Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 4363–4367, July 1984
Biochemistry

Molecular cloning of human terminal deoxynucleotidyltransferase

(terminal transferase/expression plasmids/DNA sequence determination/amino acid sequence determination/hybrid selection)

R. C. PETERSON, L. C. CHEUNG, R. J. MATTALIANO, L. M. S. CHANG, AND F. J. BOLLUM*

Department of Biochemistry, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814

Communicated by Bruce M. Alberts, April 9, 1984

ABSTRACT **A cDNA of the human terminal deoxynucleotidyltransferase (TdT; "terminal transferase," EC 2.7.7.31) was isolated from a human lymphoblastoid cell cDNA library in λgt11 by using immunological procedures. Four inserts containing 723 to 939 base pairs were reeloned in pBR322 for hybridization and preliminary sequence studies. mRNA selected by hybridization to recombinant DNA was translated to a 58-kDa peptide that specifically immunoprecipitated with rabbit antibodies to calf terminal transferase and mouse monoclonal antibody to human terminal transferase. Blot hybridization of total poly(A)$^+$ RNA from KM3 (TdT$^+$) cells with nick-translated pBR322 recombinant DNA detected a message of about 2000 nucleotides, sufficient to code for the 580 amino acids in the protein. mRNA from terminal transferase$^-$ cells gave no signal in hybrid selection or RNA blot hybridization. The complete sequence of the 939-base-pair insert sequence was obtained from deletions cloned in pUC8. The DNA sequence contains an open reading frame coding for 238 amino acids, about 40% of the protein. Three peptides isolated by HPLC from tryptic digests of succinylated 58-kDa calf thymus terminal transferase were sequenced, providing 20, 18, and 22 residues of peptide sequence. A search of the translated sequence of the 939-base-pair insert shows three regions beginning after arginine that have greater than 90% homology with the sequence determined from the calf thymus terminal transferase peptides. These results provide unambiguous evidence that the human terminal transferase sequence has been cloned.**

Terminal deoxynucleotidyltransferase (TdT; "terminal transferase," EC 2.7.7.31) is a 58-kDa protein expressed at about 100,000 molecules per cell in cortical thymocytes and primitive bone marrow lymphocytes (1). Expansion of the cell population expressing terminal transferase occurs in certain acute leukemias (2), making detection of this population an important adjunct in diagnosis. The enzyme activity of the protein is of major importance in genetic engineering (3). The normal biological function of the enzyme remains unknown and its presence in pre-T and pre-B cells remains an object of speculation. Persistence of the enzyme activity with high degree of sequence conservation in species from birds to mammals (4) may imply a role of considerable importance in DNA metabolism during lymphocyte development.

The development of procedures for cloning genes in expression plasmids provides a means for detecting DNA inserts through expression of protein sequence and measurement using immunological reagents. By using these procedures even minor mRNAs can be found, assuming patience and a truly specific polyclonal or monoclonal antibody is available. We have taken this approach to search for the terminal transferase gene, using λgt11 (5) as the expression vector and rabbit antibody to calf terminal transferase (6) with ^{125}I-labeled staphylococcal protein A (^{125}I-protein A) as the detection system. A cDNA library constructed from total poly(A)$^+$ RNA from human lymphoblastoid cells (KM3) contained several λgt11 clones that expressed terminal transferase sequence. Hybrid selection and comparison of translated DNA sequences with amino acid sequences of calf terminal transferase proved that we have isolated DNA sequence present in the human terminal transferase gene. The DNA sequences of terminal transferase now available should permit isolation of the cDNA clone of the complete message and analysis of gene structure, regulation, and evolution of this unusual protein.

MATERIALS AND METHODS

Materials. Vanadyl ribonucleoside complex, the large fragment of *Escherichia coli* DNA polymerase I, restriction endonuclease *Eco*RI and *Hin*dIII, λ phage packaging system, BAL-31 nuclease, exonuclease III, the 15-base M13 phage sequencing primer, rabbit reticulocyte lysate, and ^{125}I-labeled bovine serum albumin and ovalbumin were purchased from Bethesda Research Laboratories. Oligo(dT)-cellulose, $(dT)_{12-18}$, $(dG)_{12-18}$, and *Eco*RI and *Hin*dIII sequencing primers were from P-L Biochemicals. Avian myeloblastosis virus reverse transcriptase was from Life Sciences. Synthetic *Eco*RI and *Hin*dIII linkers were from Collaborative Research (Waltham, MA). ^{125}I-protein A, [α-^{32}P]dNTPs, and [^{35}S]methionine were from New England Nuclear. L-1-Tosylamido-2-phenylethyl chloromethyl ketone trypsin was from Worthington and Polybrene was from Aldrich.

Isolation of Poly(A)$^+$ mRNA. Total RNA was isolated from KM3 cells (TdT$^+$), a pre-B human lymphoblastic leukemia cell line, and RAMOS-1 cells (TdT$^-$), a human Burkitt lymphoma cell line, by using a precipitation method with 3 M LiCl, 0.1% $NaDodSO_4$, 6 M urea, 2 mg of heparin per ml, and 6 mM vanadyl ribonucleoside complexes as described by Adrian and Hutton (7).

Poly(A)$^+$ mRNA was isolated from total RNA by using two passes on an oligo(dT)-cellulose column essentially as described by Padgett *et al.* (8). From 3×10^9 cells, 200–350 μg of poly(A)$^+$ mRNA was recovered by this procedure.

Construction of the λgt11 Library and Screening for Terminal Transferase$^+$ Recombinant Clones. cDNA was synthesized from KM3 poly(A)$^+$ RNA with reverse transcriptase (9). A poly(dC) tail was added to the cDNA by using calf thymus terminal transferase (10) and synthesis of the second strand with reverse transcriptase used $(dG)_{12-18}$ as a primer (11). This DNA was treated with the large fragment of *E. coli* DNA polymerase I in the presence of all four dNTPs before ligation with synthetic *Eco*RI linkers and digestion with *Eco*RI. The double-stranded cDNA fragments were subjected to electrophoresis on a 1% agarose gel, and fragments greater than 600 base pairs (bp) were isolated by electroelution into hydroxyapatite (12). The resulting double-stranded

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: bp, base pair(s); TdT, terminal deoxynucleotidyltransferase.

*To whom reprint requests should be addressed.

cDNA fragments were ligated to *Eco*RI-cleaved, alkaline phosphatase-treated λgt11 DNA (5) and packaged into phage particles.

Recombinant phages were plated on *E. coli* Y1090 and screened for antigen-producing clones as described by Young and Davis (13) by using affinity-purified rabbit antibody to calf terminal transferase at 2.5 μg/ml followed by ^{125}I-protein A. Plaque purification was carried out on each positive recombinant phage obtained. The λgt11 library, plated on *E. coli* LE392, was also screened with nick-translated pT17 DNA (14) after transfer to nitrocellulose filters (15).

Recloning of Terminal Transferase cDNA Fragments in pBR322 and pUC8 and DNA Sequencing. Inserts in λgt11 phage were removed by cleavage with *Eco*RI and recloned in *Eco*RI-cleaved, phosphatase-treated pBR322 to produce derivatives called pT16, pT17, pT18, and pT19. Overlapping deletions for DNA sequencing were made by BAL-31 endonuclease digestion of pT17 DNA (16) after cleavage with either *Hin*dIII or *Pst* I endonuclease. After addition of synthetic *Hin*dIII linkers, the DNA was digested with *Hin*dIII and *Eco*RI and subjected to electrophoresis in a 2% agarose gel. DNA fragments were eluted from the gel, ligated to *Hin*dIII- and *Eco*RI-cleaved pUC8 plasmid DNA (17), and used to transform *E. coli* DH-1 cells (18).

Single-stranded templates for sequencing by the dideoxy chain termination method (19) were made from plasmid DNA by cleavage with *Eco*RI, *Hin*dIII, or *Pst* I restriction endonuclease followed by digestion with exonuclease III (20, 21). *Pst* I-cleaved DNA required a brief treatment with BAL-31 nuclease to produce exonuclease III-susceptible ends. Synthetic *Eco*RI and *Hin*dIII primers (22) were used for sequencing DNA in pBR322 and a 15-base primer for the *lac* operon was used for pUC8.

Blot Hybridization Analyses. Poly(A)$^+$ RNA was denatured in 50% (vol/vol) formamide, 2 M formaldehyde at 65°C for 15 min, subjected to electrophresis on a 1.5% agarose gel containing 2.2 M formaldehyde in 0.2 M morpholinopropanesulfonate, pH 7.0/0.05 M sodium acetate/5 mM EDTA (23, 24), and transferred to a nitrocellulose sheet (25). RNA species hybridizing to ^{32}P-labeled recombinant plasmid DNA were detected by autoradiography. Poly(A)$^-$ KM3 RNA, primarily 18S and 28S rRNA, and pBR322 DNA cleaved with restriction endonucleases (*Pst* I, *Hin*cII, *Bgl*, or *Sau*3A) were treated in the same manner and used as size markers.

Unsized double-stranded cDNA samples were electrophoresed on 1.2% agarose in 0.04 M Tris/acetate buffer at pH 7.8 in 2 mM EDTA, transferred to a nitrocellulose filter (26), hybridized to ^{32}P-labeled recombinant plasmid DNA, and detected by autoradiography. pBR322 DNAs cleaved with restriction endonucleases were used as size markers.

Hybrid Selection, *in Vitro* Translation, and Immunoprecipitation. pBR322 DNA and pT18 DNA (100 μg each) were bound separately to 1-cm^2 nitrocellulose filters as described by Kafatos *et al.* (27). One-sixth of the nitrocellulose paper was used for the individual hybrid selection reactions (9). In each experiment four hybrid selection reactions were carried out with 50 μg of poly(A)$^+$ mRNA: pBR322 DNA with KM3 mRNA, pT18 DNA with KM3 mRNA, pBR322 DNA with RAMOS-1 mRNA, and pT18 DNA with RAMOS-1 mRNA. The mRNA selected in each hybridization was divided into two aliquots and translated in the rabbit reticulocyte lysate system. Immunoprecipitation of the protein products was carried out on one aliquot with rabbit antibodies to calf terminal transferase and on the other aliquot with mouse monoclonal antibody to human terminal transferase.

Translation of mRNA with rabbit reticulocyte lysate was as described by Pelham and Jackson (28), using 0.9 mM [^{35}S]methionine (1030 Ci/mmol; 1 Ci = 37 GBq) and 0.25 mM phenylmethylsulfonyl fluoride. Incubation was at 30°C and the reactions were monitored by spotting aliquots on Whatman GF/C paper. The filters were washed with trichloroacetic acid and radioactivity was measured in a liquid scintillation counter.

When translation products were to be immunoprecipitated, a reaction volume of 50 μl was used. To terminate translation, 25 μl of puromycin at 0.6 mg/ml, 3 mM phenylmethylsulfonyl fluoride, and 30 mM methionine was added and incubation was continued for 5 min. The reaction mixture was clarified in a Beckman Airfuge for 30 min at 28 pounds/inch^2 (190 kPa) before immunoprecipitation. Each reaction mixture was diluted with 100 μl of 0.2 M Tris·HCl/0.1 M NaCl containing 20 mM methionine and 0.5% Nonidet P-40, and 2 μg of affinity-purified rabbit antibodies to calf terminal transferase [or 2 μg of mouse monoclonal antibody to human terminal transferase (29)] was added. The mixture was incubated at room temperature for 2 hr. The immune complexes formed were adsorbed onto formaldehyde-treated *Staphylococcus aureus* cells (30). Proteins in the immune complexes on *S. aureus* cells were dissociated in $NaDodSO_4$ buffer, analyzed by electrophoresis on a 12.5% polyacrylamide gel (31) in the presence of $NaDodSO_4$, and detected on the gel by fluorography.

Sequencing of Tryptic Peptides of Succinylated Terminal Transferase. The high molecular weight form of calf thymus terminal transferase (32) was modified with succinic anhydride as described by Glazer *et al.* (33). After exhaustive dialysis against 0.25 M NH_4HCO_3, pH 8.4, the modified protein (2 mg/ml) was treated with trypsin in the presence of 0.1 mM $CaCl_2$ at 37°C, using trypsin at 0.04 mg/ml for 7 hr followed by addition of trypsin at 0.04 mg/ml and incubation for 9 more hr. The tryptic peptides precipitated by titration to pH 1.0 with trifluoroacetic acid were collected by centrifugation and redissolved in 0.15 M *N,N*-bis(2-hydroxyethyl)glycine (Bicine) buffer at pH 8.1 containing 6 M guanidine·HCl and 1 mM EDTA. After reacidification to pH 1.0, tryptic peptides were separated on a 0.39 × 30 cm μBondapak C-18 column with a linear gradient of 0–80% acetonitrile in 0.1% trifluoroacetic acid (34) in a Waters HPLC system.

Automated Edman degradation was performed on 5–15 nmol of each isolated peptide, using a Beckman 890C sequenator equipped with a cold trap accessory and Sequemat P-6 autoconverter. The 0.1 M Quadrol program with S_1 and S_2 wash (Beckman) was used in all instances, except that the concentration of phenylisothiocyanate was 2.5% (35). Polybrene (6 mg) and 100 nmol of glycylglycine were subjected to six cycles of Edman degradation (36) prior to addition of sample. Phenylthiohydantoin derivatives of amino acids were identified by separation on a 0.46 × 25 cm Zorbax-octadecylsilica column at 58°C, using a linear gradient of 20–42% acetonitrile (37) in 0.02 M sodium acetate buffer at pH 5.5 in a DuPont HPLC system.

RESULTS

Analysis of Recombinants Containing Terminal Transferase DNA Sequences. DNA complementary to poly(A)$^+$ RNA from KM3 cells (TdT$^+$) was used to construct the λgt11 library. *Eco*RI sites were added to the ends of the cDNA by ligation to synthetic *Eco*RI linkers and the resulting DNA was inserted into the *Eco*RI site of λgt11. Insertion into the *Eco*RI site results in production of an inactive β-galactosidase. We obtained greater than 1.2 × 10^6 phages per μg of cDNA used in this construction, and 25% of these were recombinants (β-galactosidase-negative). If the terminal transferase sequence is inserted in phase with the β-galactosidase translational reading frame the fused protein should contain immunological determinants from the inserted sequence. We screened about 200,000 plaques from our λgt11 library and obtained 5 recombinants producing hybrid protein reactive

to rabbit anti-calf terminal transferase. Four of these inserts, recloned in pBR322, are called pT16 (768-bp insert), pT17 (939-bp insert), pT18 (789-bp insert), and pT19 (723-bp insert).

When pT17 was used to probe poly(A)$^+$ RNA from KM3 cells after gel electrophoresis and transfer to a nitrocellulose filter, an RNA slightly larger than 18S rRNA was detected (Fig. 1 *Left*, lanes C and D). This RNA is estimated to contain 2000 nucleotides, sufficient to code for the intact human terminal transferase peptide (58 kDa). Poly(A)$^+$ RNA from RAMOS-1 cells (TdT$^-$) does not contain any material hybridizable to pT17 (Fig. 1 *Left*, lanes A and B).

The DNA sequences at the ends of pT16, pT17, pT18, and pT19 were determined to establish sequence relationships. The procedure used to make the double-stranded cDNA should produce cloned fragments with a poly(dG) sequence at one end, a poly(dA) sequence at the other end, and part of an *Eco*RI linker sequence at both ends. An *Eco*RI linker sequence and a poly(dG) stretch were found at one end of each cloned cDNA fragment, corresponding to the 5′ end of the mRNA. The nucleotide sequence at the opposite ends was identical for all four fragments and lacked both a poly(dA) stretch and the sequence present in the *Eco*RI linker; thus the *Eco*RI site found at the 3′ end of the cloned fragments was probably present in the cDNA sequence. To test this hypothesis, total double-stranded cDNA to poly(A)$^+$ RNA from KM3 cells was subjected to gel electrophoresis before and after digestion with *Eco*RI endonuclease. DNA blot hybridization of total double-stranded KM3 cDNA with ^{32}P-labeled pT17 DNA detected a band at about 2000 bp, demonstrating that our original preparation of KM3 cDNA contained full-length cDNA for terminal transferase mRNA (Fig. 1 *Right*, lane A). The *Eco*RI-treated sample produced a band at about 1200 bp (Fig. 1 *Right*, lane B), indicating presence of *Eco*RI sites within the cDNA sequence.

The 1200-bp fragment detected in the DNA blot of *Eco*RI-treated KM3 cDNA should be present in the λgt11 cDNA library as constructed. Since we did not detect any inserts of this length in clones selected by the immunological screening procedure, we assume that all of the larger sequences must be in the wrong reading frame. Labeled pT17 DNA was used to estimate the total yield of recombinants containing the terminal transferase sequence. In this way we found more than 1% recombinants. We analyzed 16 of these recombinants, and all have inserts of 1200 bp, presumably as generated by *Eco*RI cleavage at these sites in the cDNA. The total frequency of TdT$^+$ clones is greater than expected, since the transferase represents less than 0.1% of total protein in KM3 cells and less than 0.1% of immunoprecipitable protein produced by *in vitro* translation of KM3 poly(A)$^+$ RNA. This indicates that the cDNA fragments that have *Eco*RI sites, notably terminal transferase sequences, are over-represented in this library. We chose to sequence these inserts that gave expression and have not examined the additional sequence present in the 1200-bp inserts.

Translation of Hybrid-Selected mRNA. Results obtained from translation of mRNA selected by hybridization to pT18 DNA immobilized on nitrocellulose are shown in Fig. 2. When KM3 (TdT$^+$ cells) mRNA was used for hybridization with pBR322 DNA, no translation product was detected in the immunoprecipitate (Fig. 2, lane A). NaDodSO$_4$/polyacrylamide gel analysis of the translated and immunoprecipitated product of KM3 mRNA selected with pT18 DNA produced a two-peptide pattern (Fig. 2, lane B). The largest product migrates as the 58-kDa terminal transferase peptide and the second band, at 56 kDa, is thought to be due to proteolytic cleavage during translation, as often occurs during purification of the high molecular weight terminal transferase species (4, 31). When the hybrid selection reactions were carried out with RAMOS-1 (TdT$^-$ cells) mRNA, no immunoreactive material was produced with either pBR322 DNA (Fig. 2, lane C) or pT18 DNA (Fig. 2, lane D), indicating the absence of terminal transferase mRNA from RAMOS-1 cells. Identical results were obtained by precipitation with mouse monoclonal antibody to human terminal transferase.

DNA Sequence of the Cloned Fragment of Human Terminal Transferase Gene. Preliminary dideoxy sequencing was performed on four inserts in pBR322—pT16, pT17, pT18, and pT19—as described above. The orientation of these four DNA fragments, an abbreviated restriction map (experimentally determined, confirmed by sequencing), and homologies between experimentally determined calf terminal transferase amino acid sequences and predicted human terminal transferase sequences (discussed later) are shown in Fig. 3.

The complete sequence of 939 nucleotides in pT17 was determined by making two sets of deletion fragments from pT17. Deletion at one end of the insert was accomplished by digestion of pT17 with *Hin*dIII followed by BAL-31 and at the other end by *Pst* I followed by BAL-31. Both sets of fragments were recloned in pUC8 and sequenced by the di-

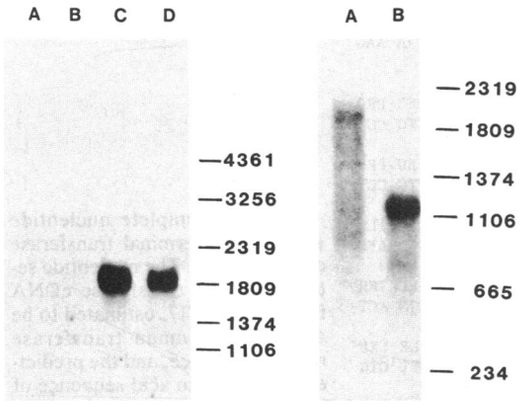

FIG. 1. Blot hybridization of RNA (*Left*) and cDNA (*Right*). For RNA blots, RAMOS-1 poly(A)$^+$ RNA (lanes A and B) and KM3 poly(A)$^+$ RNA (lanes C and D) were analyzed for terminal transferase mRNA by electrophoretic separation on an agarose gel, transfer to a nitrocellulose filter, and hybridization of nick-translated pT17 [^{32}P]DNA. Each lane contained 10 μg of poly(A)$^+$ RNA. For the DNA blots, 0.6 μg of unfractionated double-stranded KM3 cDNA was used with (lane B) and without (lane A) *Eco*RI nuclease digestion. Fragments derived from *Pst* I, *Hin*cII, *Bgl* I, or *Sau*3A digestion of pBR322 DNA were used as size markers; sizes are number of nucleotides.

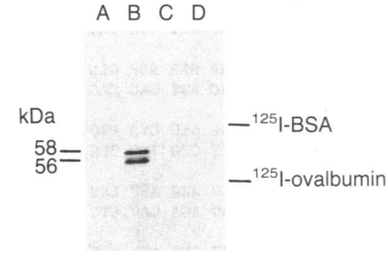

FIG. 2. Translation–immunoprecipitation products of human lymphoblastoid cell RNAs selected by hybridization to pT18 DNA. The molecular weight markers used on the NaDodSO$_4$/polyacrylamide gel are indicated on the right (BSA, bovine serum albumin). Human KM3 terminal transferase, containing 90% 58-kDa species and 10% 56-kDa species, was also separated on the same polyacrylamide gel and detected by Coomassie blue staining to provide molecular weight comparisons. Lane A, products of pBR322 DNA-selected KM3 RNA; lane B, products of pT18 DNA-selected KM3 RNA; lane C, products of pBR322 DNA-selected RAMOS-1 RNA; and lane D, products of pT18 DNA-selected RAMOS-1 RNA.

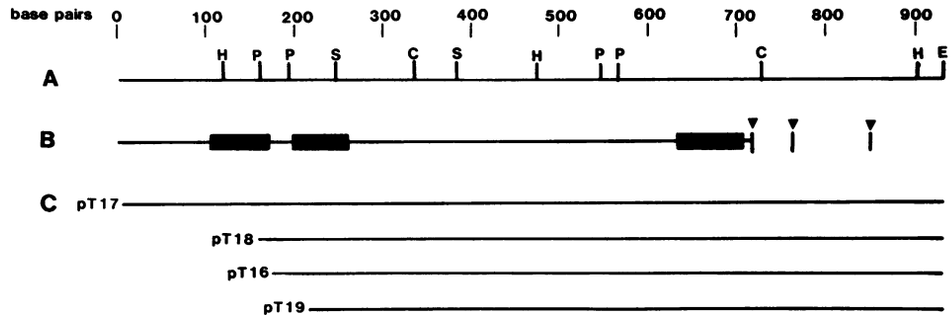

FIG. 3. Summary map of cloned terminal transferase cDNA fragments. Line A shows the positions of restriction endonuclease sites determined from the complete nucleotide sequence of the terminal transferase cDNA fragment of pT17. Line A shows cleavage sites: H, *Hae* III; C, *Hin*cII; P, *Hpa* II; S, *Sau*3A; and E, *Eco*RI. Line B shows translation of the nucleotide sequence of pT17 cDNA in one of three reading frames, providing an open reading frame in the first 238 codons. Termination codons (▼) are found beyond this region and throughout the sequence in the other two reading frames. The nucleotide sequence gives an amino acid sequence in three regions that is homologous to three tryptic peptides of calf terminal transferase (bars). Lines C show the relationship of cDNA fragments from different clones. The poly(dG) tail is found at the left end of each fragment.

deoxy method. The nucleotide sequence of both strands of the insert in pT17 was reconstructed from overlapping fragments. The nucleotide sequence of terminal transferase mRNA in pT17 and the predicted sequence for this region of human protein is shown in Fig. 4. The cDNA portion of this fragment is 928 bp and contains an open reading frame starting at position 2 (Fig. 4) and extending 238 codons. This fragment probably represents the carboxyl-terminal portion of the protein since multiple stop codons are found beyond this region. The sequence ends in an *Eco*RI site and no *Eco*RI linker sequence is found at the 3′ end. The 5′ end contains the *Eco*RI linker sequence and a homopolymer dG run (sequences not included in Fig. 4). The length of the poly(dG) tails at the 5′ end of each of the four recombinants is such that all are in phase with the β-galactosidase gene in λgt11.

Amino Acid Sequences of Calf Thymus Terminal Transferase Peptides. Three peptides that were isolated by HPLC from tryptic digests of succinylated calf terminal transferase were found to contain an amino acid sequence with close homology to the predicted human terminal transferase sequence. This comparison is also shown in Fig. 4 (lowercase sequences). Identity in 19 out of 20, 16 out of 18, and 21 out

```
    PHE ARG THR LEU SER LYS VAL ARG SER ASP LYS SER LEU LYS PHE THR ARG MET GLN LYS
1 T TTC AGA ACT CTG AGT AAA GTA AGG TCG GAC AAA AGC CTG AAA TTT ACA CGA ATG CAG AAA

    ALA GLY PHE LEU TYR TYR GLU ASP LEU VAL SER CYS VAL THR ARG ALA GLU ALA GLU ALA
62  GCA GGA TTT CTG TAT TAT GAA GAC CTT GTC AGC TGT GTG ACC AGG GCA GAA GCA GAG GCC
                                                    P-1     ala glu ala glu ala

    VAL [SER] VAL LEU VAL LYS GLU ALA VAL TRP ALA PHE LEU PRO ASP ALA PHE VAL THR MET
122 GTC [AGT] GTG CTG GTT AAA GAG GCT GTC TGG GCA TTT CTT CCG GAT GCT TTC GTC ACC ATG
    val [gly] val leu val lys glu ala val --- ala phe leu pro asp ala

    THR GLY GLY PHE ARG ARG GLY LYS LYS [MET] GLY HIS ASP VAL ASP PHE LEU ILE THR SER
182 ACA GGA GGG TTC CGG AGG GGT AAG AAG [ATG] GGG CAT GAT GTA GAT TTT TTA ATT ACC AGC
                P-2     gly lys lys [ile] gly his asp val asp phe leu ile --- ---

    PRO GLY SER [THR] GLU ASP GLU GLU GLN LEU LEU GLN LYS VAL MET ASN LEU TRP GLU LYS
242 CCA GGA TCA [ACA] GAG GAT GAA GAG CAA CTT TTA CAG AAA GTG ATG AAC TTA TGG GAA AAG
    pro gly --- [ala] glu asp glu

    LYS GLY LEU LEU LEU TYR TYR ASP LEU VAL GLU SER THR PHE GLU LYS LEU ARG LEU PRO
302 AAG GGA TTA CTT TTA TAT TAT GAC CTT GTG GAG TCA ACA TTT GAA AAG CTC AGG TTG CCT

    SER ARG LYS VAL ASP ALA LEU ASP HIS PHE GLN LYS CYS PHE LEU ILE PHE LYS LEU PRO
362 AGC AGG AAG GTT GAT GCT TTG GAT CAT TTT CAA AAG TGC TTT CTG ATT TTC AAA TTG CCT

    ARG GLN ARG VAL ASP SER ASP GLN SER SER TRP GLN GLU GLY LYS THR TRP LYS ALA ILE
422 CGT CAA AGA GTG GAC AGT GAC CAG TCC AGC TGG CAG GAA GGA AAG ACC TGG AAG GCC ATC

    ARG VAL ASP LEU VAL LEU CYS PRO TYR GLU ARG ARG ALA PHE ALA LEU LEU GLY TRP THR
482 CGT GTG GAT TTA GTT CTG TGC CCC TAC GAG CGT CGT GCC TTT GCC CTG TTG GGA TGG ACT

    GLY SER ARG PHE GLU ARG ASP LEU ARG ARG TYR ALA THR HIS GLU ARG LYS MET ILE LEU
542 GGC TCC CGG TTT GAG AGA GAC CTC CGG CGC TAT GCC ACA CAT GAG CGG AAG ATG ATT CTG

    ASP ASN HIS ALA LEU TYR ASP LYS THR LYS ARG [ILE] PHE LEU LYS ALA GLU SER GLU GLU
602 GAT AAC CAT GCT TTA TAT GAC AAG ACC AAG AGG [ATA] TTC CTC AAA GCA GAA AGT GAA GAA
                              P-3       [val] phe leu lys ala glu --- glu glu

    GLU ILE PHE ALA HIS LEU GLY LEU ASP TYR ILE GLU PRO TRP GLU ARG ASN ALA ***
662 GAA ATT TTT GCG CAT CTG GGA TTG GAT TAT ATT GAA CCG TGG GAA AGA AAT GCC TAG GAA
    glu ile phe ala --- leu gly leu asp tyr ile glu pro trp glu

722 AGT GTT GTC AAC ATT TTT TCC TAT TCT TTT CAA GTT AAA TAA ATT ATG CTT CAT ATT AGT

782 AAA AGA TGC CAT AGG AGA GTT TGG GGT TAT TTA GGT CTT ATT GAA ATG CAG ATT GCT ACT

842 AGA AAT AAA TAA CTT TGG AAA CAT GGG AAG GTG CCA CTG GTA ATG GGT AAG GTT CTA ATA

902 GGC CAT GTT TAT GAC TGT TGC ATA GAA TTC
```

FIG. 4. Complete nucleotide sequence of terminal transferase cDNA in pT17. The nucleotide sequence of the transferase cDNA fragment in pT17, estimated to be 40% of the human transferase mRNA sequence, and the predicted partial amino acid sequence of human transferase are shown in uppercase letters with asterisks at the first stop codon. Lowercase letters are the amino acid sequences determined for bovine peptides P-1, P-2, and P-3. The dashes in the amino acid sequences of bovine peptides indicate cycles of degradation in which no single amino acid could be identified. The boxed amino acids highlight differences in the bovine and human sequences.

of 22 determined residues is seen for P-1, P-2, and P-3, respectively. The total homology is 56 out of 60 residues or 93%.

DISCUSSION

The evidence that the terminal transferase sequence has been cloned begins with detection of the expression of immunological determinants in a fused hybrid protein. The sequence insert from recombinant phages was used to construct pBR322 recombinants that detect a 2000-nucleotide mRNA in TdT^+ cells and select an mRNA that translates to an immunoprecipitable 58-kDa peptide. Mouse monoclonal anti-human terminal transferase was as effective as rabbit polyclonal anti-calf terminal transferase for immunoprecipitation.

The DNA sequence of the 939-bp insert was used to predict the sequence of human terminal transferase. The amino acid sequences of three peptides (60 residues) from calf terminal transferase exhibit greater than 90% homology with the predicted human transferase sequence. All of this evidence is self-consistent and proves our isolation of partial sequence of the human terminal transferase cDNA.

We believe that success in cloning the terminal transferase sequence in the expression vector was due to a specific polyclonal antibody and an element of chance. The chance element relates to the method used to construct the cDNA library. Our original cDNA preparation contained 600 to 3000 nucleotides. The cloning procedure used *Eco*RI inserts and thus the cDNA was cut with *Eco*RI nuclease before joining. As a result, any cDNA containing *Eco*RI sites will be more efficiently inserted into the expression vector. Since we now know that the terminal transferase cDNA contains at least two *Eco*RI sites, this probably resulted in the 1.2% positive clones found compared to less than 0.1% predicted on the basis of message occurrence.

Cloning minor messages in expression vectors allows the use of alternative detection methods, such as the polyclonal antibody system used in this study, with which appropriate signal-to-noise values may be obtained in library screening. But the use of polyclonal antibodies for detection always presents an element of uncertainty due to the possible presence of trace specificities unrelated to the population presenting the major antigenic stimulus. One can break this circle of immunological logic if well-characterized monoclonal antibodies are available. Unfortunately, none of a panel of 12 mouse monoclonal anti-human terminal transferase antibodies (29) could be demonstrated to react with the TdT^+ expression recombinants we isolated in this search. This could be due to the fact that the determinants recognized by our monoclonal antibodies are not expressed in the recombinant phages found or that our detection methods were not sensitive enough. In any event, the specific precipitation of protein translated from hybrid-selected RNA with monoclonal antibody provides good evidence that the clones selected do indeed contain the terminal transferase sequence. That the polyclonal antibody selection process was specific for terminal transferase sequences provides additional evidence for the unique specificity of our rabbit polyclonal antibodies.

The most impressive evidence for our isolation of the terminal transferase sequence comes from the high degree of homology between the human sequence predicted from DNA sequence and the amino acid sequence determined from the bovine enzyme. Bulk purification and sequencing of human terminal transferase is impractical in our situation. Although our earlier studies on terminal transferase molecular weight and immunoreactivity had predicted close homology between species, the confirmation received in this study is most satisfactory. DNA probes can now be used to extend our studies over a wider range of organisms with improved confidence. All of the differences in peptide sequence observed between bovine and human terminal transferase in this study could be results of single base changes in the DNA ($A \rightarrow G$ or $G \rightarrow A$).

This work was supported by Grant GM31393 from National Institute of General Medical Sciences.

1. Bollum, F. J. (1979) *Blood* **54,** 1203–1215.
2. Janossy, G., Bollum, F. J., Bradstock, K. F., McMichael, A., Rapson, N. & Greaves, M. F. (1979) *J. Immunol.* **123,** 1525–1529.
3. Nelson, T. & Brutlag, D. (1979) *Methods Enzymol.* **68,** 41–50.
4. Bollum, F. J. & Chang, L. M. S. (1981) *J. Biol. Chem.* **256,** 8767–8770.
5. Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1194–1198.
6. Bollum, F. J. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 4119–4122.
7. Adrian, G. S. & Hutton, J. J. (1983) *J. Clin. Invest.* **71,** 1649–1660.
8. Padgett, R. A., Wahl, G. M., Coleman, P. F. & Stark, G. R. (1979) *J. Biol. Chem.* **254,** 974–980.
9. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
10. Chang, L. M. S. & Bollum, F. J. (1971) *J. Biol. Chem.* **246,** 909–916.
11. Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schütz, G. (1981) *Nucleic Acids Res.* **9,** 2251–2266.
12. Tabak, H. F. & Flavell, R. A. (1978) *Nucleic Acids Res.* **5,** 2321–2332.
13. Young, R. A. & Davis, R. W. (1983) *Science* **222,** 778–782.
14. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
15. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113,** 237–251.
16. Legerski, R. J., Hodnett, J. L. & Gray, H. B., Jr. (1978) *Nucleic Acids Res.* **5,** 1445–1464.
17. Vieira, J. & Messing, J. (1982) *Gene* **19,** 259–268.
18. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580.
19. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
20. Smith, A. J. H. (1979) *Nucleic Acids Res.* **6,** 831–848.
21. Guo, L. H. & Wu, R. (1982) *Nucleic Acids Res.* **10,** 2065–2084.
22. Wallace, R. B., Johnson, M. J., Suggs, S. V., Miyoshi, K., Bhatt, R. & Itakura, K. (1981) *Gene* **16,** 21–26.
23. Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. (1977) *Biochemistry* **16,** 4743–4751.
24. Goldberg, D. A. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5794–5798.
25. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
26. Southern, E. M (1975) *J. Mol. Biol.* **98,** 503–517.
27. Kafatos, F., Jones, C. W. & Efstratiadis, A. (1979) *Nucleic Acids Res.* **7,** 1541–1552.
28. Pelham, H. R. B. & Jackson, R. J (1976) *Eur. J. Biochem.* **67,** 247–256.
29. Bollum, F. J., Augl, C. & Chang, L. M. S. (1984) *J. Biol. Chem.* **259,** 5848–5850.
30. Kessler, S. (1975) *J. Immunol.* **115,** 1617–1624.
31. Laemmli, U. K. & Favre, M. (1973) *J. Mol. Biol.* **80,** 575–599.
32. Augl, C., Lee, S., Breviario, D., Chang, L. M. S. & Bollum, F. J. (1983) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **42,** 2147 (abstr.).
33. Glazer, A. N., DeLange, R. J. & Sigman, D. S. (1976) in *Laboratory Techniques in Biochemistry and Molecular Biology*, eds. Work, T. S. & Work, E. (Elsevier, New York), pp. 78–79.
34. Mahoney, W. C. & Hermodson, M. A. (1980) *J. Biol. Chem.* **255,** 11199–11203.
35. Brauer, A. W., Margolies, M. N. & Haber, E. (1975) *Biochemistry* **14,** 3029–3035.
36. Hunkapillar, M. W. & Hood, L. E. (1978) *Biochemistry* **17,** 2124–2133.
37. Zimmerman, C. L., Apella, E. & Pisano, J. J. (1977) *Anal. Biochem.* **77,** 569–573.