHDPELIPABETCTENTTPPRSRIEVI.	419
	T	
	•	
PTHR	TS VTNVGPRAGLSLPLSPRLLPATTNGHSQLPGHAKPGAPAIENETIPVTM	553
GRFR	TSEC	423
PTHR	TVPKDDGFLNGSCSGLDEEASGSARPPPLLQEEWETVM	591

FIG. 3. Comparison of the amino acid sequence identity and exon-intron boundaries of the mouse PTHR and GHFR genes. Similar or identical amino acids are indicated in bold type and the exon-intron boundaries are indicated by vertical bars. Predicted transmembrane domains (I-VII) and three hydrophobic regions (A-C) as designated by Abou-Samra *et al.* (2) are overlined.

cAMP are known to down-regulate expression of the PTHR (31, 32). It is notable that the proximal promoter region lacks any potential AP-1 or CRE sites, elements that mediate up-regulation by these agents.

Mapping the 3' End of the Mouse PTHR Transcript. The PTHR gene contains 2041 bp of exonic sequence up to the TGA of exon C (Fig. 1A). The PTHR mRNA detected in extracts of mouse kidney tissue with a highly similar rat cDNA probe is ≈ 2.2 kb in length and comigrates with a band detected in extracts of ROS 17/2.8 cells, a rat osteoblast-like osteosarcoma line (Fig. 1D). No minor mouse kidney transcripts were detected upon prolonged exposure of the blot, indicating that the 2.2-kb band represents the predominant transcript. However, we cannot rule out the possibility that other mRNA species are expressed within different tissues (23, 33). Taken together the above results suggest that the 3' untranslated region of the mouse PTHR mRNA is ≈ 160 bp in length. There is no consensus AAUAAA or related sequences that reportedly serve as polyadenylylation signals (34) in the first 448 bp downstream of the TGA translation stop codon (Fig. 1C, and data not shown). The 3' untranslated sequence of the mouse apparently diverges from that of the rat 115 bp downstream of the TGA codon at a series of A residues (see Fig. 1C). Though the A-rich sequence is not fully conserved in the rat, a similar sequence is found in the human cDNA (Fig. 1C). We have sequenced the products amplified by RT-PCR (35) of the 3' untranslated region and mapped the beginning of the poly(A) tail of the mouse PTHR transcript to a site 32 bp downstream of the A-rich region, or 166 bp downstream of the TGA codon (Fig. 1C, and data not shown). This gives a total length of the mouse PTHR mRNA minus the poly(A) tail of 2207 bases, in very good agreement with the results of Northern analysis (Fig. 1D). Given its position relative to the poly(A) tail, our data suggest that the A-rich sequence can replace the AAUAAA consensus in

Proc. Natl. Acad. Sci. USA 91 (1994)

Mouse	1A MGTARIAPSL	ALLLCCPVLS	SAYADVDADD	VFTKEEQIFL	50 LHRAQAQCDK
Rat	· · λ · · · · · · · ·				· · · · · · · · · · · · · · · · · · ·
Human Opossum	G.			.M	E.
	1				
Mouse	51 LLKEVLHTAA	NIMESDECHT	PASTSCKPRK	EKAPGKEYPE	100 SKENKDVPTG
Rat				s	
Human	RQRP.	S	S	DSL	D.EA
opossum	RRVE-	ELAA.D	-A.R.AKIK.		AL. SKE. SUK
Manaa	101	PHONIT	CARCELLAND		150
Rat	SRRRGRPCLP	EWDNIVCWPL	GAPGEVVAVP	CPDITIDENH	KGHAIRRCDR
Human	Y	H.L			
Opossum	LQDGF	· · · · · · · · · · · A	.vĸ	•••••	RS
	151			I	200
Mouse	NGSWEVVPGH	NRTWANYSEC	LKFMTNETRE	REVFDRLGMI	YTVGYSMSLA
Human	L		VL		·····v
Opossum	N		VL	••••	IG
	201	1		II	250
Mouse	SLTVAVLILA	YFRRLHCTRN	YIHMHMFLSF	MLRAASIFVK	DAVLYSGFTL
Rat	•••••	•••••	•••••		••••••
Opossum	G		L.V	VI.	
•			10		
Mouse	251 DEARRITEER	LHITAOVPPP	PAAAAVGYAG	CRVAVTEELY	FLATNY WIL
Rat					
Human	••••••	.RAA	TA	•••••	••••
opossum		. KAP IEP	.P.DKA.FV.		
					1
••	<u>301 III</u>			IV	35
Mouse Rat	301 III VEGLYLHSLI	FMAFFSEKKY	LWGFTIFGWG	IV LPAVFVAVWV	35 GVRATLANTG
Mouse Rat Human	301 III VEGLYLHSLI	FMAFFSEKKY	LWGFT IFGWG	IV LPAVFVAVWV	350 GVRATLANTG S
Mouse Rat Human Opossum	301 III VEGLYLHSLI	FMAFFSEKKY	LWGFT IFGWG		350 GVRATLANTG S
Mouse Rat Human Opossum	301 JJJ VEGLYLHSLI	FMAFFSEKKY	LWGFT IFGWG		350 GVRATLANTG SE TE 400
Mouse Rat Human Opossum Mouse	301 III VEGLYLHSLI	FMAFFSEKKY	LWGFTIFGWG	IV LPAVFVAVNV	350 GVRATLANTG SE TE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human	301 III VEGLYLHSLI 	FMAFFSEKKY WI IQVPILAS	LWGFT IFGWG	IV LPAVFVAVWV IIIRVLATKLR	350 GVRATLANTC SE TE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum	301 III VEGLYLHSLI 	FMAFFSEKKY WIIQVPILAS	LWGFT IFGWG	IV LPAVFVAVWV 	350 GVRATLANTG SE TE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum	301 III VEGLYLHSLI 	FMAFFSEKKY WIIQVPILAS	LWGFT IFGWG 	IV LPAVFVAVWV IIIRVLATKLR	350 GVRATLANTG SE TE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse	301 III VEGLYLHSLI 	FMAFFSEKKY WI IQVPILAS	LWGFT IFGWG VL VVLNFILFIN I.V YTVFMALPYT	IV LPAVFVAVWV IIRVLATKLR V EVSGTLWQIQ	350 GVRATLANTG SE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human	301 III VEGLYLHSLI 	FMAFFSEKKY WI IQVPILAS 	LWGFTIFGWG 	IV LPAVFVAVWV IIRVLATKLR V EVSGTLMQIQ	
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum	301 III VEGLYLHSLI 351 CWDLSSGHKK 401 QQYRKLLRST K	FMAFFSEKKY WI IQVPILAS 	LWGFT IFGWG V VVLNF1LFIN I YTVFMALPYT IT	IV LPAVFVAVWV IIRVLATKLR V EVSGTLMQIQ V IV.	35 GVRATLANTG SE TE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum Opossum	301 III VEGLYLHSLI 351 CWDLSSGHKK 	FMAFFSEKKY WI IQVPILAS VZ LVLVPLFGVH M.	LWGFT IFGWG 	IV LPAVFVAVWV IIRVLATKLR V EVSGTLWQIQ V.	35 GVRATLANTG SE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum Opossum Mouse Mouse	301 III VEGLYLHSLI 351 CWDLSSGHKK 	FMAFFSEKKY WI IQVPILAS VZ LVLVPLFGVH	LWGFT IFGWG 	IV LPAVFVAVWV 	
Mouse Rat Human Opossum Mouse Rat Human Opossum Opossum Mouse Rat	301 222 VEGLYLHSLI 351 CWDLSSCHKK N. QQYRKLLRST 	FMAFFSEKKY WI IQVPILAS 	LWGPT IFGWG 	IV LPAVFVAVWV IIRVLATKLR V EVSGTLWQIQ V I.V. LDFKRKARSG	35 GVRATLANTC SE TE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Coossum	301 III VEGLYLHSLI 351 CWDLSSCHKK 401 QQYRKLLRST 	FMAFFSEKKY WIIQVPILAS A VI LVLVPLFGVH M. FCNGEVQAEI	LWGPT IFGWG 	IV LPAVFVAVWV 	35 GVRATLANTC S TE 400 ETNAGRCDTR 450 MHYEMLFNSF 500 SSSYSYGEMV
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum	301 III VEGLYLHSLI 351 CWDLSSGHKK 401 QYRKLLRST K	FMAFFSEKKY WIIQVPILAS A VI LVLVPLFGVH M. FCNGEVQAEI	LWGPT IFGWG 	IV LPAVFVAVWV IIRVLATKLR V EVSGTLWQIQ V. LDFKRKARSG	
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum	301 III VEGLYLHSLI 	FMAFFSEKKY WI IQVPILAS A VI LVLVPLFGVH FCNGEVQAEI	V V V V V VINFILFIN I.V. VINFILFIN I.V. VINFILFIN I.V. VINFILFIN RKSWSRWTLA K. K. K. RLIP-ATT VINFILFIN	IV LPAVFVAVWV IIRVLATKLR V EVSGTLWQIQ V. LDFKRKARSG NGHSQLPCHA	350 GVRATLANTG SE TE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat	301 III VEGLYLHSLI 	FMAFFSEKKY WI IQVPILAS VI LVLVPLFGVH M. FCNGEVQAEI RAGLSLPLSP	V V V V VVLNFILFIN I I.V. I YTYFMALPYT I I.T. T RKSWSRWTLA K. K. RLLP-ATT	IV LPAVFVAVWV IIRVLATKLR V EVSGTLMQIQ LDFKRKARSG NGHSQLPCHA	35 GVRATLANTC S TE 400 ETNAGRCDTR MHYEMLFNSF 500 SSSYSYCPMV KPGAPAIEN-
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human	301 III VEGLYLHSLI 351 CWDLSSCHKK 	FMAFFSEKKY WI IQVPILAS VI LVLVPLFGVH FCNGEVQAEI	Image: Number of the second	IV LPAVFVAVWV IIRVLATKLR V EVSGTLMQIQ LDFKRKARSG NGHSQLPCHA 	
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum	301 222 VEGLYLHSLI 351 CWDLSSCHRK 401 QQYRKLLRST CGFFVAIIYC CGFFVAIIYC 501 SHTSVTNVCP	FMAFFSEKKY WI IQVPILAS A VZ LVLVPLFGVH FCNGEVQAE I RAGLSLPLSP .G.A.S.	V V V L V L V L V L V L V L V L V L V L V L V L I I I I RKSWSRWTLA K RLLP-ATT	IV LPAVFVAVWV IIRVLATKLR V EVSGTLWQIQ LDFKRKARSG NGHSQLPGHA 	350 GVRATLANTG S
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum	301 III VEGLYLHSLI 351 CWDLSSCHKK N. N. 401 QQYRKLLRST K. K. K. 	FMAFFSEKKY WIIQVPILAS A VI LVLVPLFGVH M. FCNGEVQAEI RAGLSLPLSP .G. A. S.	V V V V VVLNFILFIN I I.V. I YTVFMALPYT I I.T. T RKSWSRWTLA K. RLLP-ATT T.	IV LPAVFVAVWV 	35 GVRATLANTC S TE 400 ETNAGRCDTR
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum	301 III VEGLYLHSLI 351 CWDLSSGHKK 401 QQYRKLLRST 	FMAFFSEKKY WI IQVPILAS A VI LVLVPLFGVH M. FCNGEVQAEI RAGLSLPLSP G. A. S. KDDGFLNGSC	LWGPT IFGWG 	IV LPAVFVAVWV IIRVLATKLR V EVSGTLWQIQ V. LDFKRKARSG NGHSQLPCHA NGHSQLPCHA ARPPPLLQEE	35 GVRATLANTG S TE 400 ETNAGRCDTR 450 MHYEMLFNSF 500 SSSYSYGPMV 550 SSSYSGPMV T 550 SSSYSGPMV 550 SSSYSGPMV 550 KPGAPAIEM
Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human Opossum Mouse Rat Human	301 III VEGLYLHSLI 	FMAFFSEKKY WIIQVPILAS A VI LVLVPLFGVH FCNGEVQAEI RAGLSLPLSP .G.A.S. KDDGFLNGSC	LWGPTIFGWG 	IV LPAVFVAVWV 	350 GVRATLANTC S

FIG. 4. Comparison of the amino acid sequences of the mouse, rat, human, and opossum PTHR proteins. Predicted transmembrane domains (I-VII) and three hydrophobic regions (A-C) as designated by Abou-Samra *et al.* (2) are overlined.

serving as a polyadenylylation signal. We note that the sequence AUUAAA, which may serve as a polyadenylylation signal (34), is found 155 bp downstream of the TGA codon in the rat PTHR sequence (Fig. 1C).

Analysis of PTHR Gene Structure. Six of the introns separating coding sequence lie between codons (phase 0), whereas four are of phase 1, and three are of phase 2 (Fig. 1B). The introns separating the transmembrane domains are of all three phases, and three of the introns fall within putative membranespanning regions. The exons are heterogeneous in length. In addition, there is no evident positioning of the exon-intron boundaries within this region with respect to the beginning or end of predicted membrane-spanning regions (Figs. 1 A and B and 3; see also Fig. 4). Taken together, this suggests that these exons did not arise through duplication events. Exons E1-E4, which encode extracellular sequence, are of similar length and are separated from each other by phase 1 introns (Fig. 1B), raising the possibility that they arose by duplication events. However, no sequence identity was detected at the amino acid level between the E exons, and DNA sequence analyses did not reveal significantly more identity between pairs of extra-

Biochemistry: McCuaig et al.

cellular exons than between a given extracellular exon and either a given transmembrane exon, the cytoplasmic exon, or random DNA (data not shown).

The structure of the PTHR gene is very similar to that of the related mouse GHFR (Fig. 3; ref. 13). There is some divergence in the exons encoding the extracellular domain. For example, the PTHR gene contains an additional exon in this region. However, the positioning of introns within the coding sequence is particularly well conserved in the transmembrane region, where the two receptors share the greatest homology (Fig. 3). This provides strong evidence that the PTHR and GHFR genes diverged from a common ancestor. It is likely that genes encoding other members of the subfamily (vasoactive intestinal peptide, calcitonin, secretin, glucagon-like peptide, and glucagon) will share similar structures. The multiple introns of the PTHR gene raise the possibility that different receptor forms could be generated in different tissues by alternative splicing. A major PTHR transcript estimated at 2.3 kb has been detected in a number of tissues in the rat (31). Minor transcripts have also been detected (23, 33), although their functional significance remains to be determined.

Analysis of the Predicted PTHR Amino Acid Sequence. The sequence of the mouse PTHR translational initiation site is identical to that of the rat (2) and contains the sequence GCG ATG G (data not shown), which conforms closely to the consensus A/GCC ATG G first reported by Kozak (36). The predicted amino acid sequence of the mouse PTHR is 99%, 91%, and 76% similar to rat, human, and opossum sequences, respectively. The sequence of the mouse PTHR protein differs from that of the rat in 6 of 591 positions (Fig. 4). There are two changes (Thr-3 \rightarrow Ala-3 and Pro-84 \rightarrow Ser-84) in the extracellular region, which do not affect the potential glycosylation sites (2), and four changes in the C-terminal cytoplasmic domain (Ile-544 → Thr-544, Gln-546 → Thr-546, Ile-549 \rightarrow Lys-549, Thr-554 \rightarrow Ala-554). The transmembrane regions of the two proteins are 100% conserved.

Summary. The structure of the mouse PTHR gene is very similar to that of the related mouse GHFR, providing strong genetic evidence that they evolved from a common precursor. The proximal promoter region is (G+C)-rich and lacks either TATA box or initiator element homologies. The 3' end of the gene is unusual in that it lacks a consensus polyadenylylation signal upstream of the poly(A) tail. Our results strongly suggest that an unusual A-rich sequence serves as a polyadenylylation signal.

We are grateful to Drs. D. Goltzman and G. Hendy (Dept. of Physiology, McGill University) for the rat PTHR cDNA and thank Drs. D. Goltzman, J. Orlowski, and S. Mader (McGill) and Dr. A. Stoltzfus (Canadian Institute for Advanced Studies, Halifax) for critically reading the manuscript. This work was supported by the Medical Research Council of Canada (Grant MT-11704) and the Fonds de la Recherche en Santé du Québec. J.H.W. is a Chercheur-Boursier of the Fonds de la Recherche en Santé du Québec.

- Jüppner, H., Abou-Samra, A.-B., Freeman, M. W., Kong, 1. X.-F., Schipani, E., Richards, J., Kolakowski, L. F., Jr., Hock, J., Potts, J. T., Jr., Kronenberg, H. M. & Segre, G. (1991) Science 254, 1024-1026.
- 2. Abou-Samra, A.-B., Jüppner, H., Force, T., Freeman, M. W., Kong, X.-F., Schipani, E., Urena, P., Richards, J., Bonventre, J., Potts, J. T., Jr., Kronenberg, H. M. & Segre, G. (1992) Proc. Natl. Acad. Sci. USA 89, 2732-2736.
- Schipani, E., Karga, H., Karaplis, A. C., Hellman, P., Xie, L.-Y., Potts, J. T., Jr., Kronenberg, H. M., Segre, G., Abou-Samra, A.-B. & Jüppner, H. (1992) Endocrinology 132, 2157-2165.
- 4. Rosenblatt, M., Kronenberg, H. & Potts, J. T., Jr. (1989) in Endocrinology, ed. DeGroot, L. J. (Saunders, Philadelphia), pp. 848-891.
- 5. Kronenberg, H. (1993) in Primer on Metabolic Bone Diseases

and Disorders of Mineral Metabolism, ed. Favus, M. J. (Raven, New York), 2nd Ed., pp. 58-60.

- 6. Hendy, G. N. & Goltzman, D. (1992) in Parathyroid Hormone-Related Peptide: Normal Physiology and its Role in Cancer, eds. Halloran, B. P. & Nissenson, R. A. (CRC, Boca Raton, FL), pp. 25-55.
- 7. Strewler, G. D. & Nissenson, R. A. (1993) in Primer on Metabolic Bone Diseases and Disorders of Mineral Metabolism, ed. Favus, M. J. (Raven, New York), 2nd Ed., pp. 61-63.
- Thiede, M. A. & Rodan, G. A. (1988) Science 242, 278-280. Ikeda, K., Mangin, M., Dreyer, B. E., Webb, A. C., Posillico, J. T., Stewart, A. F., Bander, N. H., Weir, E. C., Insogna, K. L. & Broadus, A. E. (1988) J. Clin. Invest. 81, 2010–2014.
- 10. Moniz, C., Burton, P. B. J., Malik, A. N., Dixit, M., Banga, J. P., Nicolaides, K., Quirke, P., Knight, D. E. & McGregor, A. M. (1990) J. Mol. Endocrinol. 5, 259-266.
- Kremer, R., Karaplis, A. C., Henderson, J., Gulliver, W., Banville, D., Hendy, G. N. & Goltzman, D. (1991) J. Clin. 11. Invest. 87, 884-893
- 12. Iismaa, T. P. & Shine, J. (1992) Curr. Opin. Cell Biol. 4, 195-202
- Lin, S.-C., Lin, C.-R., Gukovsky, I., Lusis, A. J., Sawchenko, 13. P. E. & Rosenfeld, M. G. (1993) Nature (London) 364, 208-213.
- Strewler, G. D. & Nissenson, R. A. (1992) in Parathyroid 14. Hormone-Related Peptide: Normal Physiology and its Role in Cancer, eds. Halloran, B. P. & Nissenson, R. A. (CRC, Boca Raton, FL), pp. 145-167.
- 15. Fraser, C. M., Arakawa, S., McCombie, R. W. & Venter, J. C. (1989) J. Biol. Chem. 264, 11754–11761.
 16. Albert, P. R., Zhou, Q.-Y., Van Tol, H. H. M., Bunzow, J. R.
- & Civelli, O. (1990) J. Biol. Chem. 265, 5825-5832.
- 17. Machida, C. A., Bunzow, J. R., Searles, R. P., Van Tol, H., Tester, B., Neve, K. A., Teal, P., Nipper, V. & Civelli, O. (1990) J. Biol. Chem. 265, 12960-12965.
 18. Flordellis, C. S., Handy, D. E., Bresnahan, M. R., Zannis,
- V. I. & Gavras, H. (1991) Proc. Natl. Acad. Sci. USA 88, 1019-1023.
- 19. Lomasney, J. W., Cotecchia, S., Lorenz, W., Leung, W.-Y., Scwinn, D. A., Yang-Feng, T., Brownstein, M., Lefkowitz, R. J. & Caron, M. G. (1991) J. Biol. Chem. 266, 6365-6369.
- 20. Koo, Y. B., Ji, I., Slaughter, R. G. & Ji, T. H. (1991) Endocrinology 128, 2297-2308.
- 21. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY) pp. 2.73-2.76.
- 22. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY) pp. 7.79-7.83.
- Urena, P., Kong, X.-F., Abou-Samra, A.-B., Jüppner, H., Kronenberg, H. M., Potts, J. T., Jr., & Segre, G. (1992) 23. Endocrinology 133, 617-623.
- 24. Smale, S. T. & Baltimore, D. (1989) Cell 57, 103-113.
- Faisst, S. & Meyer, S. (1992) Nucleic Acids Res. 20, 3-26. 25.
- Akira, S., Isshiki, H., Sugita, T., Yanabe, O., Konoshita, S., Nishio, Y., Nakajima, T., Hirano, T. & Kishimoto, T. (1990) 26. EMBO J. 9, 1897-1906.
- 27. Roodman, G. D. (1992) J. Bone Min. Res. 7, 475-478.
- Thompson, C. B., Wang, C.-Y., Ho, I.-C., Bohjanen, P. R., Petriniak, B., June, C. H., Miesfeldt, S., Zhang, L., Nabel, 28. G. J., Karpinski, B. & Leiden, J. M. (1992) Mol. Cell. Biol. 12, 1043-1053.
- 29. Brown, T. A. & McKnight, S. L. (1992) Genes & Dev. 6, 2502-2512.
- 30. Bhat, N. K., Fisher, R. J., Fujiwara, S., Ascione, R. & Papas, T. (1987) Proc. Natl. Acad. Sci. USA 84, 3161-3165.
- 31. Urena, P., Iida-Klein, A., Kong, X.-F., Abou-Samra, A.-B. & Segre, G. (1992) J. Bone Min. Res. 7, S130.
- 32. Wu, S., Okano, K., Pirola, C. J., Wang, H. M., Jüppner, H., Abou-Samra, A.-B., Segre, G., Fagin, J. A. & Clemons, T. L. (1993) J. Bone Miner. Res. 8, S281.
- 33. Tian, J., Smogorzewski, M., Kedes, L. & Massry, S. G. (1993) J. Am. Soc. Nephrol. 4, 729.
- 34 Wickens, M. & Stephenson, P. (1984) Science 226, 1045-1051.
- 35. McCuaig, K., Turbide, C. & Beauchemin, N. (1992) Cell Growth Differ. 3, 165-174.
- 36. Cavener, D. R. & Ray, S. C. (1991) Nucleic Acids Res. 19, 3185-3192.