CHAPTER 2

Parathyroid Hormone, from Gene to Protein

Osnat Bell, Justin Silver and Tally Naveh-Many

Abstract

The biosynthetic pathway of parathyroid hormone (PTH) has been studied from gene expression to PTH intracellular processing.[1] The processing of PTH has been described and involves the synthesis of an initial translational product, preProPTH, and two proteolytic cleavages that in turn produce ProPTH and PTH. The genes and cDNAs from ten different species have been cloned, sequenced and characterized. This chapter will summarize the molecular biology of PTH, from the gene to the mRNA, the initial translational product, preProPTH and the processed mature secreted form of PTH. It will describe the sequences of the PTH gene and mRNA in different species and the specific elements in the PTH mRNA that determine mRNA processing, stability and translation.

The Prepro PTH Peptide

The primary form of PTH, which is stored and secreted, contains 84 amino acids.[2] PTH is initially synthesized as a precursor, preProPTH. Two proteolytic cleavages produce the ProPTH and the secreted form of PTH. The proPTH sequence contains six extra amino acids at the N-terminus.[3,4] Conversion of ProPTH to PTH occurrs about 15 to 20 min after biosynthesis at about the time ProPTH reached the Golgi apparatus.[5]

The Structure of the Pre-Peptide

Evidence that the translational product of PTH mRNA was larger than ProPTH was initially obtained by translation of a crude preparation of bovine parathyroid RNA in the wheat germ cell-free system.[6] The primary translational product migrated slower than ProPTH when analyzed by electrophoresis on either acidic-urea or sodium dodecyl sulfate-containing acrylamide gels. At that time, a similar phenomenon had been observed only for myeloma light chains.[7] In further studies, preProPTH was shown to be synthesized in cell-free systems of reticulocyte lysates.[8] Translation of human parathyroid RNA also produced an analogous preProPTH.[9]

The observation that the carboxyl terminal peptides of bovine PTH and preProPTH were identical indicated that the extra amino acids in preProPTH were at the amino terminus. This was confirmed by incorporating selected radioactive amino acids into preProPTH and determining the location of the radioactivity by automated Edman degradation.[10] By analyzing overlap of these radioactive amino acids with those in ProPTH, the length of the bovine pre-peptide was shown to be 25 amino acids. The entire sequence of the bovine pre-peptide was determined eventually by this microsequencing technique[11] and was later confirmed by

Molecular Biology of the Parathyroid, edited by Tally Naveh-Many.

structural studies of both the bovine PTH cDNA and gene.[12-14] The sequence of human pre-peptide was also partially determined by this microsequencing technique.[9] The complete amino acid sequence was derived from the human PTH cDNA sequence[15] and later confirmed by the determination of the structure of the human gene.[16] The amino acid sequence of the rat pre-peptide was derived from the sequence of the rat PTH gene[17] and partially by analysis of cloned rat PTH cDNA.[18]

The amino acid sequences of the pre-peptides show that the human and bovine pre-peptides are 80% homologous while the rat sequence is 64% homologous to the bovine and human.[1] This is somewhat lower than the homology of 89 and 77% in the Pro and PTH regions for bovine/human and rat/bovine-human, respectively (Fig. 1). The fact that the pre-peptide is less conserved than the rest of the molecule is consistent with pre-peptides or signal peptides of many eukaryotic proteins.[19] General structural features of the signal peptides are a central hydrophobic core and, in many cases, charged amino acids at the N-terminal and C-terminal ends of the central core. These features are largely retained in the pre-peptides of the three preProPTH molecules. Only conservative changes are present within the central core of uncharged amino acids from amino acids 10 to 21.[1]

Conversion of PrePro to ProPTH

The removal of the pre-peptide to produce ProPTH is mediated by an enzyme associated with microsomes.[8] In reticulocyte and wheat germ systems that contain little or no microsomal membranes, the primary transcriptional product of PTH mRNA is preProPTH.[6,8] Addition of microsomal membranes from dog pancreas or chicken oviduct results in the synthesis of ProPTH.[8,20]

The first evidence that pre or signal peptides function by binding to a limited number sites in the microsomal membrane was obtained by studies on a synthetic prePro-peptide of bovine preProPTH.[21] The identification of the signal recognition particle as a signal peptide receptor, later on, confirmed this mechanism for most secreted and membrane proteins.[22]

The pre peptide of preProPTH is rapidly degraded after its proteolytic cleavage from preProPTH. In studies of PTH biosynthesis in intact cells, no labeled pre-peptide could be detected.[23] The proteolytic removal of the pre-peptide probably occurs before completion of the ProPTH nascent chain, since preProPTH is difficult to detect in intact cells.

Homology of the Mature PTH

The mature PTH has been determined or predicted by the cDNAs in several species. The sequence of PTH of mouse, rat, man, non-human primates, horse, dog, cat, cow, pig, and chicken is shown in Figure 1. The resulting phylogenetic tree obtained from alignment of the protein sequences is shown in Figure 3A.

A comparison of the amino acid sequences of PTH from several species revealed high conservation of the protein amongst all species apart from gallus (Fig. 1). In addition, three relatively conserved regions could be observed.[17] The first two regions comprise the biologically active region of PTH and would be expected to be conserved. The addition or loss of a single amino acid at the amino terminus greatly reduces biological activity, and the region is involved in binding of PTH to the receptor. In addition there is a region of conservation at the C-terminal region that is itself of interest, particularly since this region may have a separate biological effect at least on osteoclasts.[24] Analyses of the silent changes that occur between the nucleotide sequences suggest that the conservation in the C-terminal region may be related to pre-translational events. Analysis by Perler et al[25] described replacement changes that result in changes in amino acids and silent changes that do not alter the encoded amino acid.

```
                                        pre    pro
                                         ↓      ↓
          1                                                    50
murine    MMSANTVAKV MIIMLAVCLL TQTDGKPVRK RAVSEIQLMH NLGKHLASME
rat       MMSASTMAKV MILMLAVCLL TQADGKPVKK RAVSEIQLMH NLGKHLASVE
human     MIPAKDMAKV MIVMLAICFL TKSDGKSVKK RSVSEIQLMH NLGKHLNSME
macaca    MIPAKDMAKV MIVMLAICFL TKSDGKSVKK RSVSEIQLMH NLGKHLNSME
Equine                                  K RSVSEIQLMH NLGKHLNSVE
canine    MMSAKDMVKV MIVMFAICFL AKSDGKPVKK RSVSEIQFMH NLGKHLSSME
feline    MMSAKDMVKV MVVMFAICFL AKSDGKPVKK RSVSEIQFMH NLGKHLSSVE
bovine    MMSAKDMVKV MIVMLAICFL ARSDGKSVKK RAVSEIQFMH NLGKHLSSME
porcine   MMSAKDTVKV MVVMLAICFL ARSDGKPIKK RSVSEIQLMH NLGKHLSSLE
gallus    MTSTKNLAKA IVILYAICFF TNSDGRPMMK RSVSEMQLMH NLGEHRHTVE

          51                                                   100
murine    RMQWLRRKLQ DMHNFVSLGV QMAARDGSHQ KPTKKEENVL VD........
rat       RMQWLRKKLQ DVHNFVSLGV QMAAREGSYQ RPTKKEENVL VD........
human     RVEWLRKKLQ DVHNFVALGA PLAPRDAGSQ RPRKKEDNVL VE........
macaca    RVEWLRKKLQ DVHNFIALGA PLAPRDAGSQ RPRKKEDNIL VE........
Equine    RVEWLRKKLQ DVHNFIALGA PIFHRDGGSQ RPRKKEDNVL IE........
canine    RVEWLRKKLQ DVHNFVALGA PIAHRDGSSQ RPLKKEDNVL VE........
feline    RVEWLRRKLQ DVHNFVALGA PIAHRDGGSQ RPRKKEDNVP AE........
bovine    RVEWLRKKLQ DVHNFVALGA SIAYRDGSSQ RPRKKEDNVL VE........
porcine   RVEWLRKKLQ DVHNFVALGA SIVHRDGGSQ RPRKKEDNVL VE........
gallus    RQDWLQMKLQ DVHS...... ..ALEDARTQ RPRNKEDIVL GEIRNRRLLP

          101                                  128
murine    .....GNPKS LGEGDKADVD VLVKSKSQ
rat       .....GNSKS LGEGDKADVD VLVKAKSQ
human     .....SHEKS LGEADKADVN VLTKAKSQ
macaca    .....SHEKS LGEADKADVD VLTKAKSQ
Equine    .....SHQXS LGEADKADVD VLSKTKSQ
canine    .....SYQKS LGEADKADVD VLTKAKSQ
feline    .....NHQKS LGEADKADVD VLIKAKSQ
bovine    .....SHQKS LGEADKADVD VLIKAKPQ
porcine   .....SHQKS LGEADKAAVD VLIKAKPQ
gallus    EHLRAAVQKK SIDLDKAYMN VLFKTKP~
```

Figure 1. Alignment of the amino acid sequences of PTH from the 10 different species. Alignments were obtained using the default setting of PileUp program (Accelrys Inc. Madison WI). Comparison of the amino acid sequences of PTH for mouse (mus), rat, human, non human primates (macaca), horse (equine), dog (canine), cat (feline), cow (bovine), pig (porcine) and chicken (gallus). Gaps indicated by dashes were introduced to maximize the homology to the gallus sequence. The N terminal sequence of the equus PTH is not available. The arrows indicate the protolytic cleavage sites required for the conversion of preProPTH to ProPTH and PTH.

The PTH mRNA

Bovine preProPTH mRNA was initially more extensively characterized than the mRNAs from the other species. Preparations of bovine parathyroid RNA were obtained that contained about 50% PTH mRNA as estimated by gel electrophoresis and RNA excess hybridization to radioactive cDNA.[26] The size of the mRNA was estimated to be about 750 nucleotides by sucrose gradient centrifugation. About two thirds of the translatably active mRNA was retained by oligo(dT) cellulose, and the sizes of the poly(A) extension was broadly distributed around an average size of 60 adenylate residues, though this may be an under estimation of the actual size. While not directly determined, PTH mRNA probably contains a 7-methylguanosine cap since the translation of PTH mRNA was inhibited by 7-methylguanosine-5′-phosphate. The human and bovine

PTH mRNAs appear to be heterogeneous at the 5' terminus (see section on genes). The sizes of the rat and human PTH mRNAs have been determined by Northern blot analysis to be about 800 and 850 nucleotides, respectively.[15,17] Therefore, PTH mRNAs are typical eukaryotic mRNAs that contain a 7-methyguanosine cap at the 5' terminus and a polyadenylic acid (poly A) stretch at the 3' terminus. The PTH mRNAs are twice as long as necessary to code for the primary translational product, due to 5' and 3' untranslated regions at both ends of the mRNA.

Cloning of the PTH cDNAs

To date the sequence of the full cDNA of rat,[17] man,[15] dog,[27] cat (un published), cow,[13] pig,[18] and chicken[28] and the partial sequence of horse[29] and non human primates[30] have been determined. The cDNA of mouse PTH was determined from the genomic PTH sequence.[31] Table 1 shows the Gene Bank accession number for the PTH sequences of the different species and the length of the cDNAs of each of the mRNAs as they appear in the NCBI and Gene Bank databases. In addition, the hypothalamus PTH cDNA was sequenced after the PTH mRNA had been detected in neuronal tissue.[32]

The first PTH cDNAs identified were the DNAs complementary to bovine[12,13] and human[15] PTH mRNA that had been cloned into the Pst 1 site of pBR322 by the homopolymer extension technique. The rat PTH cDNA[18] was cloned by the Okayama and Berg method. The bovine mRNA was isolated from normal parathyroid glands, and the human mRNA was isolated from parathyroid adenomas. The sequence of the rat mRNA has been derived partially from the rat cDNA and from the sequence of the cloned gene.[17]

Kronenberg et al[12] initially determined the sequence of a bovine cDNA clone, pPTHml, which contained about 60% of the PTH mRNA, including the entire region coding for pre-ProPTH. Restriction analysis of near full-length double-stranded cDNA, synthesized enzymatically from partially purified bovine PTH mRNA, indicated that about 200 nucleotides from the 3' untranslated region were missing in the clone.[33] Analysis of several additional bovine PTH cDNA clones and the sequencing of cDNA of the 5' terminus of PTH mRNA, which was synthesized by extension of a primer with reverse transcriptase, provided the full bovine DNA sequence.[34]

Nucleotide sequences of the parathyroid (PTH) gene of 12 species of non-human primates belonging to suborder Anthropoidea were characterized.[30] The deduced amino acid sequences of exons II and III of the PTH gene of the 12 species of non- human primates was compared to the human PTH and revealed no amino acid substitution in the mature PTH among orangutans, chimpanzees, and humans. The results indicated that the PTH gene is highly conserved among primates, especially between great apes and humans.[30]

The 5' end of the bovine mRNA sequence, which was determined by sequencing DNA complementary to the 5' end of PTH mRNA produced by primed reverse transcription,[34] produced multiple 5' termini of the mRNA. The heterogeneity at the beginning of the 5' end of the mRNA was confirmed by S1 nuclease mapping.[14] The longest reverse transcribed cDNA was isolated and sequenced. Surprisingly, this cDNA contained a canonical TATA sequence at the beginning, which was in the proper position to direct the transcription of the shorter mRNAs. This result suggested that a second TATA sequence would be present 5' to the one detected in the cDNA and would direct the synthesis of the longer mRNAs. The predicted second TATA sequence was discovered when the gene was sequenced. The 5' end of the rat PTH mRNA was also analyzed by S1 nuclease mapping and was less heterogeneous than the bovine mRNA.[17] The single species of rat PTH mRNA corresponded to the larger of the bovine mRNAs. The size of the human mRNA, based on the cDNA sequence, is about 100 nucleotides longer than the bovine and rat mRNAs (Table 1). Northern blot analysis of the mRNAs was consistent with these predicted sizes.[17] The 3' untranslated region (UTR) of the avian PTH mRNA is 1236 nt long, much larger than any of the PTH mRNA 3'-UTRs (Table 1). In general the difference in size in

Table 1. List of the known sequences for the PTH gene and the sizes of the mRNA, 5'-UTR, coding region and 3'-UTR

	NCBI Accession Number	mRNA	5'UTR	CDS	3'UTR
Mus musculus	Af066074: gene, exon 1 Af066075: gene, exons 2 and 3 and complete mRNA (deduced)	714	127	348	239
Rat	K01267: gene, exon 1 K01268: gene, exon 2 and 3 X05721: mRNA, complete	704	118	348	238
Canis familiaris	U15662: mRNA, complete	692	88	348	256
Felis catus	Af309967: mRNA, complete	737	63	348	326
Human	J00300:gene, 3' end J00301: gene, coding region and 3' flank V00597: mRNA, complete	772	74	348	350
Macaca fascicularis	Af130257: gene, complete cds	398*		348	50*
Bovine	K01938: gene, complete cds and flank M25082: mRNA, complete	699	127	348	224
Equus caballus	Af134233: gene, partial cds	311*		267*	44*
Porcine	X05722: mRNA, complete	698	96	348	254
Gallus gallus	M36522: mRNA, complete	1723	127	360	1236

The NCBI accession number of the different sequences and the size of the mRNAs are indicated. The asteryxes show sequences that have been partially sequenced.

the PTH mRNA of the different species primarily results from the difference in the size of the 3'-UTR (Table 1). The significance of this finding has not been studied.

The overall nucleotide compositions of the cDNAs are similar. All the sequences are A-T rich. The 3' noncoding region has a particularly large portion of A and T, ranging from 68 to 74%, making it an AU rich element (ARE). The rat sequence differs from the other sequences in that the 5' noncoding region is only 50% A and T compared to 63 to 65% for the human and bovine.

Homology of the cDNA Sequences

Alignment of the PTH cDNAs of the nine preProPTH sequences is shown in Figure 2. Gaps have been introduced in the 5' and 3' untranslated regions to maximize homology. For

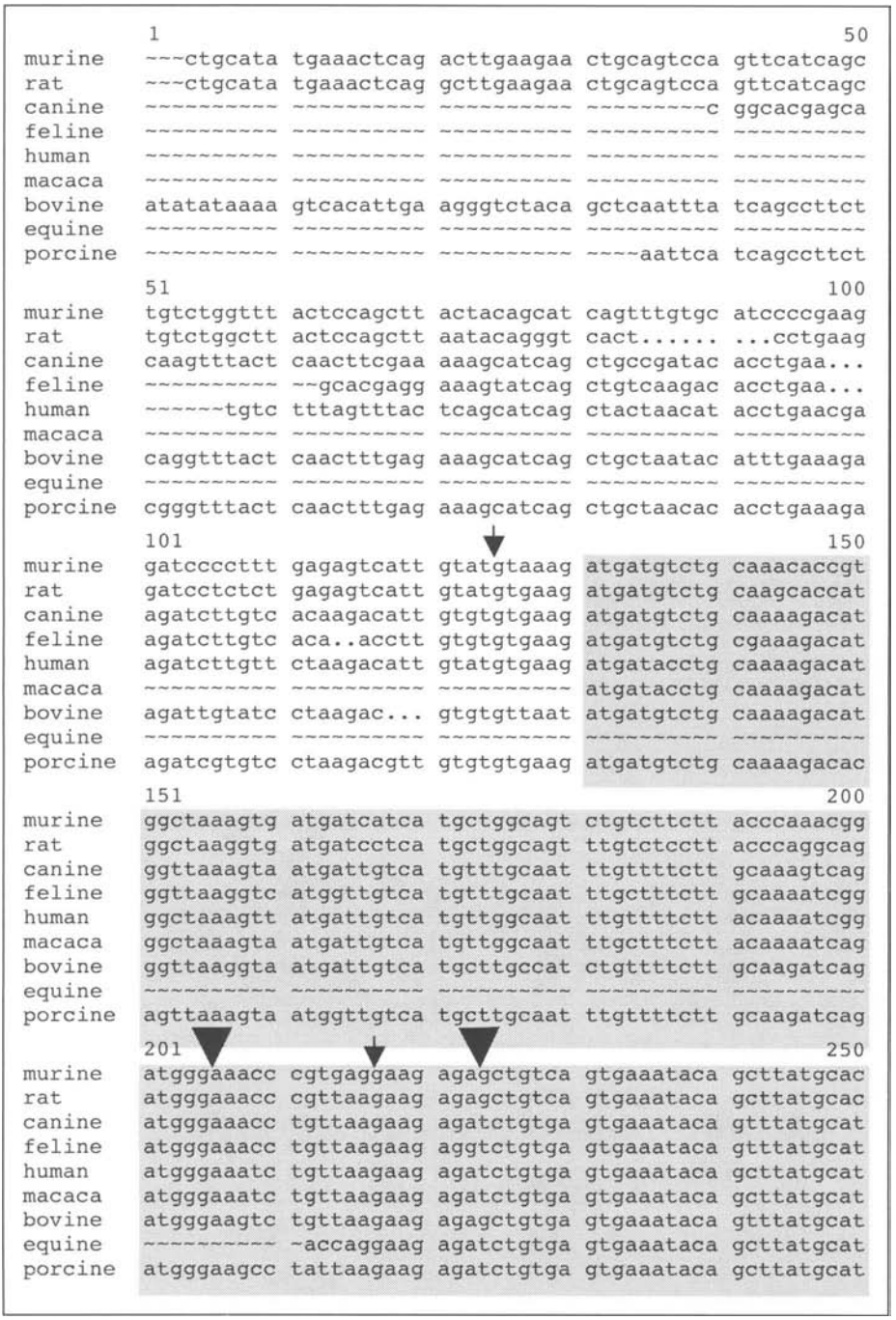


Figure 2. Part 1, see legend page 16.

```
          251                                                        300
murine    aacctgggca aacacctggc ctccatggag aggatgcaat ggctgagaag
rat       aacctgggca aacacctggc ctctgtggag aggatgcaat ggctgagaaa
canine    aacctgggca aacatctgag ctccatggag agggtggaat ggctacggaa
feline    aacctgggca agcatctgag ctccgtggag agggtagaat ggctgcggag
human     aacctgggaa aacatctgaa ctcgatggag agagtagaat ggctgcgtaa
macaca    aacctgggaa aacatctgaa ctcgatggag agagtagaat ggctgcgtaa
bovine    aacctgggca aacatctgag ctccatggaa agagtggaat ggctgcggaa
equine    aacctgggca aacatctgaa ctcagtggaa agggtggaat ggctgcggaa
porcine   aacctgggca aacacctgag ctctctggag agagtggaat ggctgcgaaa
          301                                                        350
murine    gaagctgcaa gatatgcaca attttgttag tcttggagtc caaatggctg
rat       aaagctgcaa gatgtacaca attttgttag tcttggagtc caaatggctg
canine    gaagctccag gatgtacaca actttgttgc ccttggagct ccaatagctc
feline    gaaactacag gatgtacaca actttgtcgc cctcggagct ccaatagctc
human     gaagctgcag gatgtgcaca attttgttgc ccttggagct cctctagctc
macaca    gaagctgcag gatgtgcaca attttattgc ccttggagct cctctagctc
bovine    aaagctacag gatgtgcaca actttgttgc ccttggagct tctatagctt
equine    gaagctgcag gatgtgcaca attttattgc cctcggagct cctatatttc
porcine   gaagctgcag gatgtgcaca actttgttgc cctcggagct tctatagttc
          351                                                        400
murine    ccagagatgg cagtcaccag aagcccacca agaaggagga aaatgtcctt
rat       ccagagaagg cagttaccag aggcccacca agaaggagga aaatgtcctt
canine    acagagatgg tagttcccag aggcccctaa aaaaggaaga caatgtccta
feline    acagagatgg tggttcccag aggccccgaa aaaaggaaga caatgtcccg
human     ccagagatgc tggttcccag aggccccgaa aaaaggaaga caatgtcttg
macaca    ccagagatgc tggttcccag aggccccgaa aaaaggaaga caatatcttg
bovine    acagagatgg tagttcccag agacctcgaa aaaaggaaga caatgtcctg
equine    acagagatgg tggttcccag aggcctcgaa aaaaggaaga caatgtgctg
porcine   acagagatgg tggttcccag agaccccgaa aaaaggaaga caatgtcctg
          401                                                        450
murine    gttgatggca atccaaaaag tcttggtgag ggagacaaag ctgatgtgga
rat       gttgatggca attcaaaaag tcttggcgag ggggacaaag ctgatgtgga
canine    gttgagagct atcaaaaaag tcttggagaa gccgacaaag ctgatgtgga
feline    gctgagaacc atcaaaaaag tcttggagaa gcagacaaag ctgatgtgga
human     gttgagagcc atgaaaaaag tcttggagag gcagacaaag ctgatgtgaa
macaca    gtagagagcc atgaaaaaag tcttggagag gcagacaaag ctgatgtgga
bovine    gttgagagcc atcagaaaag tcttggagaa gcagacaaag ctgatgtgga
equine    attgagagcc atcaaraaag tcttggagaa gcagacaaag ctgatgtgga
porcine   gttgagagcc atcaaaaaag tctcggagaa gcagataaag ctgctgtgga
          451                                                        500
murine    tgtattagtt aaatcaaaat ctcagtaaat gctgatttat tctagacagt
rat       tgtattagtt aaggctaaat ctcagtaaat gctgacgtat tctagaccgt
canine    tgtattaact aaagctaaat cccagtgacg ataca..... ......tcag
feline    tgtgttaatc aaagctaaat cccagtgaag acaga..... ......gcag
human     tgtattaact aaagctaaat cccagtgaaa atgaaaacag atattgtcag
macaca    tgtattaact aaagctaaat cccaatgaaa atgaaaatag atatggtcag
bovine    tgtattaatt aaagctaaac cccagtgaa. ...aacagat atgatc..ag
equine    tgtgttaagt aaaactaaat cccagtgar. ...aacagat aggatc..ag
porcine   tgtattaatt aaagctaaac cccagtgaa. ...aacacat atgatcagag
```

Figure 2. Part 2, see legend page 16.

```
          501                                                       550
murine    gcagggcact gacatatgct gctacctttt caagct.tat gaagatcacc
rat       gctgagcaat aacatatgct gctatccttt caagctccac gaagatcacc
canine    ggcactgctg tagacagcat agggcaacaa cattacaagc tgctaacatt
feline    agcactgcta tacacaggat agggcaacaa aattacatgc tgctaacatt
human     agttctgctc tagacagtgt agggcaacaa tacatgctgc taattcaaag
macaca    agttctgctc tagacagtgt agggcaac~~ ~~~~~~~~~~ ~~~~~~~~~~
bovine    atcactgttc tagacagcat agggc.aaca atattacatg ctgctaatgt
equine    agcatcgctc tagacagcat asggc.aaa~ ~~~~~~~~~~ ~~~~~~~~~~
porcine   agcactgctc tagacagcat aaggcaaaca atatttcatg ctgctaatgt
          551                                                       600
murine    aagtgctaat acttctactg taatgaaact ttggaatttt tttgattaca
rat       aagtgctaat tcttctactg taataaaagt ttgaaa.... tttgattcca
canine    ttcaagctct taagattaat aaatgccaaa atttacatgt aatccattgt
feline    ctcaagcttt gaagatcacc aaatgccaat atttacgtct aatccatggc
human     ctctattaag atttccaagt gccaatattt ctgatataac aaactacatg
macaca    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
bovine    gttcaccttc tattaagtgc cagtagttct atgaccaacc tttattgcta
equine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
porcine   tttcaatctc tattaagatt aagtgccaat atttctaata ttactaaact
          601                                                       650
murine    tttttgctca tttaaggtct ctttcaatga ttccatttca atatgctctt
rat       cttttgctca tttaaggtct cttccaatga ttccatttca atatattctt
canine    tagccatgat agctgaaatt ttaattgatt gttttgattc tagtttaatt
feline    tagccacgat agctgaaatt ctaattgatt gttttgattc tacttttatt
human     taatccatca ctagccatga taactgcaat tttaattgat tattctgatt
macaca    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
bovine    gctgtgatac ctacaatttt aattgagtat tttgattcta ctttattcat
equine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
porcine   tgatgggtaa tcattgctag ccatgattgc tgaaatttta attgatcatt
          651                                                       700
murine    ctttttaaag tactactcat ttccacttct ctccttaaat ataaataaag
rat       ctttttaaag tattacacat ttccacttct ctccttaaat ataaataaag
canine    catttaagag ctcttttaat tgttctattt ctattgttta ttctttttaa
feline    catgtaaggc ctcttttaat tattccattt ctgttgttta ttctttttaa
human     ccacttttat tcatttgagt tattttaatt atcttttcta ttgtttattc
macaca    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
bovine    ctaagagctc ttttaataat tctatttcta ttgattccaa ataaatgaag
equine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
porcine   ttgattctac ttttactcat ttaagagctt cttttaacaa ttctatttct
          701                                                       750
murine    ctttaatgct catgaatc~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
rat       ~tttaatgat catgaaccaa a~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
canine    agtatgtttt tgcataattt ataaaagaat aaaattgcac ttttt~~~~~
feline    agtatgttat tgcataattt ataaaagaat aaaattgcac tttgtaacct
human     tttttaaagt atgttattgc ataatttata aaagaataaa attcgacttt
macaca    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
bovine    ttaagtatt~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
equine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
porcine   attgattcta aataaatgaa gtatttcttc cttgtt~~~~ ~~~~~~~~~~
```

Figure 2. Part 3, see legend page 16.

```
          751                                                        800
murine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
rat       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
canine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
feline    ctctcccatc gtacactgca aaataaaaat ttaatgatca taattttaaa
human     taaacctctc ttctaccttа aaatgtaaaa caaaaatgta atgatcataa
macaca    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
bovine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
equine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
porcine   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          801                        828
murine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~
rat       ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~
canine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~
feline    aaaaaaaaaa aaaaa~~~~~ ~~~~~~~~
human     gtctaaataa atgaagtatt tctcactc
macaca    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~
bovine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~
equine    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~
porcine   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~
```

Figure 2. Alignment of the known nucleotide sequences of PTH mRNA for different species. Alignments were obtained using the default setting of PileUp program. Alignment of the nucleotide sequences for mouse (murine), rat, dog (canine), cat (feline), human, non-human primate (macaca), cow (bovine), horse (equine) and pig. The gallus sequence was not included in the alignment because of the large differences in sequence and size from all other published species. Gaps indicated by dashes were introduced to maximize the homology in the 5' and 3' -UTRs. The arrows indicate the positions of the two introns in the gene. The closed triangles indicate the protolytic cleavage sites required for the conversion of preProPTH to ProPTH and PTH. The nt in the dark gray box show the coding sequence, and the sequences 5' and 3' to this region are the 5'-UTR and 3'-UTR respectively. The nt that are surrounded by the square comprise the proximal PTH mRNA 3'-UTR protein binding element and the nt that are on a light gray background are the distal *cis* acting functional element. Nucleotides that are not identical to the bovine (proximal element) and the rat (distal element) sequence are shown in bold.

simplicity the gallus cDNA was not included in the alignment of the pre ProPTH mRNAs in Figure 2. This PTH mRNA is significantly longer than the other cloned cDNAs (Table 1) and is the least preserved compared to the other species, even in the coding sequence (Table 2 and Fig. 1).

Comparison of the sequences show that human and macaca; canis and felis; rat and mouse; and bovine and pig are the most similar to each other (Table 2). The lowest homology is seen when the sequence of gallus PTH mRNA is compared to each of the other sequences, even in the translated coding region of the mRNA that is, as expected, the most conserved region. The coding sequences of the other species are the most preserved as expected. The 5'-UTR is relatively well conserved with homologies about 15% less than the coding region. The 3'-UTR is the least conserved region (Table 2).

Interestingly, a 26 nt *cis* acting functional protein binding element at the distal region of the 3' UTR is highly conserved in the PTH mRNA 3'-UTRs of rat, mouse, man, dog and cat (Table 3, distal element). In the 26 nt element, the identity amongst species varies between 73 and 89%. In particular, there is a stretch of 14 nt within the element that is present in all five species. We have previously characterized this distal protein binding element in the rat PTH mRNA 3'-UTR as a *cis*-acting sequence that determines the stability of the PTH mRNA and its regulation by calcium and phosphate (P).

Table 2. Similarity (ratio) of the nucleotide sequences for the PTH mRNAs of different species

	mRNA	5'-UTR	CDS	3'-UTR
Bovine/Porcine	7.864	8.229	9.253	6.643
Bovine/Canine	7.341	6.364	9.224	6.049
Bovine/Human	6.682	5.541	8.879	5.571
Bovine/Rat	5.863	4.610	7.730	4.107
Human/Macaca			9.613	9.6
Human/Canine	7.542	5.824	9.052	6.605
Human/Rat	5.804	4.743	7.816	4.660
Rat/Murine	8.743	8.5	9.253	8.118
Rat/Canine	5.925	4.511	7.989	4.176
Canine/Feline	8.503	7.206	9.195	8.504
Bovine/Gallus	4.565	4.283	5.330	3.705
Human/Gallus	4.519	4.554	5.474	3.809
Rat/Gallus	4.207	3.847	5.043	3.765
Canine/Gallus	4.444	4.250	5.388	3.691

The comparisons between each two sequences were performed using the default setting of the GAP program (Accelrys Inc, Madison WI). This program considers all possible alignments and gap positions between two sequences and creates a global alignment and evaluates its significance. The average alignment score, plus or minus the standard deviation, of all randomized alignments is reported in the output file as *the 'quality'* score. Ratio is the quality divided by the number of bases in the shorter segment of each two sequences.

In addition, a 22 nt protein binding element in the 3' UTR (Table 3 proximal element) was also identified in bovine and porcine, as well as human, non-human primates, equus, canis and felis (Table 3 proximal element), but not in rat and mouse. The functionality of the proximal element remains to be determined. The conserved sequences within the 3'-UTR suggest that the binding elements represent a functional unit that has been evolutionarily conserved (see 'conserved elements in the 3' UTR).

The 3'-UTR in the human and feline sequences are more than 100 nucleotides longer than the other 3' UTR sequences, with the exception of the gallus PTH mRNA. Large gaps have to be introduced to maximize homology to the human 3'-UTR (Fig. 2). Hendy et al[15] suggested that the extra sequence in the 3' region of the human cDNA, corresponding to the large gap in the bovine sequence, might have been the result of a gene duplication since it contained some homology to the region around the polyadenylation signal, including a second consensus polyadenylation signal. Interestingly, in the rat sequence, large gaps also must be introduced in this region, but they do not coincide exactly with that of the bovine sequence. Phylogenetic trees obtained from alignment of the protein and mRNA sequences are shown in Figure 3. The same phylogenetic tree is obtained from the amino acid sequences and from the coding regions of the mRNA (Fig. 3A). Phylogenetic comparison based on nt similarity of the full PTH mRNAs or the 3'-UTRs is shown in Figure 3B. This map does not include macaca and equine PTH sequences where there are only partial sequences of the cDNA available. The gallus is very different from all the other species indicating a separate evolutionary branch. Interestingly, based on amino acid sequence and the coding region of the mRNA, the bovine and porcine were grouped closest to canis and felis but not by the full-length mRNA or 3'-UTR sequences. This mainly represents the large

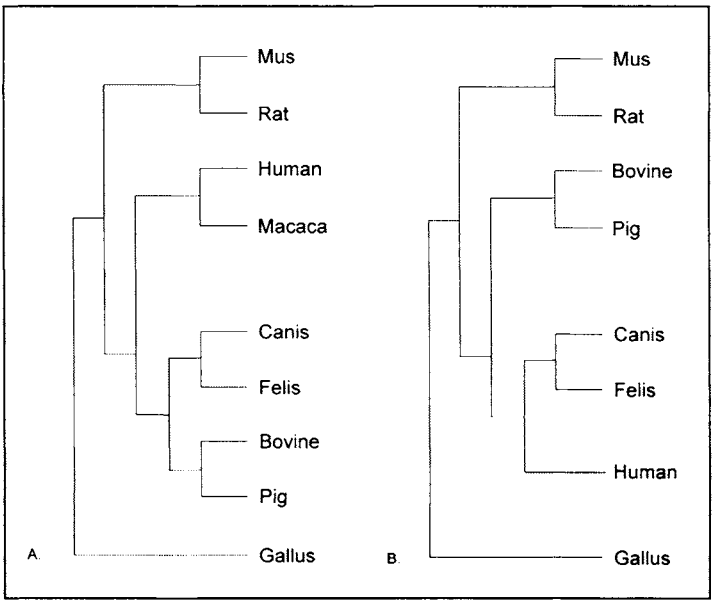


Figure 3. Phylogenetic tree obtained from alignment of the amino acid sequences and nucleotide sequences of PTH and PTH mRNA. The phylogenetic trees were obtained using the default setting of PileUp program. A) Phylogenetic tree based on amino acid similarities, or the nt sequence of the coding regions of the PTH mRNAs for mouse (mus), rat, dog (canis), cat (felis), human, non-human primate (macaca), cow (bovine), pig (porcine) and chicken (gallus) PTH. The horse PTH was not included in this study because only partial amino acid sequence is available. B) Phylogenetic tree according to nt sequence of the full-length PTH mRNAs for mouse (mus), rat, dog (canine), cat (feline), human, cow (bovine), pig (porcine) and chicken (gallus). The horse and non-human primate (macaca) PTH mRNA were not included in this study because only partial sequences of these RNAs are available. The same Phylogenetic tree is also obtained when the 3'-UTR sequences are analized separately. Interestingly, based on amino acid sequence, the bovine and pig were grouped closest to canis and felis but not by RNA sequence. This corresponds to the presence of the distal and proximal protein-binding elements in the 3'-UTRs.

differences in the 3'-UTRs and correlates with the conservation of protein-binding elements (Table 3). The mouse and rat species are separate because they only have the distal PTH mRNA 3'-UTR element. The human, canis, felis, bovine and porcine are grouped together, all containing the proximal element. But in this group, the bovine and porcine represent a separate branch expressing only the proximal element, and the human, felis and canine are a distinct branch, which corresponds with their expression of both the proximal and distal elements.

Structure of the PTH mRNA

The 5' Untranslated Region

The 5' untranslated sequence of the longer forms of the human and bovine mRNAs and rat PTH mRNA contains about 120 nucleotides, and the shorter bovine and human cDNAs contain about 100 nucleotides in the 5' noncoding region. The average length of the 5' UTR in eukaryotic mRNAs is 80-120 nucleotides.[35] As a result, the m^7G cap at the 5' terminus of the mRNA is a considerable distance from the initiator codon. In the bovine sequence, a

Table 3. The sequences of the 26 nt proximal cis acting element and the 22 nt distal element of the PTH mRNA 3'-UTR in different species

	Proximal Element	Distal Functional Element
Rat	-	ATATATTCTTCTTTTTAAAGTA
Mus musculus	-	ATAT**GC**TCTTCTTTTTAAAGTA
Bovine	TGTTCTAGACAGCATAGGGCAA	-
Porcine	TG**C**TCTAGACAGCATA**A**GGCAA	-
Equus caballus	**C**G**C**TCTAGACAGCATA*GGCAA	?
Canis familiaris	TG**C**T**G**TAGACAGCATAGGGCAA	AT**TGT**TT**A**TTCTTTTTAAAGTA
Felis catus	TG**C**T**A**TA**C**ACAG**G**ATAGGGCAA	**GTTGT**TT**A**TTCTTTTTAAAGTA
Human	TG**C**TCTAGACAG**TG**TAGGGCAA	AT**TGT**TT**A**TTCTTTTTAAAGTA
Macaca fascicularis	TG**C**TCTAGACAG**TG**TAGGGCAA	?
Gallus gallus	-	-

The sequences that are not available are indicated by a ?; species lacking a particular element are indicated by a -. The * in the equine proximal element indicates an unidentified nt in the gene bank. The nt in the proximal and distal elements that are different from the bovine and rat sequence respectively are shown in bold.

possible hairpin loop may bring the 5' end closer to the initiator codon. However, in both the human and rat sequences, deletions of 11 and 16 nucleotides respectively, largely eliminate the sequences involved in the stem of the loop. Thus, there seems to be little functional significance related to the bovine secondary structure. In the rat PTH mRNA 5' terminus the first 19 nt of the mRNA may form a stable stem loop structure that could affect PTH mRNA translation, but its function has not been determined. (T. Naveh-Many, unpublished data). However the most outstanding conclusion from a comparative analysis of the sequences is that no region in the 5' untranslated region is conserved that has any known functional significance.

The Coding Region

The actual initiator ATG codons for the human and bovine PTH mRNAs have been identified by sequencing in vitro translation products of the mRNAs.[36,10] In the bovine sequence, the first ATG codon is the initiator codon, in accord with many other eukaryotic mRNAs.[37,38] The human and rat sequences have ATG triplets prior to the probable initiator ATG, which are present ten nucleotides before the initiation codon and are immediately followed by a termination codon. In the rat, another ATG is present 115 nucleotides before the initiator codon. The designation of the third ATG codon of the rat sequence as the initiator codon is based on indirect evidence, primarily by comparison with the bovine and human cDNAs. Regardless, the presence of termination codons in phase with the earlier ATG prohibits the synthesis of a long protein initiated at these codons, as is the case in some other genes with premature ATG codons.[37,38] However in some systems, small peptides that are translational products of upstream ATGs have been shown to have regulatory functions.[39] Whether this is the case in the rat PTH mRNA is not known. The most stringent requirement for optimal initiation of synthesis is for a purine at the –3 position. Since non of the premature ATG codons in the rat and human PTH mRNAs has a purine at the –3 position, they are likely to be weak initiators. In contrast, the probable initiator ATG codon has an A at the –3 position in each sequence.

The 3' Untranslated Region

As noted above, the 3' untranslated region is the most variable region of the cDNAs requiring significant gaps to maximize homology. The termination codon in all species except for the rat and mouse is TGA and is followed closely by a second in-phase termination codon. In the rat and mouse TAA is the termination codon and no following termination codon is present (Fig. 2).

Conservation of Protein Binding Elements in the PTH mRNA 3'-UTR

We have defined the *cis* sequence in the rat PTH mRNA 3'-UTR that determines the stability of the PTH mRNA and its regulation by calcium and phosphate (P). PTH gene expression is regulated post-transcriptional by Ca^{2+} and P, with dietary induced hypocalcemia increasing and dietary induced hypophosphatemia decreasing PTH mRNA levels. This regulation of PTH mRNA stability correlates with differences in binding of *trans* acting cytosolic proteins to a *cis* acting instability element in the PTH mRNA 3'-UTR. There is no PT cell line and therefore to study PTH mRNA stability we performed in vitro degradation assays. We did this by incubating the labeled PTH transcript with cytosolic PT proteins from rats on the different diets and measuring the amount of intact transcript remaining with time. PT proteins from low Ca^{2+} rats stabilized and low P PT proteins destabilized the PTH transcript compared to PT proteins of control rat. This rapid degradation by low P was dependent upon the presence of the terminal 60 nt protein binding region of the PTH mRNA.[40] We have defined the *cis* sequence in the rat PTH mRNA 3'-UTR that determines the stability of the PTH transcript and to which the *trans* acting PT proteins bind. A minimum sequence of 26 nt was sufficient for RNA-protein binding (Table 3, distal element). One of the *trans* acting proteins that binds and prevents degradation of the PTH mRNA was identified by affinity purification. This protein is AU rich element binding protein 1 (AUF1) that is also involved in half life of other mRNAs.[41,42]

To study the functionality of the *cis* sequence in the context of another RNA, a 63 bp PTH cDNA sequence consisting of the 26 nt and flanking regions was fused to the growth hormone (GH) cDNA. Since there is no parathyroid (PT) cell line an in vitro degradation assay was used to determine the effect of PT cytosolic proteins from rats fed the different diets on the stability of RNA transcripts for GH and the chimeric GH-PTH 63 nt.[43,44] The GH transcript was more stable than PTH RNA and was not affected by PT proteins from the different diets. The chimeric GH PTH 63 nt transcript, like the full-length PTH transcript was stabilized by PT proteins from rats fed a low calcium diet and destabilized by proteins from rats fed a low phosphate diet. Therefore, the 63 nt protein binding region of the PTH mRNA 3'-UTR is both necessary and sufficient to regulate RNA stability and to confer responsiveness to changes in PT proteins by calcium and phosphate.[43] The regulation of PTH mRNA stability by calcium and phosphate is discussed in detail in the chapter by Levin et al.

Sequence analysis of the PTH mRNA 3'-UTR of different species revealed a preservation of the 26 nt core protein-binding element in rat, mouse, human, cat and canine 3'-UTRs (Table 3). The *cis* acting element identified is at the 3' distal end in all species that express it and is therefore designated the distal functional *cis* element. The conservation of the sequence suggests that the binding element represents a functional unit that has been evolutionarily conserved. Protein binding experiments by UV cross linking and RNA electrophoretic mobility shift assays showed that there is specific binding of rat and human parathyroid extracts to an in vitro transcribed probe for the rat and human PTH mRNA 3'-UTR 26 nt elements.

In contrast, the 26 nt distal *cis* element was not present in the 3'-UTR of bovine, porcine and gallus PTH mRNA. To determine the protein binding pattern of the bovine PTH mRNA,

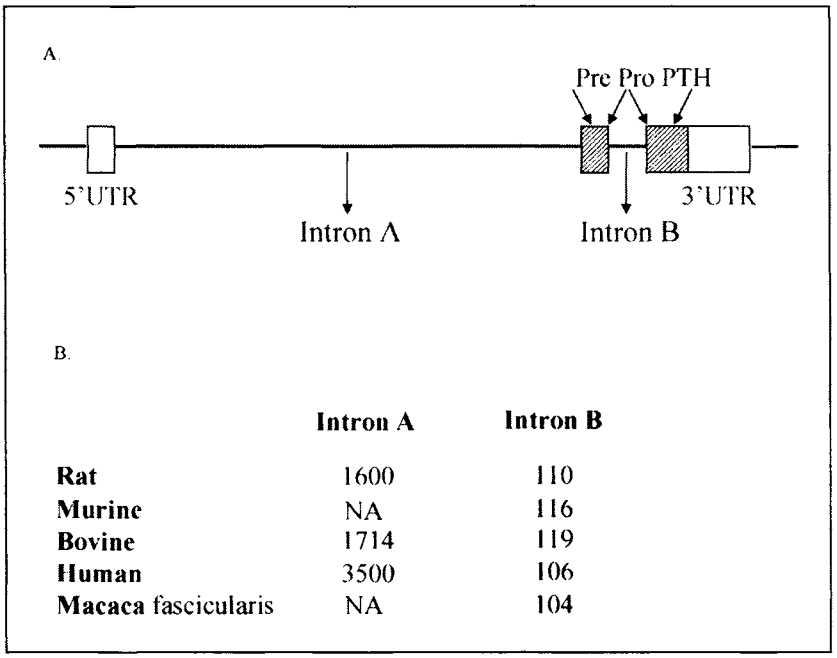
binding experiments were performed with bovine parathyroid gland extracts and RNA probes for different regions of the bovine PTH mRNA. Binding and competition experiments revealed a 22 nt minimal protein binding element in the bovine PTH mRNA 3'-UTR that was sufficient for protein binding. The 22 nt element is at the 5' portion of the 3'-UTR (Fig. 2) and is the proximal *cis* element. Interestingly this element was also present in the 3'-UTRs of man, dog, cat, non-human primates, horse and porcine PTH mRNA. Therefore the PTH mRNA 3'-UTRs of man, dog and cat have both sequences, the distal functional *cis* element of 26 nt that has been characterized in rat PTH mRNA and the 22 nt proximal protein binding element initially characterized in bovine PTH mRNA (Table 3). The bovine and porcine mRNAs only have the 22 nt element and the gallus PTH mRNA has neither of the elements. It is not known if the 26 nt element is present in the horse and macaca because there is only partial sequencing of these mRNAs. Though the 22 nt sequence is a protein binding element, its functionality remains to be determined.

The Polyadenylation Signal

Another region that is well conserved in the PTH mRNA 3'-UTR is the AATAAA polyadenylation signal. In the bovine sequence, only a single AATAAA has been detected in the 3' noncoding region, whereas in the human and rat sequences two potential polyadenylation sites are found. The second AATAAA region in the human sequence is about 60 bases upstream from the first and has been suggested to have resulted from a gene duplication;[15] however, other than the AATAAA regions, there is little homology surrounding the two sites. Sequences analogous to the human upstream AATAAA are missing in both the bovine and rat sequences. No cDNAs were detected in which the upstream AATAAA was utilized as a polyadenylation signal; however the probability that these sites function as a polyadenylation signal cannot be ruled out. The rat sequence also has a second AATAAA site about 115 nucleotides earlier than the functional one. A single rat PTH mRNA was detected by Northern blot analysis ,[17] suggesting that only one polyadenylation site is used, and the size of the mRNA was consistent with the second AATAAA being the site. There is no direct evidence for the location of the 3' end by analysis of the rat PTH mRNA or cDNA.

The PTH Gene

The genes for human,[16] bovine,[14] rat[17] and mouse[31] PTH have each been cloned and characterized from genomic libraries in lambda phage. The human gene was isolated from a total human fetal DNA library prepared in λ phage Charon 4A. The library was screened initially by filter hybridization with human cloned cDNAs as a probe and later by the recombination selection method. The structure of the human gene was determined by the analysis of two overlapping clones. For the bovine gene, Southern analysis of the total bovine DNA showed that the PTH gene was present on an 8000 bp EcoR1 fragment. To clone the gene, bovine liver DNA was digested with EcoR1 and fragments in the range of 5,000 to 10,000 bp were isolated by sucrose gradient centrifugation. A partial library was then constructed by ligating the EcoR1 fragments to λ phage Charon 31 arms, which had been isolated after digestion with EcoR1. Several independent clones were isolated by plaque filter hybridization using cloned bovine PTH cDNA as probes. The rat gene was isolated from a λ phage Charon 4A rat liver DNA library produced by partial EcoR1 digestion of the rat DNA. Two independent positive plaques were obtained. The insert of each of the two phages contained the entire rat PTH gene. The sequence of the mouse gene was determined from a mouse genomic library.[31] One recombinant clone contained 14 kb of DNA, encompassing the entire PTH gene. The transcriptional unit spans 3.2 kb of genomic DNA, analogous to the human PTH gene.



	Intron A	Intron B
Rat	1600	110
Murine	NA	116
Bovine	1714	119
Human	3500	106
Macaca fascicularis	NA	104

Figure 4. Schematic representation of the rat and bovine PTH gene structure and the length of the introns. A) The PTH gene including exons 1-3 and introns A and B. Exons are indicated by the rectangles and the shaded areas indicate regions in the gene that code for preProPTH. B) The known sizes, in bp, of introns A and B in rat, mouse (murine), bovine, human and non-human primate (macaca fascicularis). The full length sequence of intron A is not available (NA) for mouse and non-human primate.

Structure of the Gene

The overall structure of the bovine or rat PTH gene is shown schematically in Figure 4A. In the human gene the larger intron A is aproximatly twice as long as this intron in bovine and rat (Fig. 4B). All sequenced genes contain two introns. The exact location of the bovine and human gene introns was determined by comparing the sequence of the gene to the previously determined cDNA structure. The location of intron A of the rat was determined by comparing the gene sequence with the sequence of cDNA to the 5' end of the mRNA. The cDNA was synthesized with reverse transcriptase using a synthetic pentadecamer as primer. Intron B in the rat gene was determined indirectly by homology of the sequence to the human and bovine cDNA sequences.

The locations of the introns are identical in each case as has been found with most other genes.[45] Intron A splits the 5' untranslated sequence five nucleotides before the initiator methionine codon (Fig. 2). Intron B splits the fourth codon of the region that codes for the pro sequence of preProPTH. The three exons that result, thus, are roughly divided into three functional domains. Exon I, 95 to 121 nucleotides long, contains the 5' untranslated region. ExonII, has 91 nucleotides and codes for the pre sequence, or signal peptide and exon III, 375 to 486 nucleotides long, codes for PTH as well as the 3' untranslated region. The structure of the PTH gene is thus consistent with the proposal that exons represent functional domains of the mRNA.[45]

The length of the introns in the species where the sequence is available is shown in Figure 4B. Although the introns are at the same location, the size of the large intron A in human is

about twice as large as those in the rat and bovine (Fig. 4B). It is of interest that the human gene is considerably longer in both intron A and the 3' untranslated region of the cDNA compared to the bovine, rat and mouse. Knowledge of the structures of other PTH genes from other species will be necessary in order to determine whether the extra sequence was inserted or is less susceptible to deletion in the human gene.

Both introns have the characteristic splice site elements. They have the GT-AG nucleotides at the 5' and 3' ends of the intron and the pyrimidine tract at the 3' end of the intron. The large intron A, has about 75% homology between the bovine and human PTH genes in over 200 bp of the intron, similar to the homology in the other non-translated regions of the genes. The rat intron A is only 55-57% homologous to the other species. The sequences of introns are generally only conserved at the *cis* elements essential for splicing and the relatively large homology for the PTH genes suggests that there may be some constrains on the basis of changes some distance from the intron /exon border.

The second exon, containing 106 and 121 nt in the human and bovine pre-mRNA, is much smaller and more homologous in size among the genes than intron A. The sequence of intron B is well conserved with homology of 74 and 68% of bovine/human and human/rat, respectively, but is relatively poorly conserved between the rat and bovine genes, with a homology of 49%.[1]

In each of the species, only a single PTH gene appears to be present. Extensive Southern blot analysis of bovine DNA with cloned PTH cDNA as probe produced single hybridizing bands for restriction enzymes that do not cut within the probe sequence.[13] The restriction map determined from the Southern analysis of bovine DNA was consistent with that of the cloned gene. With the exception of a single nucleotide in the 3' untranslated region, the sequence of the cloned cDNA was identical to the sequence of the exons in the gene. Less extensive Southern blot analysis of the human[16] and rat[17] genes also were consistent with a single gene per haploid genome. Furthermore, in the human studies, probes from the 5' and 3' ends of the cDNA both hybridized to the same sized fragment, and the strength of the signal from the genomic DNA was about the same as one gene-equivalent of the gene cloned in λ phage. Thus, the PTH gene is a single gene. The genes for PTH and PTHrP (PTH-related protein) are located in similar positions on sibling chromosomes 11 and 12. It is therefore likely that they arose from a common precursor by chromosomal duplication.

Initiation Site for RNA Transcription

As noted above in the discussion on the cDNA, the 5' termini of bovine PTH mRNA are heterogeneous. The large mRNAs contain a TATA sequence in the appropriate location to direct the synthesis of the smaller mRNAs. It was postulated that a second TATA would be found in the gene sequence 5' of the first one.[13] In both the human and bovine gene sequences, a second TATA sequence is present in the 5' flanking region about 25 base pairs from the first one in the appropriate position to direct the synthesis of the larger mRNAs. The heterogeneity of the 5' end of the bovine PTH mRNA, originally detected by reverse transcription of the mRNA,[13] was confirmed by S1 nuclease mapping.[14] The initiation sites for human PTH mRNA have not been determined directly, but were proposed[17] on the basis of analogy with the bovine sequence and the consensus TATA sequences. The presence of multiple functional TATA sequences has been reported for several other eukaryotic genes. The rat mRNA appears to be relatively homogeneous at the 5' terminus on the basis of both primed reverse transcriptions near the 5' end of the mRNA and S1 nuclease mapping.[17] The single initiation site for the rat mRNA can be explained by the changes in the rat sequence which alter the second downstream TATA sequence. The sequence, TATATATAAAA, in the human and bovine genes, is changed to TGCATATGAAA in the rat gene,[1] which is no longer a consensus TATA sequence. While this change seems the most likely explanation for the difference in length at the 5' termini

between the mRNAs, there are other changes that also occur in this region of the gene and may play a role.[17]

The smaller bovine PTH mRNAs are also heterogeneous with initiation occurring over a range of about eight nucleotides at the 5' terminus. The second TATA sequence in the bovine sequence is unusual since the sequence TA is repeated five times, and thus the TATA-like sequence is spread over 12 base pairs. This may result in a less rigorous delineation of the appropriate start site.

The conclusion that the 5' end of bovine mRNA is heterogeneous has not been conclusively proven. Both the S1 nuclease mapping and the primed reverse transcriptase techniques require that the mRNA is intact and not degraded. Since in the studies described above, it was not demonstrated that all the mRNA had a 5' methylguanosine cap and thus was intact, the possibility that heterogeneity was introduced during isolation of the mRNA cannot be excluded. However, the additional indirect evidence provided by the presence of two TATA sequences considerably strengthens the theory that two regions are utilized for initiation of transcripts.

The 5' Flanking Region

The three PTH genes of human, bovine and rat show homology in the first 200 bp upstream of the RNA initiation site of the 5' flanking region.[1] The homology in this region is similar to that in the 5' untranslated region of the mRNAs.[1] There are few stretches of sequence in the 5' flanking region that are completely conserved in all three sequences except for the TATA sequences. A C-rich sequence, GCACCGCCC, about 75 bp to the 5' side of the upstream TATA sequence is present in all three sequences, and an AT-rich region of about 25 bp immediately prior to this C-rich region is strongly conserved. A sequence, CAGAGAA, about 25 bp to the 5' side of the TATA sequence, is also present in all three sequences. No CAAT sequence is present 5' of the TATA sequences. In the bovine gene, an extraordinary stretch of almost 150 nucleotides, located from 250 to 400 nucleotides before the transcript initiator, consists primarily of alternating AT.[1] A similar region is not present in the rat gene, suggesting it is not critical for the function of the gene. There are defined functional response elements in the 5'-flanking region that regulate PTH gene transcription, such as the vitamin D response element (VDRE) and the cyclic AMP response element (CRE) that are discussed in detail in the chapters by Kel et al and Silver et al.

The 3' Flanking Region

In the 3' flanking region, again there is also considerable homology between the bovine and human sequences. A small inverted repeat region, that could form a hairpin loop in the transcript, is followed by a stretch of 7 Ts. There is no direct evidence that this region serves as a transcriptional stop signal in the PTH genes. A difference in the stem in the human compared to bovine is matched by a second change in the human that maintains the base pairing in the stem. A similar sequence is not present in the approximately 110 bp of 3' flanking sequence reported for the rat sequence. The rat in fact has little homology with either of the other two sequences beyond the polyadenylation signal.[1] This is surprising in view of the homology retained between the rat PTH gene and the other genes in the 5' flanking and intron regions. Perhaps the polyadenylation signal for the rat sequence is derived from a different region of the gene, which was moved into its present position by a deletion of sequence or translocation. Large gaps must be introduced into the bovine and rat sequences just prior to the polyadenylation signal supporting the idea that this may be a relatively unstable region of the gene.[1]

Overall, the PTH genes are typical eukaryotic genes that contain the consensus sequences for initiation of RNA synthesis, RNA splicing, and polyadenylation. The PTH genes appear to

be represented only once in the haploid genome. Perhaps the most striking characteristic of the DNA in the region of the genes is its stability. In addition, regions that diverge rapidly in other genes are relatively stable in the PTH genes, particularly between the human and bovine sequences and to a lesser extent with the rat sequence. Thus, considerable homology is observed between 5' and 3' flanking and untranslated regions, internal regions of introns, and potential sites for silent changes in the coding region. Since these regions that do not change the amino acid sequence have been estimated to diverge at a rate of 1% 10^6 years, relatively low homologies would be expected from these sequences that diverged about 60 to 80 x 10^6 years ago.[1] Whether this conservation of sequence occurs because the genes happen to be present in a region of the chromosome that is usually stable or reflects some functional constraints inherent in the PTH gene, remains to be elucidated.

The rat and mouse sequences are considerably less homologous to the human and bovine sequences than these sequences are to each other. This observation is difficult to explain, since evolutionarily each of the sequences is about equidistant from another. Potentially, differences in the physiology or nutrition of calcium in the rat and mouse compared to the other two species may have resulted in increased acceptance of mutations in the rat PTH gene.

Chromosomal Location of the Human PTH Gene

The location of the human PTH gene on chromosome 11 has been determined independently by two groups. The assignments were made by screening panels of human-mouse [46] or human-mouse and human-Chinese hamster cell [47] hybrids with a human cDNA clone or a cloned fragment of human genomic DNA. The PTH gene was further localized to the short arm of the chromosome 11 by analysis of human-mouse hybrids with various translocations.[46] The short arm of chromosome 11 contains several other polymorphic genes including the β globin gene cluster, insulin, and the human oncogene Harvey *ras* (C-Ha-ras-1).[48,49] The polymorphisms in these genes and PTH were used to determine whether the genes are genetically linked and their order on the chromosome. In addition to these genes, the gene for calcitonin has also been mapped to the short arm of chromosome 11. Thus, the short arm contains genes for both of the polypeptide hormones that regulate calcium metabolism. Whether this is a mere coincidence or is somehow related to the evolution or regulation of these calcium regulating genes remains a matter of speculation. The porcine PTH gene was localized to chromosome 14q25-q28 by in situ hybridization[50] and the equus gene is on chromosome 11p15.3.[29]

Summary

The PTH genes and cDNAs have been isolated and characterized in 10 species. The gene contains two introns, which are in the same position in each species, and dissect the gene into 3 exons that code, respectively, for the 5' untranslated region, the signal peptide, and PTH plus the 3' untranslated region. The mRNAs contain a 7-methyl quanosine cap at the 5' terminus and a polyadenylation signal at the 3' terminus. They are about twice as long as necessary to code for preProPTH. The 5' termini of the bovine and human mRNAs are heterogeneous at the 5' terminus, the basis of which is two TATA sequences in the 5' flanking regions of the gene. In contrast, the mouse and rat gene contain a single TATA sequence and the mRNA has a single 5' terminus. The initial translational product of the mRNA is preProPTH, and the pre-peptide of 25 amino acids and the pro sequence of 6 amino acids are removed by two proteolitic cleavages.

The mRNAs are very homologous in the region that codes for preProPTH. But substantial homology is also retained in the mRNA untranslated regions and flanking regions and introns, where sequences are available. The gallus PTH mRNA is the most distant sequence of the PTH mRNAs. In the PTH mRNA the 3'-UTR is the region less conserved

amongst species. However two protein binding elements in the 3'-UTR were identified and show high homology. One of these elements is the distal 26 nt *cis* acting functional element that has been shown to mediate the regulation of PTH mRNA stability in response to changes in serum calcium and phosphate. This element is expressed in the 3'-UTR of rat, man, dog, cat and mouse. An additional proximal element of 22 nt is present in the 3'-UTR of bovine, pig, macaca, horse and also in man, cat, dog. This element binds cytosolic proteins but its function has not been demonstrated. The conservation of such elements in the 3'-UTR suggests that they represent an evolutionary conserved function. PTH is central to normal calcium homeostasis and bone strength and the PTH peptide is highly conserved amongst species apart from Gallus. This conservation is evident in the coding sequence but also, to a less extent in the 5'- and 3'-UTRs.

Acknowledgements

This chapter quotes widely, with the author's permission, from the outstanding review by Byron Kemper where there is a detailed analysis of the bovine, rat and human PTH genes that were published at that time (ref. 1). We are extremely grateful to him both for his contribution and his generosity.

References

1. Kemper B. Molecular biology of parathyroid hormone. CRC Crit Rev Biochem 1986; 19:353-379.
2. Brewer HBJr, Ronan R. Bovine parathyroid hormone: Amino acid sequence. Proc Natl Acad Sci USA 1970; 67:1862-1869.
3. Cohn DV, MacGregor RR, Chu LL et al. Calcemic fraction-A: Biosynthetic peptide precursor of parathyroid hormone. Proc Natl Acad Sci USA 1972; 69:1521-1525.
4. Kemper B, Habener JF, Potts JT Jr et al. Proparathyroid hormone: Identification of a biosynthetic precursor to parathyroid hormone. Proc Natl Acad Sci USA 1972; 69:643-647.
5. Habener JF, Amherdt M, Ravazzola M et al. Parathyroid hormone biosynthesis: Correlation of conversion of biosynthetic precursors with intracellular protein migration as determined by electron microscope autoradiography. J Cell Biol 1979; 80:715-731.
6. Kemper B, Habener JF, Mulligan RC et al. Pre-proparathyroid hormone: A direct translation product of parathyroid messenger RNA. Proc Natl Acad Sci USA 1974; 71:3731-3735.
7. Milstein C, Brownlee GG, Harrison TM et al. A possible precursor of immunoglobulin light chains. Nat New Biol 1972; 239:117-120.
8. Dorner A, Kemper B. Conversion of preproparathyroid hormone to proparathyroid hormone by dog pancreatic microsomes. Biochemistry 1978; 17:5550-5555.
9. Habener JF, Kamper B, Potts JT Jr et al. Preproparathyroid hormone identified by cell-free translation of messenger RNA from hyperplastic human parathyroid tissue. J Clin Invest 1975; 56:1328-1333.
10. Kemper B, Habener JF, Ernst MD et al. Pre-proparathyroid hormone: analysis of radioactive tryptic peptides and amino acid sequence. Biochemistry 13-1-1976; 15:15-19.
11. Habener JF, Rosenblatt M, Kemper B et al. Pre-proparathyroid hormone; amino acid sequence, chemical synthesis, and some biological studies of the precursor region. Proc Natl Acad Sci U S A 1978; 75:2616-2620.
12. Kronenberg HM, McDevitt BE, Majzoub JA et al. Cloning and nucleotide sequence of DNA coding for bovine preproparathyroid hormone. Proc Natl Acad Sci USA 1979; 76:4981-4985.
13. Weaver CA, Gordon DF, Kemper B. Nucleotide sequence of bovine parathyroid hormone messenger RNA. Mol Cell Endocrinol 1982; 28:411-424.
14. Weaver CA, Gordon DF, Kissil MS et al. Isolation and complete nucleotide sequence of the gene for bovine parathyroid hormone. Gene 1984; 28:319-329.
15. Hendy GN, Kronenberg HM, Potts JTJ et al. Nucleotide sequence of cloned cDNAs encoding human preproparathyroid hormone. Proc Natl Acad Sci USA 1981; 78:7365-7369.

16. Vasicek TJ, McDevitt BE, Freeman MW et al. Nucleotide sequence of the human parathyroid hormone gene. Proc Natl Acad Sci USA 1983; 80:2127-2131.
17. Heinrich G, Kronenberg HM, Potts JT Jr et al. Gene encoding parathyroid hormone. Nucleotide sequence of the rat gene and deduced amino acid sequence of rat preproparathyroid hormone. J Biol Chem 1984; 259:3320-3329.
18. Schmelzer HJ, Gross G, Widera G et al. Nucleotide sequence of a full-length cDNA clone encoding preproparathyroid hormone from pig and rat. Nucleic Acids Res 1987; 15:6740-6741.
19. Watson ME. Compilation of published signal sequences. Nucleic Acids Res 11-7-1984; 12:5145-5164.
20. Mead DA, Skorupa ES, Kemper B. Single stranded DNA SP6 promoter plasmids for engineering mutant RNAs and proteins: synthesis of a 'stretched' preproparathyroid hormone. Nucleic Acids Res 1985; 13:1103-1118.
21. Majzoub JA, Rosenblatt M, Fennick B et al. Synthetic pre-poroparathyroid hormone leader sequence inhibits cell-free processing of placental, parathyroid, and pituitary pre-hormones. J Biol Chem 1980; 255:11478-11483.
22. Jackson RC, Blobel G. Post-translational cleavage of presecretory proteins with an extract of rough microsomes from dog pancreas containing signal peptidase activity. Proc Natl Acad Sci USA 1977; 74:5598-5602.
23. Habener JF, Rosenblatt M, Dee PC et al. Cellular processing of pre-proparathyroid hormone involves rapid hydrolysis of the leader sequence. J Biol Chem 1979; 254:10596-10599.
24. Divieti P, John MR, Juppner H et al. Human PTH-(7-84) inhibits bone resorption in vitro via actions independent of the type 1 PTH/PTHrP receptor. Endocrinol 2002; 143:171-176.
25. Perler F, Efstratiadis A, Lomedico P et al. The evolution of genes: the chicken preproinsulin gene. Cell 1980; 20:555-566.
26. Stolarsky L, Kemper B. Characterization and partial purification of parathyroid hormone messenger RNA. J Biol Chem 1978; 253:7194-7201.
27. Rosol TJ, Steinmeyer CL, McCauley LK et al. Sequences of the cDNAs encoding cannine parathyroid hormone-related protein and parathyroid hormone. Gene 1995; 160:241-243.
28. Khosla S, Demay M, Pines M et al. Nucleotide sequence of cloned cDNAs encoding chicken preproparathyroid hormone. J Bone Miner Res 1988; 3:689-698.
29. Caetano AR, Shiue YL, Lyons LA et al. A comparative gene map of the horse (Equus caballus). Genome Res 1999; 9:1239-1249.
30. Malaivijitnond S, Takenaka O, Anukulthanakorn K et al. The nucleotide sequences of the parathyroid gene in primates (suborder Anthropoidea). Gen Comp Endocrinol 2002; 125:67-78.
31. He B, Tong TK, Hiou-Tim FF et al. The murine gene encoding parathyroid hormone: genomic organization, nucleotide sequence and transcriptional regulation. J Mol Endocrinol 2002; 29:193-203.
32. Nutley MT, Parimi SA, Harvey S. Sequence analysis of hypothalamic parathyroid hormone messenger ribonucleic acid. Endocrinology 1995; 136:5600-5607.
33. Gordon DF, Kemper B. Synthesis, restriction analysis, and molecular cloning of near full length DNA complementary to bovine parathyroid hormone mRNA. Nucleic Acids Res 1980; 8:5669-5683.
34. Weaver CA, Gordon DF, Kemper B. Introduction by molecular cloning of artifactual inverted sequences at the 5' terminus of the sense strand of bovine parathyroid hormone cDNA. Proc Natl Acad Sci USA 1981; 78:4073-4077.
35. Suzuki Y, Ishihara D, Sasaki M et al. Statistical analysis of the 5' untranslated region of human mRNA using "Oligo-Capped" cDNA libraries. Genomics 2000; 64:286-297.
36. Habener JF, Kemper B, Potts JT Jr et al. Parathyroid mRNA directs the synthesis of preproparathyroid hormone and proparathyroid hormone in the Krebs ascites cell-free system. Biochem Biophys Res Commun 1975; 67:1114-1121.
37. Kozak M. An analysis of vertebrate mRNA sequences: intimations of translational control. J Cell Biol 1991; 115:887-903.
38. Kozak M. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem 1991; 266:19867-19870.
39. Kwon HS, Lee DK, Lee JJ et al. Posttranscriptional regulation of human ADH5/FDH and Myf6 gene expression by upstream AUG codons. Arch Biochem Biophys 2001; 386:163-171.

40. Moallem E, Silver J, Kilav R et al. RNA protein binding and post-transcriptional regulation of PTH gene expression by calcium and phosphate. J Biol Chem 1998; 273:5253-5259.
41. Sela-Brown A, Silver J, Brewer G et al. Identification of AUF1 as a parathyroid hormone mRNA 3'-untranslated region binding protein that determines parathyroid hormone mRNA stability. J Biol Chem 2000; 275:7424-7429.
42. Guhaniyogi J, Brewer G. Regulation of mRNA stability in mammalian cells. Gene 2001; 265:11-23.
43. Kilav R, Silver J, Naveh-Many T. A conserved cis-acting element in the parathyroid hormone 3'-untranslated region is sufficient for regulation of RNA stability by calcium and phosphate. J Biol Chem 2001; 276:8727-8733.
44. Naveh-Many T, Bell O, Silver J et al. Cis and trans acting factors in the regulation of parathyroid hormone (PTH) mRNA stability by calcium and phosphate. FEBS Lett 2002; 529:60-64.
45. Gilbert W. Why genes in pieces? Nature 1978; 271:501.
46. Naylor SL, Sakaguchi AY, Szoka P et al. Human parathyroid hormone gene (PTH) is on short arm of chromosome 11. Somatic Cell Genet 1983; 9:609-616.
47. Mayer H, Breyel E, Bostock C et al. Assignment of the human parathyroid hormone gene to chromosome 11. Hum Genet 1983; 64:283-285.
48. Gerhard DS, Kidd KK, Kidd JR et al. Identification of a recent recombination event within the human beta-globin gene cluster. Proc Natl Acad Sci USA 1984; 81:7875-7879.
49. Fearon ER, Antonarakis SE, Meyers DA et al. c-Ha-ras-1 oncogene lies between beta-globin and insulin loci on human chromosome 11p. Am J Hum Genet 1984; 36:329-337.
50. Ji Q, Zhang X, Chen Y. [Localization of PTH gene on pig chromosome 14q25-q28 by in situ hybridization]. Yi Chuan Xue Bao 1997; 24:218-222.horse and non-human primate (macaca) PTH mRNA were not included in this study because only partial sequences of these RNAs are available. The same Phylogenetic tree is also obtained when the 3'-UTR sequences are analized separately. Interestingly, based on amino acid sequence, the bovine and pig were grouped closest to canis and felis but not by RNA sequence. This corresponds to the presence of the distal and proximal protein-binding elements in the 3'-UTRs.Macaca fascicularis Af130257: gene, complete cds mRNA (partial, deduced).

http://www.springer.com/978-0-306-47847-5

Molecular Biology of the Parathyroid
Naveh-Many, T. (Ed.)
2005, XIV, 199 p. 71 illus., 5 illus. in color., Hardcover
ISBN: 978-0-306-47847-5