Proc. Natl. Acad. Sci. USA
Vol. 78, No. 7, pp. 4073–4077, July 1981
Biochemistry

Introduction by molecular cloning of artifactual inverted sequences at the 5′ terminus of the sense strand of bovine parathyroid hormone cDNA

(DNA sequence determination/reverse transcriptase/mRNA sequence determination)

CHRISTINE A. WEAVER, DAVID F. GORDON, AND BYRON KEMPER*

Department of Physiology and Biophysics and School of Basic Medical Sciences, University of Illinois, Urbana, Illinois 61801

Communicated by Donald F. Steiner, March 16, 1981

ABSTRACT **To study the structure and function of the gene for parathyroid hormone, we obtained recombinant plasmids containing bovine parathyroid hormone cDNA. The nucleotide sequence at the 5′ terminus (relative to the sense strand) of the cDNA insert in a recombinant plasmid, pPTHi4, was different from that previously reported for the bovine parathyroid hormone cDNA insert of another recombinant plasmid, pPTHm1 [Kronenberg, H. M., McDevitt, B. F., Majzoub, J. A., Nathans, J., Sharp, P. A., Potts, J. T., Jr. & Rich, A. (1979) *Proc. Natl. Acad. Sci. USA* 76, 4981–4985]. The first 50 nucleotides of the pPTHm1 insert were an inverted complement of nucleotides 2–51 of the pPTHi4 insert. The cDNA insert of another plasmid, pPTHi8, contained a sequence identical to nucleotides 2–51 of the pPTHi4 insert but also contained an additional 42 bases at the 5′ terminus. The first 41 bases of the pPTHi8 insert were an inverted repeat of an internal sequence of the pPTHi4 insert corresponding to nucleotides 184–224. Restriction endonuclease analysis of pPTHi8 indicated that the internal sequence corresponding to this region was retained. The nucleotide sequence of a restriction fragment hybridized to parathyroid hormone mRNA and extended toward the 5′ terminus of the mRNA with reverse transcriptase confirmed that the sequence at the 5′ terminus of the pPTHi4 insert was an accurate copy of the parathyroid hormone mRNA sequence. These data suggest that two types of sequence rearrangements may occur during molecular cloning of cDNA: (*i*) an inversion of a sequence at the 5′ terminus, as occurred in pPTHm1, and (*ii*) an inverted repeat of an internal sequence, as occurred in pPTHi8.**

The construction of recombinant plasmids containing cDNA to produce large amounts of pure nucleotide sequence corresponding to mRNA is an important and common procedure. The accuracy of the enzymatic reactions involving avian myeloblastosis virus (AMV) reverse transcriptase and *Escherichia coli* DNA polymerase was demonstrated in early studies of cloned globin cDNA, in which the sequences determined for the cloned cDNA agreed with the sequence of the mRNA (1). In these studies, however, transversions of dAT base pairs were observed in the dA-T homopolymer extensions added during the cloning procedures. Recently, more substantial rearrangements near the 5′ terminus (relative to the sense strand) of the cloned cDNA inserts have been reported. Restriction analyses of several recombinant plasmids containing cDNA inserts for human interferon (2) or mouse immunoglobin heavy chains (3) indicated that inverted repeats occur near the 5′ terminus of the cDNA. Comparison of the sequences of the cDNA inserts for chicken β-globin (4) and fibronectin (5) in recombinant plasmids also showed rearrangement of sequence near the 5′ terminus. A rat insulin cDNA insert (6) and a chicken ovalbumin cDNA insert (7) contained inversions at the 5′ terminus, when compared with the sequences of the cloned genomic genes.

As part of our studies on the regulation of the biosynthesis of parathyroid hormone (PTH), we have isolated several recombinant plasmids containing bovine PTH cDNA inserts. Comparison of the sequences of these plasmids with each other and with the sequence previously reported for a bovine PTH cDNA insert in the plasmid, pPTHm1 (8), showed different sequences at the 5′ terminus of the cDNA inserts. In the present study, we compare the nucleotide sequences of these plasmids at the 5′ terminus with the sequence of PTH mRNA and show that artifactual inversions or inverted repeats occur in some of the cDNA inserts.

METHODS

Enzymes. AMV reverse transcriptase was obtained from J. W. Beard (Life Sciences, St. Petersburg, FL) and the Division of Program Resources Biological Carcinogenesis Branch, National Cancer Institute. Restriction endonucleases were obtained from Bethesda Research Laboratories (Rockville, MD), New England BioLabs, or Miles. Terminal deoxynucleotidyltransferase was obtained from PL Biochemicals. T4 polynucleotide kinase was supplied by O. Uhlenbeck (University of Illinois, Urbana, IL), and bacterial alkaline phosphatase was from Sigma.

Isolation of PTH cDNA Clones. PTH mRNA was isolated to ≈50% purity as described (9). Double-stranded PTH cDNA was synthesized by the method of Wickens *et al.* (10) with modifications (11) and inserted into the *Pst* I site of pBR322 by dG-dC homopolymer extension techniques (12). Plasmids containing PTH cDNA were detected by restriction analysis of the cloned cDNA inserts and comparison with the restriction map of PTH cDNA (11). Before December 1979, experiments involving recombinant plasmids were carried out in accordance with the P2/EK2 containment guidelines of the National Institutes of Health. After that time, P1/EK1 containment procedures were used.

Nucleic Acid Sequence Determination. The nucleotide sequences of the 5′-terminal regions of the cDNA inserts were determined as described by Maxam and Gilbert (13). The sense strands were sequenced from a radiolabeled *Hpa* II site in the pBR322 DNA sequence, ≈50 nucleotides from the 5′ terminus of the cDNA inserts. The antisense strands were sequenced from a radiolabeled *Sau* 3A site in the cDNA, 160 nucleotides from the 5′ terminus of the cDNA insert. To more directly sequence the mRNA, a restriction fragment was isolated from the

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: PTH, parathyroid hormone; AMV, avian myeloblastosis virus.

* To whom reprint requests should be addressed.

cDNA insert of pPTHi4 that extended from nucleotide 76, an *Mbo* II cleavage site, to nucleotide 160, a *Sau* 3A site. The fragment was labeled with $^{32}PO_4$ at the 5'-terminal nucleotide of the *Sau* 3A site. About 2.5 pmol of the cDNA fragment and 4.0 pmol of PTH mRNA (50% pure) in 36 μl of 10 mM 1,4-piperazinediethanesulfonic acid, pH7/50 mM EDTA/40 mM NaCl in 70% formamide were heated to 80°C for 2 minutes and then incubated at 52°C, 45°C, and 40°C for 20 min each. The DNA·RNA hybrid was collected by ethanol precipitation. The pellet was dissolved in 25 μl of 50 mM Tris·HCl, pH 8.3 (42°C)/40 mM KCl/10 mM $MgCl_2$/30 mM 2-mercaptoethanol/200 μM each of dATP, dCTP, dGTP, and TTP/4 mM sodium pyrophosphate containing actinomycin D at 156 μg/ml and 30 units of AMV reverse transcriptase. After incubation at 42°C for 30 min, the reaction mixture was extracted with phenol saturated with 50 mM Tris·HCl, pH 8.0; residual phenol in the aqueous phase was removed by extraction with diethyl ether, and nucleic acids were precipitated twice with ethanol. The pellet was washed with 95% ethanol and dried under reduced pressure. The sequence of the extended fragment was determined as described by Maxam and Gilbert (13).

RESULTS

Detection of an Inversion at the 5' Terminus of Cloned PTH cDNA. The plasmid, pPTHi4, contains a PTH cDNA insert of ≈800 bases including the homopolymer extensions of ≈50 nucleotides added during the cloning procedures (11). As PTH mRNA contains ≈750 bases (9), this plasmid contains a near full-length PTH cDNA insert. The nucleotide sequence of the first 66 nucleotides at the 5'-terminal noncoding region of the sense strand is shown in Fig. 1*A*. The initiator codon begins at nucleotide 93. The sequence of this region was determined unambiguously three times, twice for the sense strand and once for the antisense strand. Also shown in Fig. 1*A* is the sequence reported earlier for the PTH cDNA insert of pPTHm1 (8). Beginning at nucleotide 44 for the pPTHi4 insert and nucleotide 54 for the pPTHm1 insert, the two sequences are identical as shown and in fact are identical through the coding region to the last nucleotide in pPTHm1, nucleotide 470 (not shown). The 5'-terminal regions of the two sequences, however, are quite different. Although the two cloned cDNAs could have been derived from two variants of PTH mRNA, the identity of most of the sequence suggested that an artifact had been introduced at the 5'-terminal region of one clone.

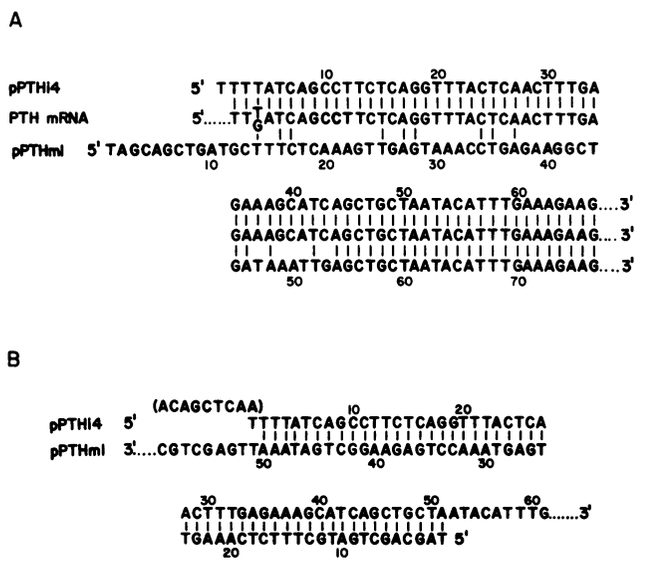

FIG. 1. Comparison of nucleotide sequences of the sense strands of the cDNA inserts of pPTHm1 and pPTHi4 with the sequence of PTH mRNA. The sequences of the inserts were determined from restriction fragments by the method of Maxam and Gilbert (12). PTH mRNA sequence was determined from a restriction fragment hybridized to PTH mRNA and extended with AMV reverse transcriptase as described in *Methods*. (*A*) identical bases between the cDNA inserts and PTH mRNA are indicated by vertical lines. (*B*) pPTHm1 sequence drawn antiparallel to that of pPTHi4; complementary bases are indicated by vertical lines. The nucleotide sequence in parenthesis at the 5' terminus of pPTHi4 is the probable sequence of the sense strand, based on identical sequences in pPTHi7 and pPTHi8.

To determine which of these two sequences was representative of PTH mRNA, we determined the sequences of the 5'-terminal regions of four additional plasmids containing PTH cDNA inserts and of a restriction fragment hybridized to PTH mRNA and extended toward the 5' terminus with reverse transcriptase. The cDNA inserts of these additional plasmids contained some or all of the disputed sequence between pPTHm1 and pPTHi4; in each case, the sequences agreed with that of pPTHi4 (not shown). To sequence PTH mRNA more directly, a restriction fragment corresponding to nucleotides 76–160 of the pPTHi4 insert that was labeled with $^{32}PO_4$ at nucleotide 160 was isolated. The restriction fragment was hybridized to PTH mRNA, extended with reverse transcriptase toward the 5' terminus of the mRNA, and then sequenced by the Maxam and Gilbert technique. The sequence could be determined to nucleotide 2 of pPTHi4, although larger fragments were present, indicating that additional nucleotides are present in the mRNA at the 5' terminus. The sequence as shown in Fig. 1*A* was in agreement with that of the pPTHi4 insert; thus, the 5' terminal sequence reported for the pPTHm1 insert is incorrect. As shown in Fig. 1*B*, nucleotides 1–50 of the pPTHm1 insert are a perfect inverted complement of nucleotides 2–51 of the pPTHi4 insert. If the inversion generated the artifactual differences, it should extend to at least guanine-53 of pPTHm1, which should be a cytosine. Shown in parenthesis above the sequence for pPTHi4 is an additional sequence of PTH mRNA based on a common sequence in this region for two other cloned cDNA inserts, pPTHi7 and pPTHi8. From these data, the inverted complement potentially extends to guanine-58. It is interesting that the region adenine-54 through guanine-58 could be generated either by inversion or by correctly copying the mRNA. In addition, the sequence present in the pPTHm1 insert could not be generated simply by an inversion of sequence, as the first three nucleotides at the 5' terminus are an inverted repeat of nucleotides 59–61 of the pPTHm1 sequence. The artifact in pPTHm1 thus is an inversion of 53–58 nucleotides with an inverted repeat of an internal sequence of 3–8 nucleotides at the 5' end, depending on exactly where the inversion occurred.

Detection of an Inverted Repeat at the 5' Terminus of Cloned PTH cDNA. In the process of analyzing other cloned PTH cDNAs to determine whether pPTHm1 or pPTHi4 contained the correct sequence at the 5' terminus, we found that the pPTHi8 cDNA insert had 42 more bases at the 5' terminus than the pPTHi4 cDNA insert. As shown in Fig. 2, the first 41 nucleotides of the pPTHi8 insert are an inverted repeat of an internal sequence of the pPTHi4 insert corresponding to nucleotides 184–224. The sequences are identical from thymine-2 of the pPTHi4 insert and thymine-44 of the pPTHi8 insert as far as the sequence of the pPTHi8 insert has been determined. By comparison with the sequence of the insert of another plasmid, pPTHi7, the sequence beginning with adenine-35 in the pPTHi8 insert appears to be correct with respect to PTH mRNA (not shown). The region adenine-35 to cytosine-41 thus can be generated either by an inverted repeat of the internal sequence

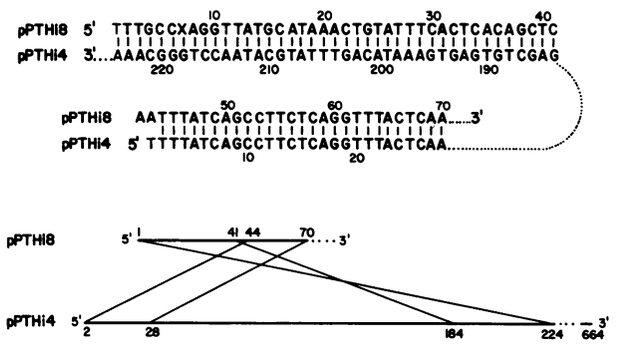

FIG. 2. Comparison of sequences of the sense strands of the cDNA inserts of pPTHi4 and pPTHi8. The upper set of sequences are drawn antiparallel to the lower to illustrate the inverted repeat at the 5′ terminus of the pPTHi8 insert, which results in complementarity between nucleotides 1–41 of the pPTHi8 cDNA insert and nucleotides 184–224 of the pPTHi4 insert. The lower set of sequences illustrates the identity of sequence beginning with nucleotides 2 of pPTHi4 and 44 of pPTHi8. The diagram at the bottom illustrates the relative location of the sequences of pPTHi8 in the sequence of pPTHi4. The crossed lines indicate an inversion.

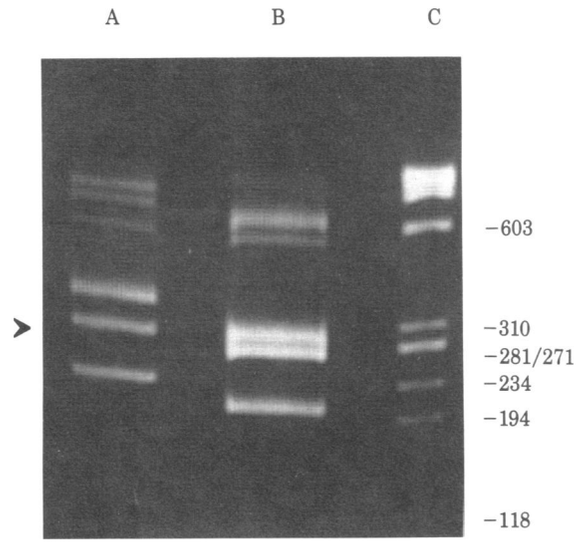

FIG. 3. Restriction analysis of cDNA inserts of pPTHi4 and pPTHi8. The two plasmids were digested with *Hpa* II, and the fragments containing the cDNA inserts were isolated. The fragments were digested with *Sau* 3A and analyzed by electrophoresis in 5% acrylamide gels. The DNA bands were stained with ethidium bromide. The marker indicates the 300-base-pair internal fragment of the pPTHi4 insert that contains the sequence complementary to the first 41 nucleotides of the pPTHi8 insert. The numbers indicate the number of base pairs in fragments of ϕX174 DNA digested with *Hae* III. Lanes: A, pPTHi4; B, pPTHi8; C, *Hae* III-digested ϕX174 DNA.

or by correctly copying the mRNA.

To distinguish whether the inversion at the 5′ terminus of the pPTHi8 insert resulted from a rearrangement of the sequence from the internal region to the 5′ terminus or was an inverted repeat or copy of the internal sequence, the cDNA insert was analyzed by restriction analysis. Inserts of pPTHi4 and pPTHi8 with ≈50 nucleotides of plasmid sequence at either end were isolated after digestion of the plasmids with *Hpa* II. Cleavage of the pPTHi4 insert fragment with *Sau* 3A produces fragments of 230, 300, and 340 base pairs, with the internal 300-base-pair fragment containing the region that was inverted at the 5′ terminus of the pPTHi8 insert. The sizes of the expected fragments for the pPTHi8 insert were 270 and 200 base pairs at the 5′ and 3′ termini, respectively, and either 300 base pairs if the inverted sequence was still present in its normal internal location or 260 base pairs if it was absent. As shown in Fig. 3, a 300-base-pair fragment was present in the pPTHi8 insert, which was the same size as the corresponding fragment in pPTHi4, confirming that the sequence at the 5′ terminus of pPTHi8 is an inverted repeat of an internal sequence and not a rearrangement.

DISCUSSION

We have shown that segments of bovine PTH cDNA may be rearranged when the cDNA is cloned by recombinant DNA techniques. These rearrangements are either primarily inversions of sequences at the 5′ terminus or inverted copies of internal sequences. We have examined the sequences of the 5′ terminus of the bovine PTH cDNA inserts of several plasmids. In addition, the structure of PTH mRNA at the 5′ terminus was determined more directly by sequencing a restriction fragment hybridized to the mRNA and extended from the 5′ terminus with AMV reverse transcriptase. The sequence of these plasmids and that of PTH mRNA at the 5′ terminus disagreed with the sequence reported previously for a bovine PTH cDNA insert in pPTHm1 (8). The first 50 nucleotides of the sequence of the sense strand of the pPTHm1 insert were an inverted complement of the corresponding sequence of the pPTHi4 insert. In addition, another plasmid, pPTHi8, contained an insert with an inverted repeat of an internal sequence at the 5′ terminus, which was also present in its normal internal location. These two kinds of inversions share the property that several nucleotides at the 3′ terminus of the inverted sequence could result either from the inversion process or from correct reverse transcription of the mRNA.

Two conclusions about the structure of PTH mRNA based on the pPTHm1 insert sequence are no longer valid because of the inverted 5′-terminal sequence. A stable secondary structure at the 5′-terminal region predicted from the pPTHm1 insert involved the inverted repeat of three nucleotides at the 5′ terminus and is not possible in the correct sequence derived from the pPTHi4 insert, although another stable secondary structure can be drawn (not shown). In the pPTHm1 insert, an AUG codon was present before the actual initiator codon for preproparathyroid hormone. This premature AUG codon was present in the inversion of the pPTHm1 insert, but no premature AUG codons are present in the sequence of the pPTHi4 insert. For the sequence of PTH mRNA currently known, the initiator codon of PTH mRNA, like that of nearly all other eukaryotic mRNAs, is the first AUG codon and is therefore consistent with the hypothesis that protein initiation starts at the first AUG codon from the 5′ terminus (14).

Similar rearrangements at 5′ termini of cloned cDNA sequences for other mRNAs have been reported, as have small inverted sequences in chicken β-globin cDNA inserts. However, these cDNAs were inserted into plasmids by blunt-end ligation techniques, and the error was postulated to be introduced by DNA polymerase during the process of making the ends of the cDNA perfectly blunt. The artifacts described in this paper and for the cDNAs described below involved incorporation of the cDNA into plasmids by techniques that did not involve the use of DNA polymerase to produce blunt ends. A rat insulin cDNA insert had an inversion of 39 nucleotides and a chicken ovalbumin cDNA insert had an inversion of 56 nucleotides, both of which were similar to that in the insert of pPTHm1 (6, 7). Larger inversions (>100 bases) have been detected by restriction analysis of interferon cDNA inserts (2). Detailed analysis of fibronectin cDNA (5) showed inverted re-

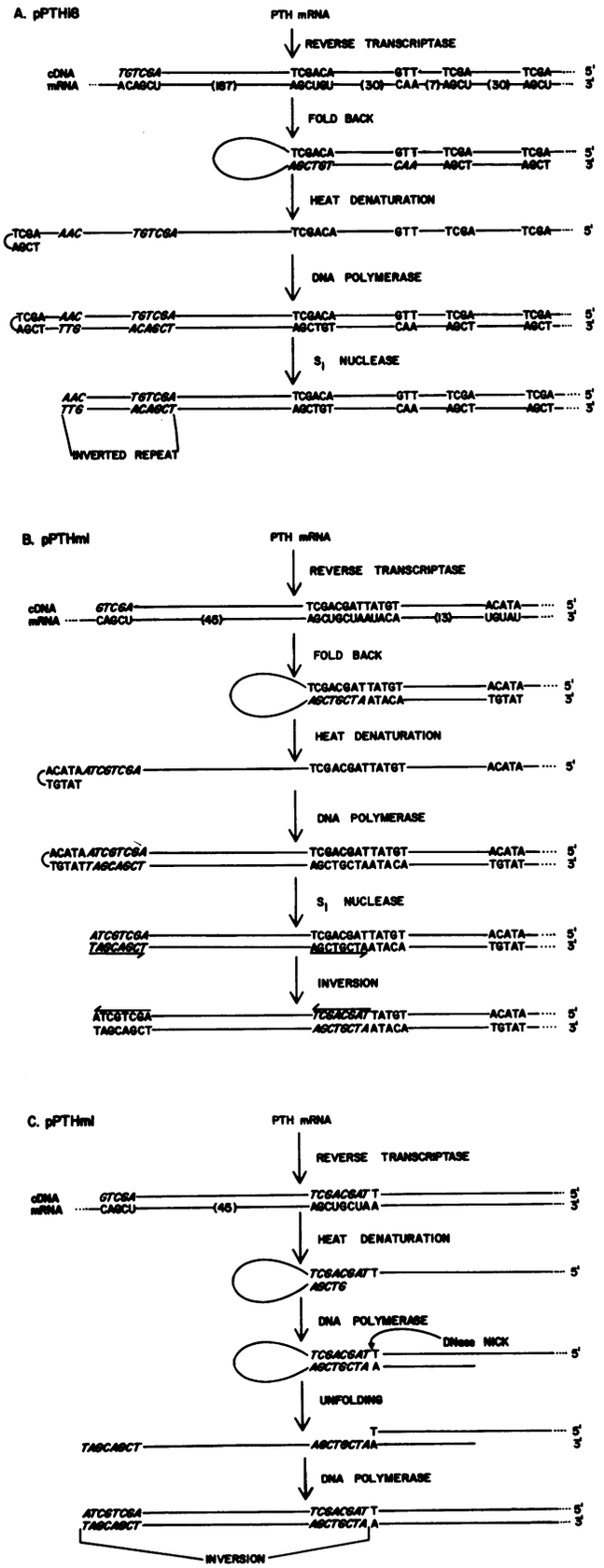

FIG. 4. Possible models for generation of inverted sequences in pPTHm1 and pPTHi8. The key step in models A and B is the folding back of the single-stranded DNA during reverse transcription with the generation of some double-stranded DNA. This is followed by heat denaturation, synthesis of the second strand by DNA polymerase, and cleavage of the hairpin loop and some nibbling of the double-stranded cDNA by nuclease S1. The numbers in parenthesis in the mRNA indicate the number of bases between the sequences shown. The italicized bases in *A* for pPTHi8 represent the entire inverted repeat of the 5′ terminus. In *B*, for pPTHm1 the italicized bases represent the inverted repeat generated during reverse transcription and preserved after nuclease S1 digestion. The resulting eight-nucleotide inverted repeats at the ends of the sequence that is inverted in pPTHm1 are underlined by the half arrows. To generate the inversions in pPTHm1, a second step involving an inversion of this sequence must occur, possibly during the transformation of the bacteria. In model *C*, single-stranded cDNA is synthesized by reverse transcription and the mRNA is removed. During synthesis of the second strand by DNA polymerase, a nick by a contaminating nuclease may be introduced in the first strand. The unfolding of the loop after nicking, followed by extension of the first strand by DNA polymerase by using the now inverted first-strand sequence as a template, generates the observed inverted sequence in pPTHm1.

peats similar to those of the interferon cDNA inserts, with the repeats at either end of the insert separated by a central portion of a different size. The last two examples of inverted repeats are similar to that of pPTHi8, except that the inverted repeat of pPTHi8 contains fewer nucleotides.

The mechanism by which these inversions occur cannot be determined conclusively from the data available. Proposals to explain these artifacts have included double-stranded DNA formation during the reverse transcription step or breathing and slippage of the DNA during the DNA polymerase reaction to synthesize the second strand (3, 5, 7). An important similarity in the two types of inversions described in this report is that the beginning or 3′ terminus of the inverted sequence can be generated either by correctly copying the mRNA or by the inversion process. In other words, the beginning of the inversion corresponds to a sequence toward the 5′ terminus of the mRNA that is complementary to an internal sequence. The same phenomenon occurs in the inversions in the ovalbumin and insulin cDNAs, although, in the latter, only six of the first seven nucleotides are complementary to an internal sequence. Such inversions could be generated during the synthesis of cDNA by reverse transcriptase if the self-complementary region of the cDNA permitted the DNA to fold back and hybridize to the internal sequence, and this was followed by continued DNA synthesis. Models incorporating this idea to explain the inverted sequences in pPTHi8 and pPTHm1 are shown in Fig. 4. Generation of the cDNA insert of pPTHi8 (Fig. 4*A*) would involve the synthesis of the sequence -A-G-C-T-G-T- near the 5′ terminus of the mRNA, which can loop back and hybridize to a complementary internal sequence. For the DNA molecule to loop back on itself and hybridize, the RNA template would have to be removed. This potentially could be mediated by the RNase H activity associated with reverse transcriptase, as has been postulated for similar loop-back transcription during reverse transcription of avian sarcoma virus RNA (15). Continued reverse transcription then produces an inverted copy at the 5′ terminus if, after denaturation of the DNA, a smaller loop forms at the 5′ terminus and serves as primer for the extension of the second strand by DNA polymerase. One possible small loop of 30 bases is shown in the figure. The incubation of the double-stranded cDNA with nuclease S1 then results in removal of the single-stranded loop and some nibbling of the double-stranded region to produce the final sequence of the pPTHi8 insert. The initial step in the formation of the inversion in the cDNA insert of pPTHm1 may be similar to that described above, except that the complementary internal sequence is much nearer the 5′ end, as shown in Fig. 4*B*. In this case, after synthesis of the sequence -A-G-C-T-G-, looping back, hybridization to an internal sequence, and continued synthesis of DNA can occur. After denaturation, a shorter loop, possibly as shown in the figure, can be formed. After the DNA polymerase reaction and

cleavage of the hairpin loop, eight nucleotides at the 5′ terminus are an inverted copy of an internal sequence. Generation of the inversion seen in the cDNA insert of pPTHm1, requires a second step, an inversion of the region of the DNA containing the inverted repeats at its ends, presumably during the transformation procedure in the bacteria. Inverted repeats are characteristic of the ends of DNA sequences that undergo rearrangement, such as transposons. A serious drawback to this hypothesis is the fact that such inversions would be expected to occur at a very low rate in recA$^-$ strains, such as the *E. coli* X1776 used to obtain pPTHm1.

An alternative hypothesis proposed to explain the inversion in an insulin cDNA insert similar to the inversion in pPTHm1 is shown in Fig. 4*C* (S. J. Chan, personal communication). In this mechanism, a DNA copy of the mRNA is synthesized by reverse transcription. After removal of the mRNA, the DNA folds back on itself to form a hairpin loop. During the DNA polymerase reaction to form the second strand, contaminating DNase, present in many commercial DNA polymerase preparations, may nick the DNA on the first strand side of the hairpin loop. If the loop then unfolds, DNA polymerase could then "repair" the first strand by using the now inverted first-strand sequence as a template, yielding the inversion observed in pPTHm1. This model has the advantage of not requiring an inversion to occur in the bacteria. On the other hand, inverted repeats as occurred in pPTHi8 are not easily explained by this mechanism.

Models based on "slippage" of DNA during the polymerase reaction (3, 5, 7) have also been proposed to explain the artifactual inversions. This explanation seems less likely because it would require slippage of 40–50 base-paired nucleotides at the relative low temperature of the DNA polymerase reaction. If the reverse transcriptase model is correct, artifactual inversions might be expected to occur more frequently for mRNAs that have considerable secondary structure because this would increase the potential for looping back. These inversions may be reduced under conditions in which the looping back and continued reverse transcription is inhibited, as in the presence of actinomycin D or where RNase H activity is inhibited by addition of fluoride. If nicking by DNase is required, then such artifacts should be reduced if DNase-free DNA polymerase is used. It is clear that these types of artifacts are reasonably frequent and that the sequence of a mRNA, particularly at the 5′ terminus, cannot be unambiguously determined on the basis of a single cloned cDNA sequence.

We thank Dr. Jeffrey Gardner for invaluable assistance with the DNA sequence determinations, Drs. Ed Morgan and David Shapiro for suggestions on cloning procedures, and Drs. Shu J. Chan and Donald F. Steiner for helpful discussions about possible models for the generation of the inversions. We thank Calihan and Co. Meats, Peoria, Il, for supplying parathyroid glands. This research was supported by Grant AM 18866 from the National Institutes of Health.

1. Efstratiadis, A., Kafatos, F. C. & Mantiatis, T. (1977) *Cell* **10**, 571–586.
2. Derynck, R., Content, J., DeClercq, E., Volckaert, G., Taverneir, J., Devos, R. & Fiers, W. (1980) *Nature (London)* **285**, 542–549.
3. Adams, J. M., Gough, N. M., Webb, E. A., Tyler, B. M., Jackson, J. & Cory, S. (1980) *Biochemistry* **19**, 2711–2719.
4. Richards, R. I., Shine, J., Ullrich, A., Wells, J. R. E. & Goodman, H. M. (1979) *Nucleic Acids Res.* **7**, 1137–1146.
5. Fagan, J. B., Pastan, I. & deCrombrugghe, B. (1980) *Nucleic Acids Res.* **8**, 3055–3064.
6. Chan, S. J., Noyes, B. E., Agarwal, K. L. & Steiner, D. F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5036–5040.
7. O'Hare, K., Breathnach, R., Benoist, C. & Chambon, P. (1979) *Nucleic Acids Res.* **7**, 321–334.
8. Kronenberg, H. M., McDevitt, B. F., Majzoub, J. A., Nathans, J., Sharp, P. A., Potts, J. T., Jr. & Rich, A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4981–4985.
9. Stolarsky, L. & Kemper, B. (1978) *J. Biol. Chem.* **253**, 7194–7207.
10. Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) *J. Biol. Chem.* **253**, 2483–2495.
11. Gordon, D. F. & Kemper, B. (1980) *Nucleic Acids Res.* **8**, 5669–5683.
12. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3727–3731.
13. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
14. Kozak, M. (1981) *Cell* **22**, 7–8.
15. Swanstrom, R., Varmus, H. E. & Bishop, J. M. (1981) *J. Biol. Chem.* **256**, 1115–1120.