PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) **International Patent Classification** 6 : **C12N 9/12, 15/54**

A1

(11) **International Publication Number:** **WO 95/14770**

(43) **International Publication Date:** 1 June 1995 (01.06.95)

(21) **International Application Number:** PCT/NZ94/00135

(22) **International Filing Date:** 23 November 1994 (23.11.94)

(30) **Priority Data:**
250288 25 November 1993 (25.11.93) NZ

(71) **Applicant** *(for all designated States except US)*: PACIFIC ENZYMES (1993) LIMITED [NZ/NZ]; UniServices House, 58 Symonds Street, Auckland 1001 (NZ).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: BERGQUIST, Peter, Leonard [NZ/NZ]; 3a Pukerangi Crescent, Ellerslie, Auckland 1005 (NZ). DAY, Darren, John [GB/NZ]; 3 Symonds Street, Auckland 1001 (NZ). GIBBS, Moreland, David [NZ/NZ]; 30 Laurie Avenue, Parnell, Auckland 1001 (NZ). REEVES, Rosalind, Alison [NZ/NZ]; 30 Laurie Avenue, Parnell, Auckland 1001 (NZ). SAUL, David, James [GB/NZ]; 5/4 Cowie Street, Parnell, Auckland 1001 (NZ).

(74) **Agents:** PIPER James, William et al.; James W. Piper & Co., P.O. Box 5298, Wellesley Street, Auckland 1036 (NZ).

(81) **Designated States:** AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) **Title:** IMPROVED POLYMERASE

(57) **Abstract**

The invention comprises a thermophilic enzyme, more specifically a heat-stable DNA polymerase, isolated from the New Zealand *Thermus* species *Thermus filiformis*. The invention also includes recombinant plasmids and transformed host cells capable of producing the enzyme. The enzyme is classified into class EC 2.7.7.7; a DNA nucleotidyltransferase DNA-directed type.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

IMPROVED POLYMERASE

FIELD OF INVENTION

The invention comprises a thermophilic enzyme, more specifically a heat-stable DNA polymerase, isolated from the New Zealand *Thermus* species *Thermus filiformis.* The invention also includes recombinant plasmids and transformed host cells capable of producing the enzyme. The enzyme is classified into class EC 2.7.7.7; a DNA nucleotidyltransferase DNA-directed type.

DEFINITIONS

In referring to a peptide chain as being comprised of a series of amino acids "substantially or effectively" in accordance with a list offering no alternatives within itself , we include within that reference any versions of the peptide chain bearing substitutions made to one or more amino acids by similar amino acids in such a way that the overall structure and the overall function of the protein composed of that peptide chain is substantially the same as - or undetectably different to - that of the unsubstituted version. For example it is generally possible to exchange alanine and valine without greatly changing the properties of the protein, especially if the changed site or sites are at positions not critical to the morphology of the folded protein.

The term "thermostable" as applied to an enzyme means that the enzyme is relatively unaffected by heat. Normally such enzymes are used in aqueous solutions and the upper limits

of the temperature range are determined by the boiling point of water at the relevant environmental pressures. Preferably a thermostable enzyme remains active for a long period at a high temperature and preferably it also has an enhanced K_m at a high temperature.

Other definitions are used in a manner consistent with the art.

BACKGROUND OF THE INVENTION

Heat stable DNA polymerases (EC 2.7.7.7, DNA nucleotidyltransferase DNA-directed) have been isolated from numerous thermophilic organisms (for example: Kaledin *et al.* 1980. *Biokimiya* **44**, 644-651; Kaledin *et al.* 1981. *Biokimiya* **46**, 1247-1254 ; Kaledin *et al.* 1982. *Biokimiya* **47**, 1515-1521; Ruttimann, *et al.* 1985. *Eur. J. Biochem* **149**, 41-46; Neuner *et al.* 1990. *Arch. Microbiol.* **153**, 205-207). For some organisms, the polymerase gene has been cloned and expressed (Lawyer *et al.* 1989. *J. Biol. Chem.* **264**, 6427-6437; Engelke *et al.* 1990. *Anal. Biochem.* **191**, 396-400; Lundberg *et al.* 1991. *Gene* **108**, 1-6; Perler *et al.* 1992. Proc. Natl. Acad. Sci. USA **89**, 5577-5581). (References to cited literature are provided at the end of the description of the preferred embodiments.)

Thermophilic DNA polymerases are increasingly becoming important tools for use in molecular biology and there is growing interest in finding new polymerases which have more suitable properties and activities for use in diagnostic detection, cloning and DNA sequencing. For example, the use of thermostable enzymes in the PCR reaction, which is used to amplify existing nucleic acid sequences by a very large ratio. At present, three polymerases have become available from *Thermus* species. *Taq* polymerase from *T. aquaticus* (Yellowstone Park, USA, Brock *et al.* 1969. J. Bacteriol. **98**, 289-297); *Tth* polymerase from *"T. thermophilus"* (Japan, Oshima and Imahori. 1974. J. Syst. Bacteriol. **24**, 104-112) and *Tfl* polymerase from *"T. flavus"* (Japan, Saiki *et al.* 1972. Agric. Biol. Chem. **34**, 2357-2366). The present patent concerns a fourth DNA polymerase from *T. filiformis,* a species isolated

from Waimangu Hot Springs, New Zealand (Hudson *et al.* 1987. Int. J. Syst. Bacteriol. **37**, 431-436). Neither *"T. thermophilus"* nor *"T. flavus"* are yet accepted as validly named species, whereas *T. filiformis* and *T. aquaticus* are accepted as distinct, valid species. Phylogenetic analysis of the *Thermus* genus (as described later in this document) indicates that the New Zealand *Thermus* isolates (including *T. filiformis*) are evolutionarily well separated from the American and Japanese isolates (Saul *et al.* 1993. Int. J. Syst. Bacteriol. **43**, 754-760).

Currently, *Taq* polymerase is the preferred enzyme for DNA sequencing using automated thermal sequencing machines such as the Applied Biosystems 373A DNA Sequencer. The advantages of using a thermophilic enzyme are that less template is required due to the linear amplification that occurs with thermal cycling and also the elevated temperatures help to melt secondary structures in the DNA that may be problematic with conventional sequencing techniques. *Taq* polymerase is currently the only enzyme used for the polymerase chain reaction (PCR). For performing a PCR on RNA rather than DNA templates (RT/PCR), a polymerase must be capable of first producing a DNA copy of an RNA template before normal PCR can commence. This activity, known as reverse transcription, has been noted to varying degrees with *Tth* and *Taq* DNA polymerases (Jones, *et al.* 1989. *Nucl. Acids Res.* **17**, 8387-8388; Meyers and Gelfand 1991. *Biochem.* **30**, 7661-7666). Reverse transcription by thermophilic DNA polymerases has advantages over mesophilic viral reverse transcriptases such as Moloney murine leukemia virus reverse transcriptase (RT) and avian myeloblastosis virus RT which are commonly used for cDNA synthesis, because the higher reaction temperatures possible with thermophilic polymerases helps to destabilise RNA secondary structures which can pose significant problems for the mesophilic viral RTs.

The wide range of uses for DNA polymerases means that it is advantageous to have available a variety of enzymes. For instance, the reverse transcriptase activity of *Taq* polymerase is considerably lower than that of *Tfil* or *Tth* polymerases and is consequently less suitable for

RT/PCR. In automated DNA sequencing, an enzyme is required which can incorporate fluorescently-labelled dideoxynucleosides efficiently. *Taq* polymerase performs this task best at reduced temperatures and with high concentrations of the base analogs. This is disadvantageous for templates which tend to form secondary structures which require high temperatures for denaturation.

Accordingly there is a desire in the art to produce alternative DNA polymerases to the most frequently used *Taq* polymerase, in order to improve both the cost-effectiveness and the quality of the results obtained. It is well known in the art that use of an improved enzyme capable of operating at a higher temperature improves the specificity and selectivity of the primer-directed extension process within the PCR reaction.

OBJECT

It is an object of this invention to provide a polymerase enzyme (EC 2.7.7.7) and/or cloned DNA and/or vectors and/or transformed host cells capable of producing this enzyme, or at least to provide the public with a useful choice.

STATEMENT OF THE INVENTION

In one aspect the invention provides a purified heat-stable DNA polymerase enzyme (EC 2.7.7.7, a DNA nucleotidyltransferase DNA-directed enzyme), characterised in that it has reverse transcriptase activity in the presence of magnesium ions.

In another aspect the invention comprises a recombinant DNA sequence that encodes DNA polymerase activity of the microorganism *Thermus filiformis.*

In a related aspect, the DNA sequence is that shown in Fig 1.

In a second related aspect the invention comprises a recombinant DNA sequence that encodes amino acid residues 1 to 833 as also shown in Fig 1.

In a further aspect the invention comprises recombinant DNA plasmids that comprise the DNA sequence of the invention inserted into plasmid vectors and which can be used to drive the expression of the thermostable DNA polymerase of *Thermus filiformis* in a host cell transformed with the plasmid.

In a related aspect the invention includes a recombinant plasmid comprising the vector pUC18 carrying the *Thermus filiformis* DNA polymerase gene and designated pNZ2300.

Also in a related aspect the invention includes a recombinant plasmid comprising the vector pT7-7 carrying the *Thermus filiformis* DNA polymerase gene and designated pNZ2303.
In a further related state the invention comprises a host cell transformed with the recombinant DNA of the invention.

As one option, the invention comprises such a host cell, being *Escherischia coli*/pNZ2300, the strain being designated PB5900.

As another option, the invention comprises such a host cell, being *Escherischia coli*/pNZ2303, the strain being designated PB5904.

In a further aspect the invention comprises a DNA polymerase having a molecular mass of about 93.4 KDa.

In a further aspect the invention comprises a DNA polymerase having a pH-dependent activity as shown in Fig 3.

In a yet further aspect the invention comprises a DNA polymerase having a dependency on magnesium ion concentration, as shown in Fig 4.

In yet another aspect the invention comprises a DNA polymerase having Reverse Transcriptase activity in the substantial absence of manganese ions.

In a still further aspect the invention comprises a DNA polymerase having an activity, referenced to *Taq* polymerase, as illustrated in Fig 5.

Preferably the invention provides a thermostable DNA polymerase which has been purified from a recombinant *Escherichia coli* strain containing the gene encoding for this enzyme, which has been isolated from *Thermus filiformis.*

The coding region for the polymerase was isolated by PCR amplification from *T. filiformis* genomic DNA. The PCR product was cloned in the vector pUC18 and active thermophilic polymerase was expressed as a fusion protein with the amino-terminus of ß-galactosidase. Higher expression was achieved by cloning the gene into the expression vector pT7-7 where the polymerase was expressed as a fusion with the T7 gene 10 protein. Analysis of the sequence of the polymerase showed that it was clearly distinguishable from similar enzymes from *Thermus aquaticus*, *"Thermus flavus"* and *"Thermus thermophilus"*.

The DNA polymerase has been purified to homogeneity by a simple three-step procedure involving heat precipitation of *E. coli* proteins followed by Q-sepharose and heparin sepharose chromatography. The purified protein has a relative molecular mass of 94 kDa as determined by SDS polyacrylamide gel electrophoresis (PAGE) using 12% and 7.5% polyacrylamide gels, whereas the predicted molecular mass, from the polypeptide sequence of Figure 1, is 93.2 kDa. The purified enzyme exhibits no endonuclease, exonuclease or ribonuclease activity,

shows maximal activity in low salt buffers at alkaline pHs and has an absolute requirement for Mg^{2+} ions (Km 1.5 mM). The DNA polymerase shows reverse transcriptase activity on primed RNA templates. The polymerase is highly processive and has an extension rate of approximately 1000 bases/min.

DRAWINGS

These and other aspects of this invention, which should be considered in all its novel aspects, will become apparent from the following description, which is given by way of example only, with reference to the accompanying drawings, in which:

FIG. 1 (3 sheets, fig 1-1, fig 1-2 and fig 1-3) shows the DNA sequence of the polymerase gene of *Thermus filiformis* and the derived peptide sequence for *Tfil* polymerase. The one-letter abbreviations for the amino acids are shown here for convenience.

F =	Phenylalanine	T =	Threonine	D =	Aspartic acid
L =	Leucine	A =	Alanine	E =	Glutamic acid
I =	Isoleucine	Y =	Tyrosine	C =	Cysteine
M =	Methionine	H =	Histidine	W =	Tryptophan
V =	Valine	Q =	Glutamine	R =	Arginine
S =	Serine	N =	Asparagine	G =	Glycine
P =	Proline	K =	Lysine		

FIG. 2 Alignment of the deduced amino acid sequences for *Tfil* , *Taq* and *Tfl* DNA polymerases. Only the residues differing from the *Tfil* sequence are shown for *Taq* and *Tfl*.

FIG. 3. Relative yield of PCR product at differing pHs. PCR was performed using primers P4 and P5 and a pUC18 plasmid containing the sheep *myo*D gene as described in the General

Methods section. The relative yield at each pH value was estimated by electrophoresis of 10 ml of reaction product on a 2% agarose gel which was stained with ethidium bromide and photographed (panel A). The intensity of the bands was estimated by densitometer scanning of the negative (negatives used in figure) and the values plotted (panel B) relative to the darkest and lightest area of the photograph.

FIG. 4 The Km of *Tfil* polymerase for Mg^{2+} ions. The Km of *Tfil* polymerase for $MgCl_2$ in the solid phase assay was determined by varying the amount of $MgCl_2$ from the standard 5 mM, and measuring the incorporation of radio-nucleotide after a 1h reaction at 70°C as in the standard solid phase assay. The plot shows the number of incorporated counts at each $MgCl_2$ concentration tested and the inset presents the data as an Eadie-Hofstee plot.

FIG. 5 The activity of *Tfil* polymerase and *Taq* polymerase at different KCl concentrations. The activity of 0.05 units of each polymerase was determined in the solid phase assay described below except that various amounts of KCl were included in the reaction at the indicated concentrations.

FIG. 6. PCR amplification using *Tfil* DNA polymerase with different buffers and templates. Panel (A) shows the PCR amplification of the 16S rRNA genes from *Thermus flavus* using genomic DNA and primers P6 and P7, and panel (B) shows amplification of a 280 bp fragment of *myoD* from a plasmid template using primers P4 and P5. The amplifications were performed as described in the General Methods section with differing concentrations of $MgCl_2$, Tween 20 and $(NH_4)_2SO_4$ as indicated. For amplification of the rRNA genes the PCR buffer was 50 mM Tris-HCl pH 8.8 containing 1.5 mM $MgCl_2$, 400 mM dNTPs and either 0%, 0.01%, 0.02% or 0.04% Tween 20 (lanes 2 to 6 respectively). *Taq* polymerase was used in the same buffer with 0.01% Tween 20. Amplifications from plasmid DNA were also as described in the General Methods section. The PCR buffer was 50 mM Tris-HCl pH 8.8 with the following

additives: lanes 8, 9, 10 and 11 each contained, 15 mM $(NH_4)_2SO_4$, 0.01% Tween 20 and 1, 1.5, 2.0, or 2.5 mM $MgCl_2$ respectively; lane 12 contained 1.5 mM $MgCl_2$, 0.01% Tween 20; lane 13 contained 1.5 mM $MgCl_2$, 25 mM $(NH_4)_2SO_4$, 0.01% Tween 20; lane 14 contained 1.5 mM $MgCl_2$ and 15 mM $(NH_4)_2SO_4$. The reaction products from both PCR reactions were analysed by electrophoresis of 5 ml of the 50 ml reaction on an 0.8% (panel A) or 2% (panel B) agarose gel stained with ethidium bromide. The markers are the BRL 1 kb ladder.

FIG. 7. Sensitivity of RT/PCR amplifications. The sensitivity of RT/PCR was determined by serially diluting an RNA transcript of α-lactalbumin and using the dilutions as template for the RT/PCR reaction. RT/PCR amplification of α-lactalbumin was performed by using both primers in the RT reaction and 2 units of *Tfil* DNA polymerase as described in the General Methods section. Lanes 1 to 5 correspond to 32 pg, 160 pg, 800 pg, 4 ng, 20 ng and 100 ng of RNA template respectively. The reaction products were analysed by electrophoresis of 5 ml of product on a 0.8% agarose gel which was stained with ethidium bromide. Lane 6 is the BRL 1 kb ladder. The gels are illustrated in negative form, for clarity.

FIG. 8. RT/PCR amplification from total cellular RNA. Panel A shows the amplification of a 200 bp region of topoisomerase IIa from total cell RNA isolated from Jurkat cell lines that were amsacrine-resistant, adriamycin-resistant, or normally sensitive (lanes 1, 2 and 3 respectively). Approximately 400 ng of RNA was used for each amplification and 5 units of polymerase. The sense primer was added with the PCR buffer after the RT as described in methods. Panel B shows the amplification of α-lactalbumin (lane 4) from 500 ng of total cell RNA as described in Experimental Procedure. Reaction products were analysed by electrophoresis of 10 ml on a 2% (panel A) or a 0.8% (panel B) agarose gel and ethidium bromide-staining. Molecular weight markers are the BRL 1 kb ladder.

FIG. 9. Phylogenetic tree of *Thermus* isolates *ex* Saul *et al.* (1993. Int. J. System. Bacteriol. **43**, 754-760). The tree is based on 16S rRNA sequence data and generated by the method of

maximum parsimony. The scale bar represents an expected nucleotide substitution rate of 0.01 per site. The values on the branches indicate the levels of support from 100 bootstrapped trees. The shaded boxes delineate clades supported by more than 80% of the bootstrapped trees. Also included are the geographic origins of the isolates.

PREFERRED EMBODIMENTS

GENERAL METHODS USED

Assay for polymerase activity. Thermophilic DNA polymerase activity was measured using a solid phase assay as described by Day *et al.* 1993 Anal Biochem, **211**:174-176.

Polyacrylamide gel electrophoresis. SDS PAGE was performed essentially as described by Laemmli (1970) *Nature* **227**, 680-685; using 12.5% and 7.5% SDS polyacrylamide gels and a Pharmacia "Phastgel" system. Gels were silver stained as described in the Pharmacia "PhastSystem" Users Manual.

Detection of endonuclease, exonuclease and ribonuclease activities. The purified polymerase was tested for the presence of endonuclease and exonuclease activities by incubating enzyme with linearised and supercoiled plasmid or an RNA transcript. Ten units of DNA polymerase were incubated with 200 ng of linearised plasmid or uncut plasmid, or 500 ng of RNA transcript, in 20 mM Tris-HCl buffer pH 8.0 or pH 8.8 containing 5 mM Mg^{++} for 3 hr at 70°C or 37°C. Degradation of RNA or plasmid DNA, or relaxation of the plasmid; was determined by agarose gel electrophoresis and staining with ethidium bromide.

Dye-labelled primer sequencing. Sequencing using the ABI dye-labelled M13 forward sequencing primer was performed on a single stranded M13 template that contained a G:C rich insert from a 16S RNA gene from a thermophilic organism (pNZ2201). The sequencing

reactions were performed exactly as described in the Applied Biosystems *Taq* Dye Primer Cycle Sequencing Kit Manual except differing buffer conditions for use with the *Tfil* polymerase.

Dye-labelled dideoxynucleotide terminator sequencing. The same template as for dye-labelled primer sequencing was used. The protocol suggested by ABI was followed with some modifications. Buffers containing varying amounts of MgCl2 and salts were tried in place of the recommended "TACS" buffer. The ratio of deoxynucleotide to dye labelled terminator was also varied by adding increasing amounts of the ABI supplied nucleotide mix (750 mM dITP, 150 mM dATP, 150 mM dTTP and 150 mM dCTP). The standard sequencing cycle of 25 cycles of 96°C for 30s, 50°C for 15s, 60°C for 4 min, was varied by increasing the extension temperature from 60°C to either 65°C or 70°C for 4 min.

RNA preparation. Total cellular RNA was isolated from cell lines derived from MA104 cells (α-lactalbumin cell lines) or Jurkat cells (topoisomerase IIa cell lines) as described by Sambrook *et al.* 1989 in "Molecular Cloning: A Laboratory Manual". Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY. RNA run-off transcripts of the cloned a-lactalbumin gene have been described elsewhere (L'Hullier *et al.* 1992. *EMBO J.* **11**, 4411-4418).

Oligodeoxynucleotide primers used.

P1	5'-CGGATCCCAAATCAGGCTTTTATTCGG-3'
P2	5'-ATTTTCCATGATCTGCTTATGAG-3'
P3	5'-TATTCCCAAACTGGATGATGCTAAT-3'
P4	5'-GCGGGCCCTTAGGCTACTACGG-3'
P5	5'-GCTCGATATCCCGGCGTGGGGC-3'
P6	5'-AAAGGAGGTGATCCAGCCGCACCTTC-3'

P7 5'-TTGTTGGAGAGTTTGATCCTG-3'

T7 5'-AATACGACTCACTATAG-3'

PCR and RT/PCR coupled amplifications. All PCR amplifications (50 ml) were performed in 50 mM Tris-HCl buffer pH 8.8 containing 1.5 mM $MgCl_2$, 0.01% Tween 20, 400 mM each of dATP, dCTP, dGTP and dTTP (dNTPs), 10 pmoles each of forward and reverse primer, 2 units of DNA polymerase and the indicated amount of template unless otherwise stated. Reactions were overlaid with 50 ml of mineral oil prior to amplification. RT/PCR amplifications had the initial reverse transcription performed in a reaction volume of 25 ml and contained 50 mM Tris-HCl buffer pH 8.8, 2 mM $MgCl_2$, 0.05% Tween 20, 0.05% Nonidet P40, 400 mM of each dNTP, and 100 ng of reverse primer P1 or p2. The amount of enzyme used was varied between 1 and 5 units, and often, the forward primer T7 or P3, which are necessary for the PCR amplification step, was also included in the reverse transcription mixture. Reverse transcriptions were performed at 60°C (under mineral oil) with the enzyme added after the 60°C reaction temperature had been attained. Reverse transcription was allowed to proceed for 5 min after which the reaction temperature was raised to 94°C and the reaction diluted with 75 ml of PCR buffer (50 mM Tris-HCl buffer pH 8.8 containing 1.5 mM $MgCl_2$, 0.01% Tween 20 and 400 mM dNTPs) that had been preheated to 94°C. The forward primer was added at this point if it had been omitted from the initial reverse transcription reaction. The reaction was overlaid with more mineral oil and then amplified in a thermal cycler using 35 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 2 min for both α-lactalbumin and topoisomerase IIa templates. This 'hot start' procedure helped prevent mis-priming during the reverse transcription and subsequent PCR amplification.

Optimisation of PCR buffer conditions. The buffer conditions that gave maximum yield of PCR product were determined using primers P4 and P5, and 10 ng of a plasmid template containing the sheep *myoD* gene (Huynen, *et al.* 1991. *Nucl. Acids Res.* **20**, 374) and also with primers P6 and P7 (16S rRNA gene primers), with 50 ng of *T. flavus* genomic DNA as

template. The PCR buffer was systematically varied by altering the amount of $MgCl_2$, $(NH_4)_2SO_4$ and Tween 20 added to the 50 mM Tris-HCl buffer pH 8.8. Each dNTP was present at 400 mM and 1 unit of polymerase was used per reaction. The PCR cycle used was 25 cycles of 94°C for 1 min., 60°C for 1 min. and 72°C for 1 min. for the plasmid amplification and 35 cycles of 94°C for 1 min., 53°C for 1 min. and 72° for 2 min., with an initial denaturing step of 4 mins at 94°C.

Determination of the pH optimum for PCR. The above primers were used (100 ng each) to determine the yield of PCR product obtained with different buffers. The reactions all contained 400 mM of each dNTP, 1.5 mM $MgCl_2$, 0.01% Tween 20, 10 ng of template DNA, 1 unit of polymerase and 50 mM of either Tris, Bis-Tris propane (1,3bis[tris(Hydroxymethyl)methlyamino]propane) or EPPS (N-[2-Hydroxyethyl]-piperazine-N'-[3-propanesulphonic acid]) buffer in the pH range 7.4 to 9.5. The reactions were thermally cycled for 25 cycles of 94°C for 1 min., 60°C for 1 min. and 72°C for 1 min. The yield of PCR product was determined by electrophoresis of a 10 ml portion of the reaction product followed by densitometer scanning of a photographic negative of the ethidium bromide stained agarose gel.

GENE CLONING OF THE *THERMUS FILIFORMIS* DNA POLYMERASE GENE.

Gene isolation. The gene encoding the DNA polymerase was isolated from *Thermus filiformis* genomic DNA using the technique of the Polymerase Chain Reaction (PCR). Two oligonucleotide primers were used for the amplification:

Primer A, 5'-CAC***GAATTC***GGGGATGCTGCCCCTCTTTGAGCCCAAG-3'

Primer B, 5'-GTG***GGATCC***ATCACTCCTTGGCGGAGAGCCAGT-3'.

The primers contained an *Eco*RI site and a *Bam*HI site at their 5' ends respectively for A and B and are derived from the primers used by Engelke *et al.* (1990) Anal. Biochem. **191**: 396-400; to clone the *Taq* polymerase gene from *Thermus aquaticus.* The PCR amplification was performed in 50 ml of buffer containing 10 mM Tris-HCl pH 8.8, 2.5 mM $MgCl_2$, 50 mM KCl, 400 µM dNTPs, 10 pmoles of each primer and 2.5 units of *Taq* polymerase (AmpliTaq; Cetus Corp). The target sequence was amplified by first denaturing at 94°C for 4 min followed by 35 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 4 mins. Thermal cycling was performed in a Perkin Elmer Cetus thermal cycler.

Cloning and expression in pUC18 The PCR product was purified by electrophoresis of 100 ml of the PCR mixture on a 0.8% Tris-Acetate agarose gel. The 2.6 kb band of the polymerase coding region was purified from the agarose by binding and elution from glass powder (Geneclean, Bio 101, San Diego, CA). The product was prepared for cloning by treating with 1 unit each of T4 polynucleotide kinase, T4 DNA polymerase and Klenow fragment in 66 mM Tris-HCl pH 7.6 containing 6 mM $MgCl_2$ and 50 µM dNTPs for 15 mins at 37°C. The DNA was then extracted with phenol/chloroform and then chloroform prior to precipitation with ethanol. The pellet was resuspended and digested with *Eco*RI to give one cohesive end for directional cloning. The DNA was ligated into pUC18 that had been digested with *Eco*RI and *Sma*I, and the ligated products were transformed into *E. coli* strain DH. Transformants were grown on L-agar containing 100 µg/ml ampicillin and IPTG/Xgal (5 mM and 33 µg/ml respectively) to allow selection of recombinants. White colonies, in which the LacZ gene had been insertionally inactivated, were picked and grown in L-broth containing 100 µg/ml ampicillin, and plasmid DNAs were prepared by alkaline lysis. The plasmids were checked for insertions by digestion with *Pvu*II. Those recombinants containing inserts were grown in L-broth containing ampicillin and tested for the expression of thermophilic DNA polymerase by induction of exponentially growing culture with 0.5 mM IPTG and assaying the heat-treated

extracts for DNA polymerase activity. A recombinant expressing polymerase was obtained. The strain was designated PB5900 and the plasmid pNZ2300.

Cloning and expression in pT7-7 The *Tfil* polymerase coding region was excised from pNZ2300 by digestion at the unique *Eco*RI and *Hin*dIII sites of the multiple cloning site of the vector. The *Eco*RI - *Hin*dIII restriction fragment was purified and ligated into the expression vector pT7-7 that had been similarly digested with *Eco*RI and *Hin*dIII. The ligated DNA was transformed into *E. coli* strain JM101 and transformants were selected by plating on L-agar containing ampicillin (60 µg/ml). Colonies were grown and plasmid prepared by alkaline lysis. Colonies containing insertions were identified by digestion with *Eco*RI. The new recombinant strain was designated PB5902 and the new plasmid pNZ2302.

The DNA polymerase coding sequence in the plasmid pNZ2302 was out of frame with respect to the phage T7 gene 10 product. The coding region for the DNA polymerase was placed in frame with the ATG of the gene 10 protein by digestion of pNZ2302 with *Eco*RI, followed back-filling of the *Eco*RI site with Klenow fragment and religation. This new construct was designated pNZ2303.

Plasmid pNZ2303 was used to transform competent *E. coli* K12 containing plasmid pGP1-2. Transformants were selected by plating on L-agar containing ampicillin and kanamycin (60 µ g/ml each) and were grown at 32°C. Colonies were picked and grown in L-broth containing ampicillin and kanamycin (60 µg/ml each) at 32°C, then induced and tested for the expression of active thermophilic DNA polymerase in heat treated induced extracts, as described above. A recombinant strain expressing active thermostable DNA polymerase (PB5905) was obtained.

PURIFICATION OF *Tfil* POLYMERASE

Growth of the recombinant *E.coli* strain PB5905 for the production of *Tfil* DNA polymerase. Ten litres of medium (L-broth containing 60 µg/ml of ampicillin) were inoculated with a 1/40 volume stationary phase culture of PB5905 and grown in batch culture. The cells were grown at 32°C to an absorbance at 600 nm of 0.5, then induced by rapidly raising the growth temperature to 42°C, and maintained at this temperature for a further 45 min. The temperature was then reduced to 37°C and the cells grown for a further 2 h before harvesting and washing in 10 mM Tris-HCl buffer pH 8.0 containing 1 mM EDTA (TE buffer). Approximately 20 g wet weight of cells were obtained and may be frozen prior to use.

Purification of *Tfil* DNA polymerase. The cells were thawed and resuspended in 100 ml of TE buffer. The protease inhibitor phenylmethylsulphonylfluoride was added to a final concentration of 0.5 mM and the cells broken by one passage through a french pressure cell. The broken cells were incubated in a 70°C water bath in a preheated 1 litre flask for 30 min after which they were cooled on ice. Precipitated protein and cell debris were removed by centrifugation at 32,600x g for 15 min at 4°C. The supernatant was decanted and dialysed for 12 h against 3 changes of 5 litres of TE buffer and filtered through a 0.22 µm filter to remove particulate matter.

A Pharmacia "Q-Hiload 26/10" was washed with 1M KCl then equilibrated with buffer A (20 mM Tris-HCl pH 8.0 containing 1 mM EDTA and 10% v/v glycerol) at 4°C. The heat-treated extract (85 ml) was applied to the column using a peristaltic pump (Pharmacia P-1) at a flow rate of 2 ml/min. Unbound material was washed from the column with 50 ml of buffer A. The column was then attached to a Pharmacia FPLC system and bound protein eluted using a linear KCl gradient in buffer A (0 - 400 mM KCl in 300 ml) at a flow rate of 5 ml/min. The elution of protein was monitored at 280 nm and 7.5 ml fractions collected and placed immediately on ice.

Fractions containing DNA polymerase activity were pooled and dialysed against 3 changes of 5 litres of 20 mM potassium phosphate pH 7.5 (buffer B). A 5 ml heparin sepharose column (Pharmacia HiTrap heparin, 5 ml) is equilibrated with buffer B at 4°C and all of the dialysed extract applied in two identical runs, each using half of the extract (70 ml). The sample was applied at a flow rate of 2 ml/min, and the column washed with 10 ml of buffer B. Bound polymerase was eluted with a linear gradient of KCl in buffer B (0 - 500 mM KCl in 30 ml). Pure DNA polymerase eluted as a single peak at about 350 mM KCl. Active fractions were pooled and diluted with an equal volume of glycerol and stored at -20°C. Approximately 20 mg of purified polymerase can be recovered by this purification scheme.

CHARACTERISTICS OF *Tfil* POLYMERASE

Analysis of the sequence of the *Tfil* polymerase gene. The sequence of the DNA polymerase gene was determined by inserting and sequencing restriction enzyme fragments of the PCR isolated *Tfil* gene cloned into M13. These were sequenced by using the standard M13 forward sequencing primers with dye-primer chemistry and an Applied Biosystems 373A Automated DNA Sequencer. The DNA sequence of *Tfil* polymerase and the derived amino acid sequence of the enzyme are shown in Fig 1. When compared with other available *Thermus* polymerase gene sequences, the *Tfil* polymerase sequence shows the following percentages of identical amino acids: *Tfil* vs. *Taq* = 84%; *Tfil* vs. *Tfl* = 82%. This represents a greater difference between *Tfil* polymerase and the others than is seen between *Taq* and *Tfl* (*Taq* vs. *Tfl* = 86% identity). An alignment of the deduced amino acid sequences of these enzymes is Fig 2.

The *Tfil* polymerase possesses several advantages over other polymerases, some of which advantages are outlined here. Characterisation of *Tfil* polymerase shows that its properties are clearly different from *Taq* polymerase, as *Tfil* polymerase has maximum activity in low salt

buffers, possesses reverse transcriptase activity and has a conveniently greater affinity for dye-labelled dideoxynucleotide analogues. Using the 5'-3' exonuclease assay described in the methods section, no detectable exonuclease activity can be measured for *Tfil* polymerase or *Taq* polymerase.

Tfil polymerase has maximum activity between pH8.5 and 9.0 as demonstrated by comparing the yield of PCR products in different buffers (Fig 3). The enzyme has maximal activity in low salt buffers and has an absolute requirement for Mg^{2+} ions (Km 1.5mM) (Fig 4). The enzyme is completely, but reversibly inhibited by the addition of EDTA. The most suitable buffer for both DNA polymerase and Reverse Transcriptase activities is 50 mM Tris-HCl buffer pH 8.8 containing 1.5 mM $MgCl_2$ and 0.01% Tween 20. In this buffer *Tfil* polymerase has a half-life at 94°C of approximately 28 minutes.

Reverse Transcriptase activity. The reverse transcriptase activity of thermophilic DNA polymerases has been described previously (Jones and Foulkes 1989. *Nucl. Acids Res.* **17**, 8387-8388; Meyers and Gelfand 1991. *Biochem.* **30**, 7661-7666) with the activity of *Taq* polymerase being poor compared with that of *Tth* polymerase. *Tth* polymerase shows similar reverse transcriptase activity to that of *Tfil* polymerase but is different in that *Tth* polymerase has greatest activity in buffers containing high salt (100 mM KCl). Furthermore, both *Taq* polymerase and *Tth* polymerase require Mn^{2+} supplied as $MnCl_2$ for activity. *Tfil* polymerase shows the same high level of reverse transcriptase activity as *Tth* pol but differs in that no activity is obtained when $MnCl_2$ was used instead of $MgCl_2$ for reverse transcription. The use of Mg^{2+} as opposed to Mn^{2+} is a significant difference between *Tth* pol and *Tfil* polymerase. Reverse transcription in the presence of Mg^{2+} ions is preferable for two-step reactions where RNA is to be copied followed by DNA extension as Mn^{2+} ions are known to lower the fidelity of DNA synthesis (Beckman *et al.* 1985. *Biochemistry* **24**, 5810-5817). Low fidelity DNA synthesis is likely to lead to mutated copies of the original template. In addition, Mn^{2+} ions

have been implicated in an increased rate of RNA degradation, particularly at high temperatures and this can cause the synthesis of shortened products in a reverse transcription reaction.

Salt tolerance. *Tfil* polymerase shows greatest activity at low KCl concentrations but retains a greater percentage activity at high salt concentrations than does *Taq* polymerase which has maximal activity at 25 mM KCl in the solid phase DNA polymerase assay but loses activity rapidly as the concentration of KCl increase (Fig 5). *Tfil* polymerase was much more tolerant of high salt concentrations, maintaining 35% of its activity at 150 mM KCl whilst *Taq* polymerase showed 2% activity at this concentration. This result indicates that *Tfil* polymerase has far greater utility that *Taq* polymerase when high salt buffers are required.

DNA sequencing. There are two sequencing chemistries that can be used on the ABI 373A Automated DNA Sequencer (currently the most utilised DNA sequencer world-wide). Both methods incorporate fluorescent dyes to label terminated reactions and generate sequencing ladders which are detected by a laser/photomultiplier excitation/detection system. The chemistries use either a dye-labelled primer and standard dideoxynucleotides to terminate the sequence or an unlabelled primer and dye-labelled dideoxynucleotide terminators. The sequence obtained using dye-labelled primers is usually superior in both quality and read length than that acquired using dye-labelled terminators but dye-labelled terminators offer the advantage that any primer may be used to generate sequence data.

Tfil polymerase was able to generate long stretches of unambiguous sequence data when using dye-labelled primer and M13 single stranded templates. The quality of sequence obtained with *Tfil* polymerase with this chemistry similar to that produced by *Taq* polymerase. Good sequence data was obtained using 50 mM Tris-HCl buffer pH 8.8 containing 1.5 mM MgCl2 and 0.01% Tween 20 (v/v) with the ABI recommended reaction cycle. Good sequence data

was obtained also with higher Mg^{2+}, Tween 20 and Tris-HCl concentrations (up to 4 mM, 0.1% v/v and 135 mM respectively).

Greater differences were found between *Taq* polymerase and *Tfil* polymerase when sequencing data was generated using dye-labelled dideoxynucleotide terminators. *Tfil* polymerase incorporated the terminators more efficiently than *Taq* polymerase at the 60°C extension temperature recommended for *Taq* polymerase in dye-terminator sequencing, such that the sequence terminated prematurely and piled up at the beginning. Full length sequence was obtained by reducing the terminator:dNTP ratio which would offer an advantage of economy as the cost of the dye-labelled terminators represent the major portion of the process cost.

PCR amplification of DNA. The polymerase was suitable for the efficient PCR amplification using plasmid or genomic DNA as template (Fig 6). As was found for the polymerase assay, increasing KCl concentrations in the PCR buffer caused a decrease in yield of PCR product. By varying the KCl, $(NH_4)_2SO_4$, $MgCl_2$ and Tween 20 concentrations in the PCR buffer (Fig 6) and measuring the yield of PCR product, it was found that the most suitable buffer for PCR amplification using *Tfil* polymerase was 50 mM Tris-HCl buffer pH 8.8 containing 1.5 mM $MgCl_2$ and 0.01% Tween 20. For some templates, a higher $MgCl_2$ and Tween 20 concentration (2 mM and 0.02% respectively) improved the PCR amplification by reducing the number and intensity of unwanted products.

PCR product was obtained over a wide range of pH values using the primers P4 and P5, with the greatest yield obtained at pH 8.5 to 9.0 (Fig 3). During PCR, the polymerase was estimated to have an extension rate of about 1000 bases/min. as determined by a primer extension assay (Carballeira *et al.* 1990. *Biotechniques* **9**, 274-281) using the nucleotide and buffer conditions employed for PCR.

Reverse transcription/ PCR amplifications. *Tfil* polymerase also shows very efficient reverse transcriptase activity on a variety of RNA templates. Using the protocol described in methods, *Tfil* polymerase was able to reverse transcribe and PCR amplify (RT/PCR) as little as 32 pg of template RNA (Fig 7). Greater sensitivity could be achieved by using more PCR cycles or by analysing a larger proportion of the reaction product on an agarose gel. The polymerase was also able to RT/PCR amplify target present in more complex samples. Fig 8 shows RT/PCR amplification from total cellular RNA of a 200 bp fragment from topoisomerase IIa message and the entire 800 bp message of a-lactalbumin. Possible contamination of the RNA template with DNA that potentially could be a target in the PCR reaction was tested by attempting to PCR amplify without initial reverse transcription. The absence of any PCR product using either *Taq* polymerase or *Tfil* polymerase with conditions that were optimal for PCR amplification showed that there was no contamination of the RNA template with DNA and that the product formed was from PCR amplification of cDNA generated in the first phase of the RT/PCR. Also, the buffer used for reverse transcription was found not to be suitable for PCR amplification (no product was obtained using the primers P4 and P5 with the plasmid template) and that the buffer used for PCR failed to give product in RT/PCR when it was used for the reverse transcription step.

DEPOSIT

Thermus filiformis is a named strain already deposited in the American Type Culture Collection under ATCC accession number 43280.

PHYLOGENY OF THE SPECIES *Thermus filiformis.*

The species *Thermus filiformis* is distinct from *Thermus aquaticus*, *"Thermus flavus"* and *"Thermus thermophilus"* and this has been demonstrated by a number of methods. Morphologically, *Thermus filiformis* is atypical of the genus in that it is filamentous and analysis of the 16S rRNA gene from this organism and others shows that it is part of a group

of organisms unique to New Zealand (Fig 9). The complete 16S rRNA genes of twenty *Thermus* isolates have been sequenced (Saul *et al.* 1993. Int. J. System. Bacteriol. **43**, In press) and the data subjected to a phylogenetic analysis using the method of maximum parsimony. The data suggest a close relationship between *"T. flavus"* and *"T. thermophilus"* but that *T. aquaticus* and *T. filiformis* are more distally related to these organisms and to each other. The validity of the branches within the tree were assessed by "bootstrapping". Felsentein, 1985. *Evolution* **39**, 783-791, suggests that confidence limits of the data should be set so that a branch should only be accepted if it occurs in more than 95% of the bootstrapped trees. *T. filiformis, T. aquaticus* and the *"T. flavus"/"T. thermophilus"* pair are separated in 100% of the trees (Fig 9).

These results suggest that the cloned enzyme of this invention is a versatile thermostable DNA polymerase that is suitable for DNA synthetic activity from both DNA and RNA templates, as well as being highly suitable for automated DNA sequencing. *Tfil* polymerase performs both of these reaction in the same buffer, has a high salt tolerance and is suitable for use in cycled fluorescent DNA sequencing using both dye-primers and dye-terminators. Its greater affinity than existing enzymes for dye-terminators means that a reduction in consumption of these costly chemicals can be attained. All these properties in a single enzyme make *Tfil* polymerase a very useful tool for the molecular biologist.

CITED LITERATURE

Kaledin *et al.* 1980. *Biokimiya* **44**, 644-651
Kaledin *et al.* 1981. *Biokimiya* **46**, 1247-1254
Kaledin *et al.* 1982. *Biokimiya* **47**, 1515-1521
Ruttimann, *et al.* 1985. *Eur. J. Biochem* **149**, 41-46
Neuner *et al.* 1990. *Arch. Microbiol.* **153**, 205-207
Lawyer *et al.* 1989. *J. Biol. Chem.* **264**, 6427-6437
Engelke *et al.* 1990. *Anal. Biochem.* **191**
Lundberg *et al.* 1991. *Gene* **108**, 1-6
Perler *et al.* 1992. Proc. Natl. Acad. Sci. USA **89**, 5577-5581

Brock *et al.* 1969. J. Bacteriol. **98**, 289-297
Oshima and Imahori. 1974. J. Syst. Bacteriol. **24**, 104-112
Saiki *et al.* 1972. Agric. Biol. Chem. **34**, 2357-2366
Hudson *et al.* 1987. Int. J. Syst. Bacteriol. **37**, 431-436
Saul *et al.* 1993. Int. J. Syst. Bacteriol. **43**, 754-760
Jones, *et al.* 1989. *Nucl. Acids Res.* **17**, 8387-8388
Meyers and Gelfand 1991. *Biochem.* **30**, 7661-7666
Day *et al.* 1993 Anal. Biochem., **211;** 174-176
Engelke *et al.* (1990) Anal. Biochem. **191**: 396-400

Finally it will be appreciated that various alterations and modifications may be made to the foregoing without departing from the scope of this invention as claimed.

CLAIMS:

1. A purified heat-stable DNA polymerase enzyme (EC 2.7.7.7, a DNA nucleotidyltransferase DNA-directed enzyme), characterised in that it has reverse transcriptase activity in the presence of magnesium ions.

2. A DNA polymerase as claimed in claim 1, characterised in that it has reverse transcriptase activity in the substantial absence of manganese ions.

3. A DNA polymerase as claimed in claim 2, characterised in that it has been isolated from the eubacterium *Thermus filiformis*.

4. A DNA polymerase as claimed in claim 1, having an amino acid sequence *Tfil* substantially as shown in Fig 2.

5. A DNA polymerase as claimed in claim 1, having a pH-dependent activity as shown graphically in Fig 3.

6. A recombinant DNA sequence capable of encoding polymerase activity of the microorganism *Thermus filiformis*.

7. A recombinant DNA sequence as claimed in claim 6 wherein the DNA sequence is a recombinant DNA sequence capable of encoding amino acid residues 1 to 833 as shown in figure 1.

8. A recombinant DNA plasmid that comprises the DNA sequence substantially as shown in figure 1, inserted into plasmid vectors and which can be used to drive the expression of the thermostable DNA polymerase of *Thermus filiformis* in a host cell transformed with the plasmid.

9. A recombinant plasmid as claimed in claim 8, comprising the vector pUC18 carrying the *Thermus filiformis* DNA polymerase gene and designated pNZ2300.

10. A recombinant plasmid as claimed in claim 8, comprising the vector pT7-7 carrying the *Thermus filiformis* DNA polymerase gene and designated pNZ2303.

11. A host cell transformed with the recombinant DNA of the invention as claimed in claim 8.

12. The host cell as claimed in claim 11, being *Escherischia coli* / pNZ2300, the strain being designated PB5900.

13. The host cell as claimed in claim 11, being *Escherischia coli* / pNZ2300, the strain being designated PB5904.

14. A purified heat-stable DNA polymerase enzyme (EC 2.7.7.7, a DNA nucleotidyltransferase DNA-directed enzyme), isolated from the eubacterium *Thermus filiformis*.

15. A DNA polymerase as claimed in claim 14, having a molecular mass of about 93.4 KDa.

16. A DNA polymerase as claimed in claim 14, that is substantially in its native form.

17. A DNA polymerase as claimed in claim 14, that is substantially in its recombinant form.

18. A DNA polymerase as claimed in claim 14, having a dependency on magnesium ion concentration, as shown in Fig 4.

19. A DNA polymerase as claimed in claim 18, having reverse transcriptase activity in the substantial absence of manganese ions.

20. A DNA polymerase as claimed in claim 19, further comprising a DNA polymerase having an activity, referenced to Taq polymerase, as illustrated in Fig 5.

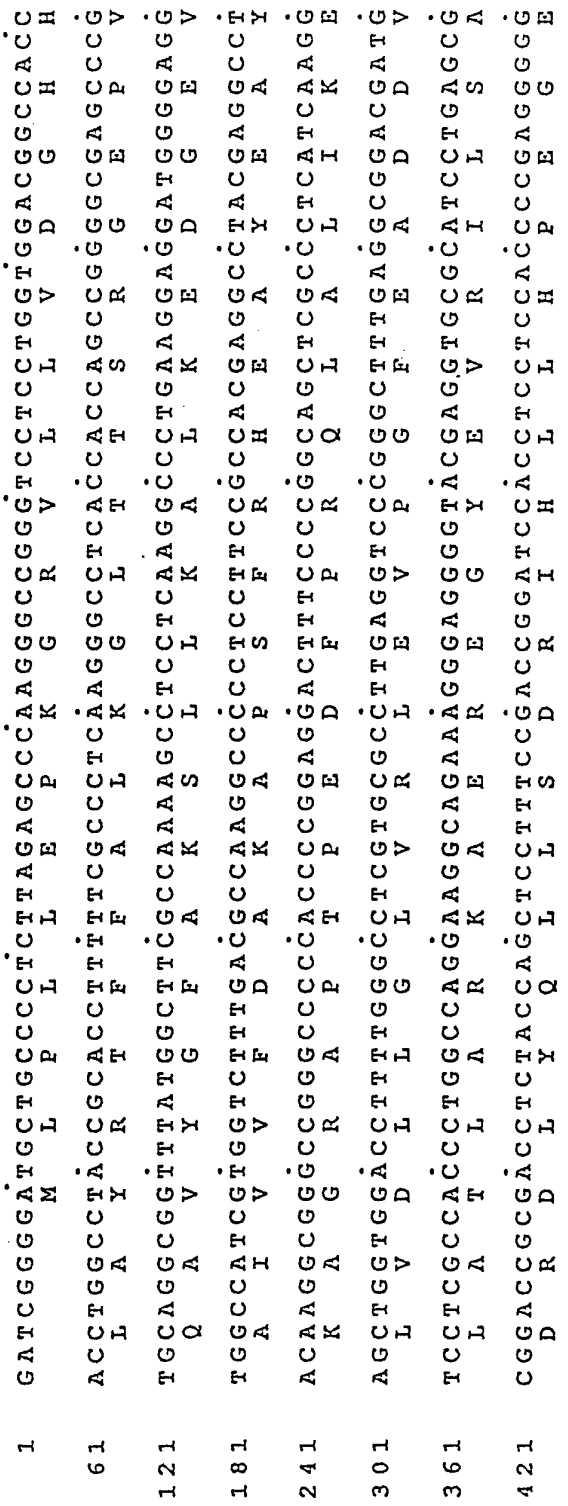


FIG 1-1

```
 841  GCCTCCTCCACGAGTTCGGCCTTCTGGAGGCCCCCAAGGAGGCGGAGGAGGCCCCCTGGC
        L   L   H   E   F   G   L   L   E   A   P   K   E   A   E   E   A   P   W   P

 901  CCCCGCCTGGAGGGGCCTTTTTGGGCTTCCTCCTCTCCCGCCCCGAGCCCATGTGGGCGG
        P   P   G   G   A   F   L   G   F   L   L   S   R   P   E   P   M   W   A   E

 961  AGCTTTTGGCCCTGGCGGGGGCCAAGGAGGGGCGGGTCCATCGGGCGGAAGACCCCGTGG
        L   L   A   L   A   G   A   K   E   G   R   V   H   R   A   E   D   P   V   G

1021  GGGCCCTAAAGGACCTGAAGGAGATCCGGGGCCTCCTCGCCAAGGACCTCTCGGTCCTGG
        A   L   K   D   L   K   E   I   R   G   L   L   A   K   D   L   S   V   L   A

1081  CCCTGAGGGAGGGCCGGGAGATCCCGCCGGGGGACGACCCCATGCTCCTCGCCTACCTCC
        L   R   E   G   R   E   I   P   P   G   D   D   P   M   L   L   A   Y   L   L

1141  TGGACCCGGGGAACACCAACCCCGAGGGGGTGGCCCGGCGGTACGGGGGGGAGTGGAAGG
        D   P   G   N   T   N   P   E   G   V   A   R   R   Y   G   G   E   W   K   E

1201  AGGACGCCGCCGCCCGGGCCCTCCTTTCGGAAAGGCTCTGGCAGGCCCTTTACCCCCGGG
        D   A   A   A   R   A   L   L   S   E   R   L   W   Q   A   L   Y   P   R   V

1261  TGGCGGAGGAGGAAAGGCTCCTTTGGCTCTACCGGGAGGTGGAGCGGCCCCTCGCCCAGG
        A   E   E   E   R   L   L   W   L   Y   R   E   V   E   R   P   L   A   Q   V

1321  TCCTCGCCCACATGGAGGCCACGGGGGGTGCGGCTGGATGTGCCCTACCTGGAGGCCCTTT
        L   A   H   M   E   A   T   G   V   R   L   D   V   P   Y   L   E   A   L   S

1381  CCCAGGAGGTGGCCTTTGAGCTGGAGCGCCTCGAGGCCGAGGTCCACCGCCTGGCGGGCC
        Q   E   V   A   F   E   L   E   R   L   E   A   E   V   H   R   L   A   G   H

1441  ACCCCTTCAACCTGAACTCTAGGGACCAGCTGGAGCGGGTCCTCTTTGACGAGCTCGGCC
        P   F   N   L   N   S   R   D   Q   L   E   R   V   L   F   D   E   L   G   L

1501  TACCCCCCATCGGCAAGACGGAGAAGACGGGCAAGCGCTCCACCAGCGCCGCCGTCCTGG
        P   P   I   G   K   T   E   K   T   G   K   R   S   T   S   A   A   V   L   E

1561  AGCTCTTAAGGGAGGCCCACCCCATCGTGGGGCGGATCCTGGAGTACCGGGAGCTCATGA
        L   L   R   E   A   H   P   I   V   G   R   I   L   E   Y   R   E   L   M   K

1621  AGCTCAAGAGCACCTACATAGACCCCCTCCCCAGGCTGGTCCACCCCAAAACCGGCCGGC
        L   K   S   T   Y   I   D   P   L   P   R   L   V   H   P   K   T   G   R   L
```

FIG 1-2

```
1681  TCCACACCCGCTTCAACCAGACGGCCACCGCCACGGGCCGCCTCTCCAGCTCCGACCCCA
        H   T   R   F   N   Q   T   A   T   A   T   G   R   L   S   S   S   D   P   N
1741  ACCTGCAGAACATCCCCGTGCGCACCCCCTTAGGCCAGCGCATCCGCAAGGCCTTCATTG
        L   Q   N   I   P   V   R   T   P   L   G   Q   R   I   R   K   A   F   I   A
1801  CCGAGGAGGGCCATCTCCTGGTGGCCCTGGACTATAGCCAGATCGAGCTCCGGGTCCTCG
        E   E   G   H   L   L   V   A   L   D   Y   S   Q   I   E   L   R   V   L   A
1861  CCCACCTCTCGGGGGACGAGAACCTCATCCGGGTCTTCCGGGAAGGGAAGGACATCCACA
        H   L   S   G   D   E   N   L   I   R   V   F   R   E   G   K   D   I   H   T
1921  CCGAGACCGCCGCCTGGATGTTCGGCGTGCCCCCCGAGGGGGTGGACGGGGCCATGCGCC
        E   T   A   A   W   M   F   G   V   P   P   E   G   V   D   G   A   M   R   R
1981  GGGCGGCCAAGACGGTGAACTTCGGGGTGCTCTACGGGATGTCCGCCCACCGCCTCTCCC
        A   A   K   T   V   N   F   G   V   L   Y   G   M   S   A   H   R   L   S   Q
2041  AGGAGCTCTCCATCCCCTACGAGGAGGCGGCGGCCTTCATCGAGCGCTACTTCCAGAGCT
        E   L   S   I   P   Y   E   E   A   A   A   F   I   E   R   Y   F   Q   S   F
2101  TCCCCAAGGTGCGGGCCTGGATCGCCAAAACCTTGGAGGAGGGGCGGAAGAAGGGGTACG
        P   K   V   R   A   W   I   A   K   T   L   E   E   G   R   K   K   G   Y   V
2161  TGGAGACCCTCTTCGGCCGCCGCCGCTACGTGCCCGACCTCAACGCCCGGGTGAAGAGCG
        E   T   L   F   G   R   R   R   Y   V   P   D   L   N   A   R   V   K   S   V
2221  TgcGGGAGGCGGCGGAGCGCATGGCCTTCAACATGCCCGTGCAGGGCACCGCCGCGGACC
        R   E   A   A   E   R   M   A   F   N   M   P   V   Q   G   T   A   A   D   L
2281  TCATGAAGCTGGCCATGGTGAAGCTCTTCCCCAGGCTCAGGCCCTTGGGCGTTCGCATCC
        M   K   L   A   M   V   K   L   F   P   R   L   R   P   L   G   V   R   I   L
2341  TCCTCCAGGTGCACGACGAGCTGGTCTTGGAGGCCCCAAAGGCGCGGGCGGAGGAGGCCG
        L   Q   V   H   D   E   L   V   L   E   A   P   K   A   R   A   E   E   A   A
2401  CCCAGCTGGCCAAGGAGACCATGGAAGGGGTTTACCCCCTCTCCGTCCCCCTGGAGGTGG
        Q   L   A   K   E   T   M   E   G   V   Y   P   L   S   V   P   L   E   V   E
2461  AGGTGGGGATGGGGGAGGACTGGCTCTCCGCCAAGGAGTGATGGATCCCACG   2512
        V   G   M   G   E   D   W   L   S   A   K   E   *
```

FIG 1-3

Fig 2


1018
500
396
344
220

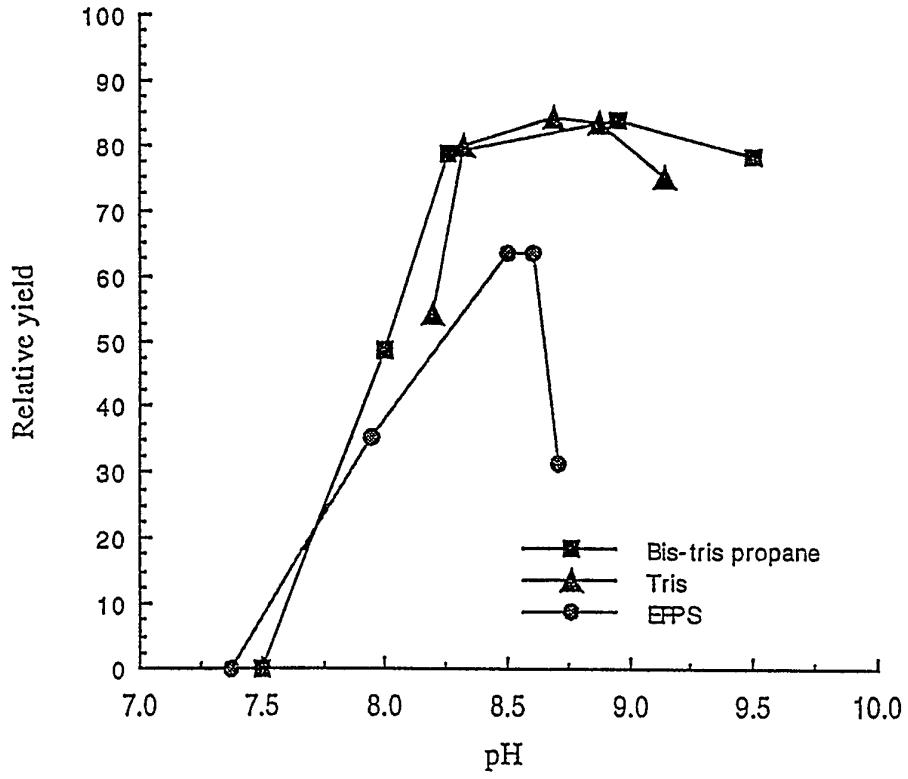
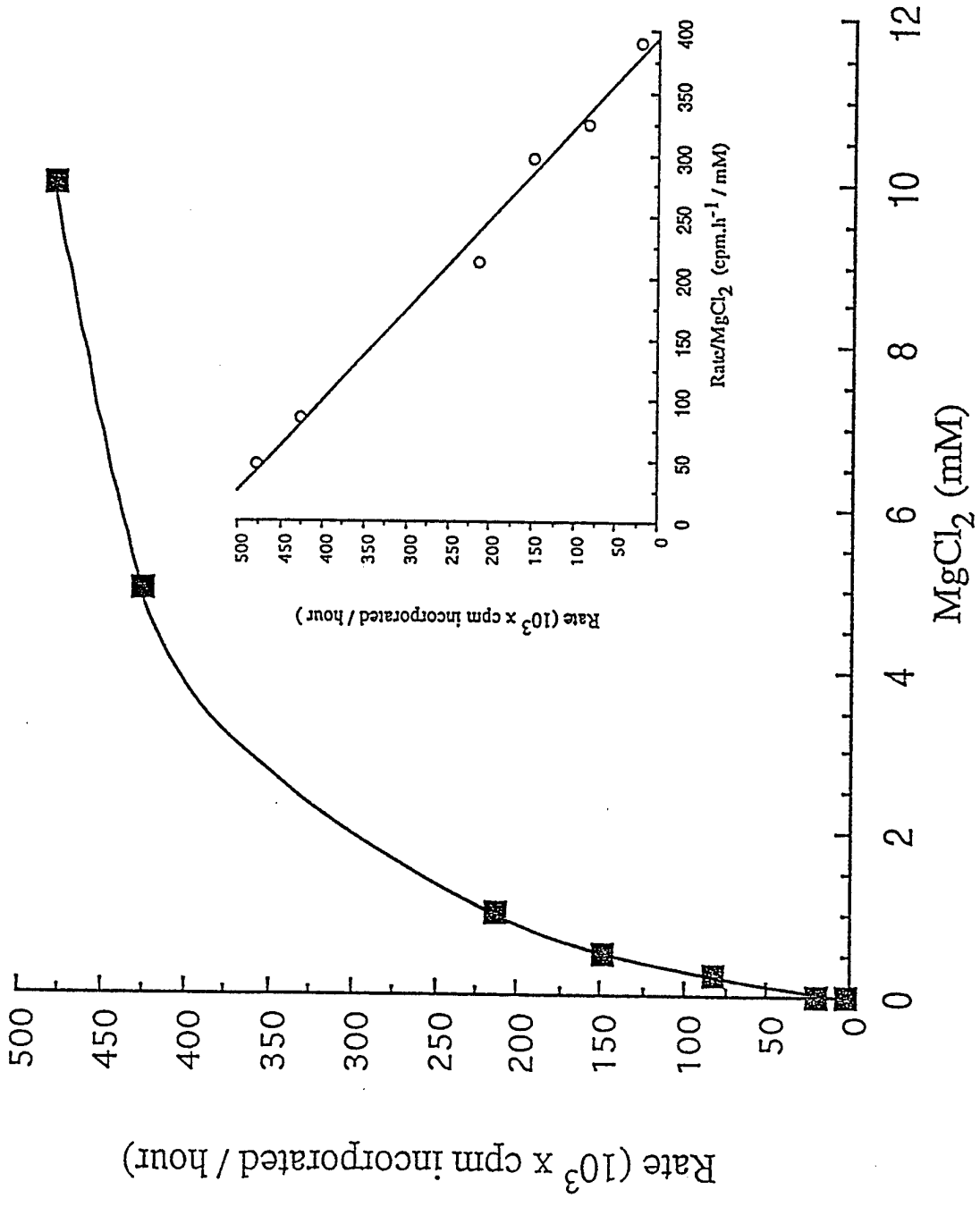

Relative yield
100
90
80
70
60
50
40
30
20
10
0
7.0
7.5
8.0
8.5
9.0
9.5
10.0
pH
Bis-tris propane
Tris
EPPS

Fig 3

Fig 4


Rate (10^3 x cpm incorporated / hour)
500
450
400
350
300
250
200
150
100
50
0
0
2
4
6
8
10
12
$MgCl_2$ (mM)
Rate (10^3 x cpm incorporated / hour)
Rate/$MgCl_2$ (cpm.h^{-1} / mM)

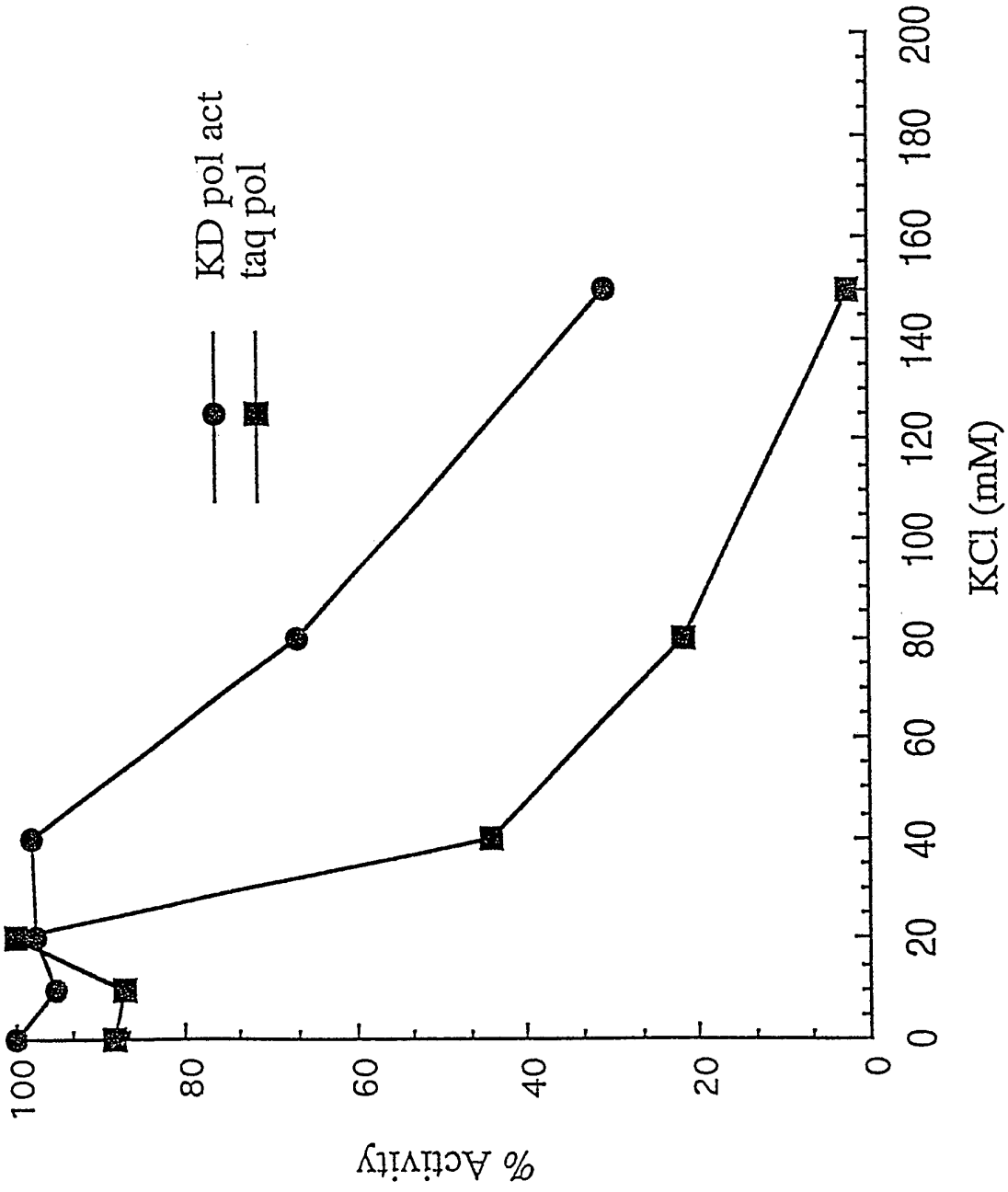

KD pol act
taq pol
% Activity
100
80
60
40
20
0
0
20
40
60
80
100
120
140
160
180
200
KCl (mM)

Fig 5

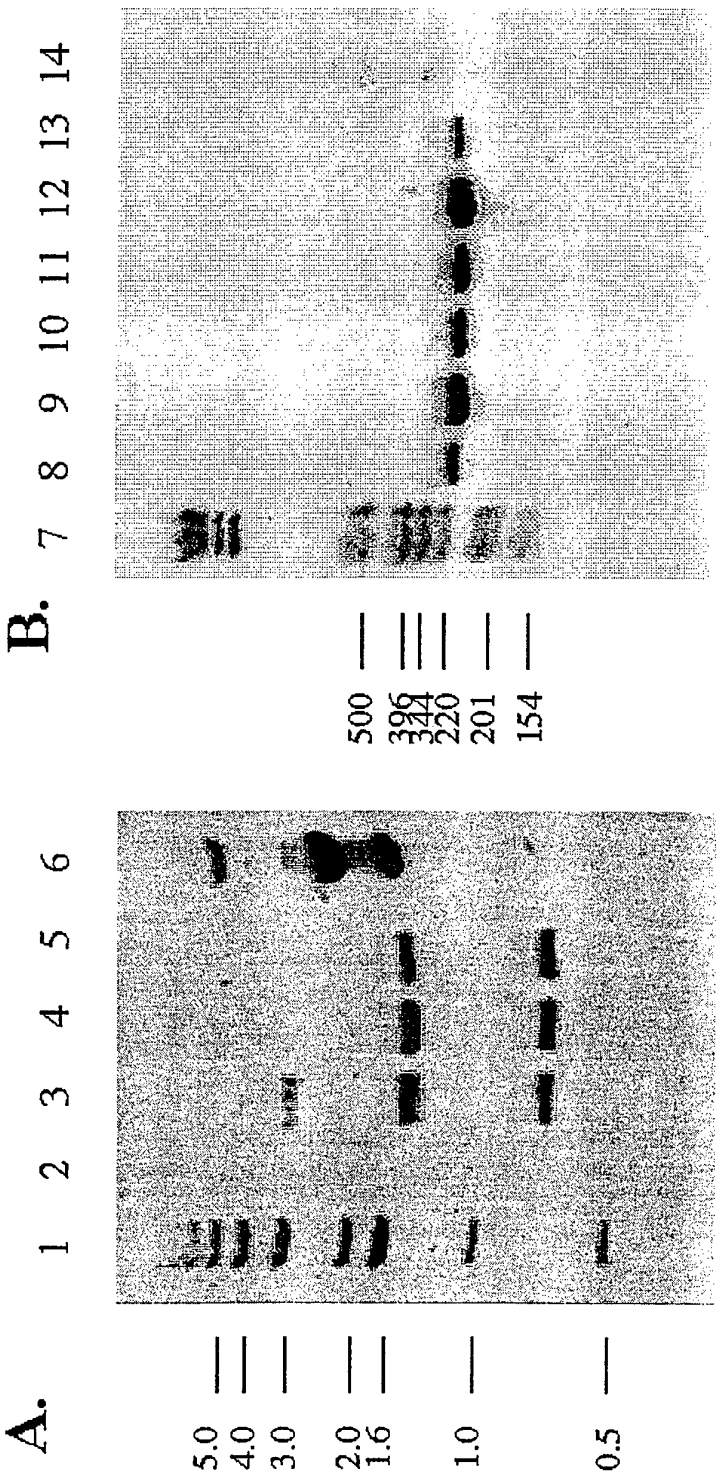

A.
1 2 3 4 5 6
5.0
4.0
3.0
2.0
1.6
1.0
0.5
B.
7 8 9 10 11 12 13 14
500
396
344
220
201
154

Fig 6

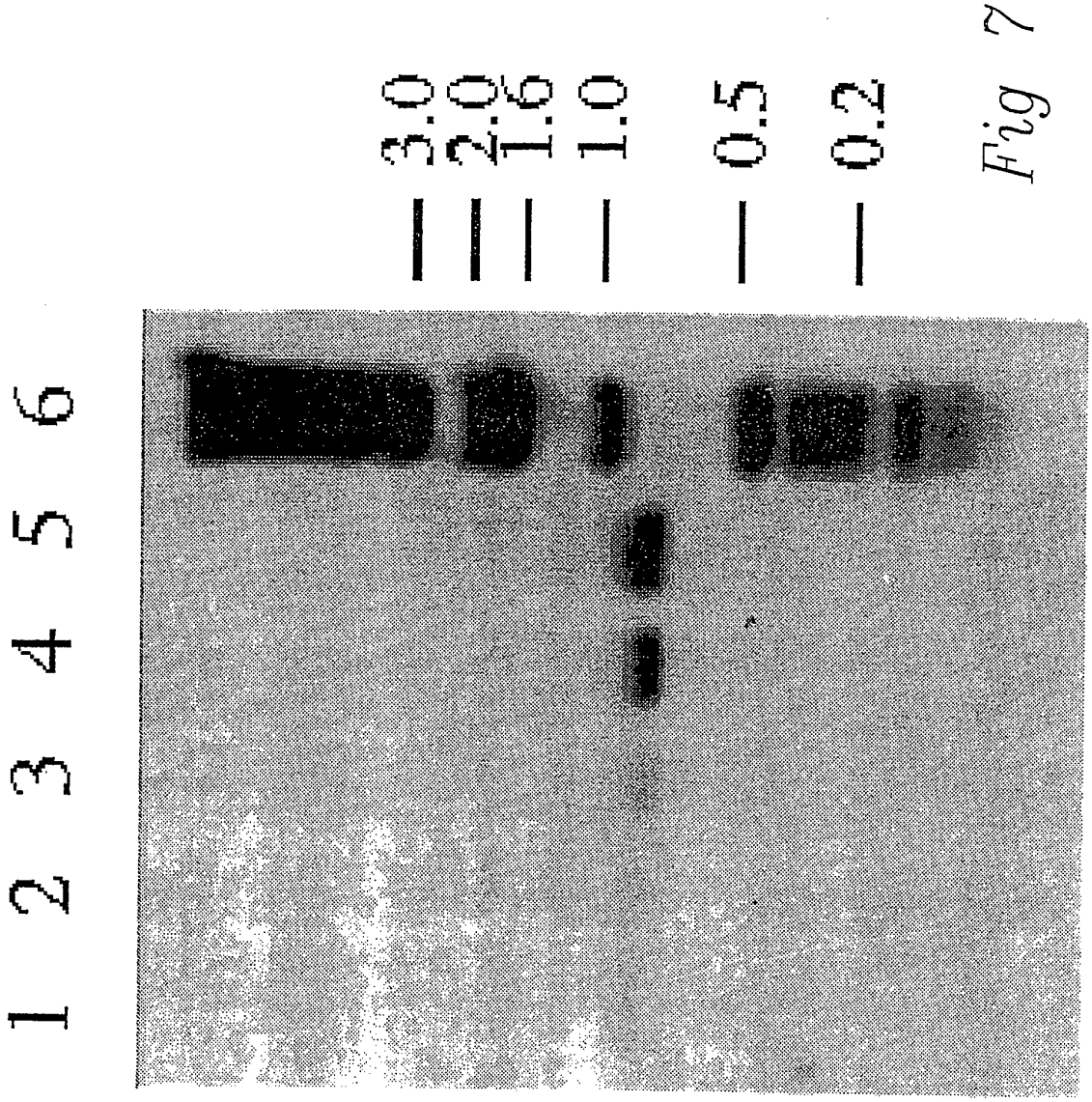
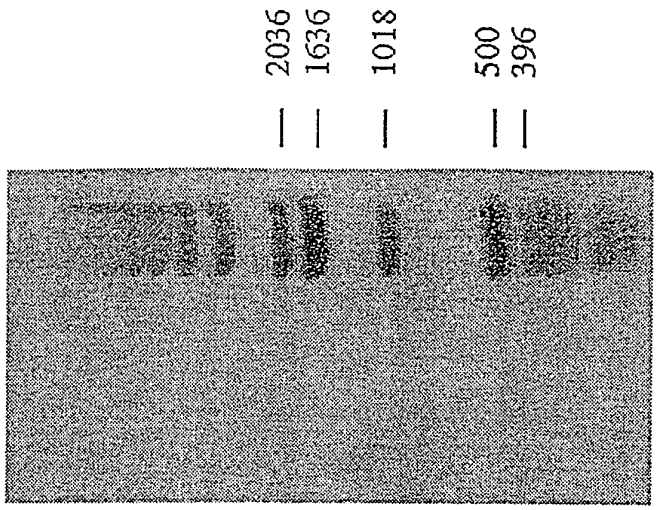
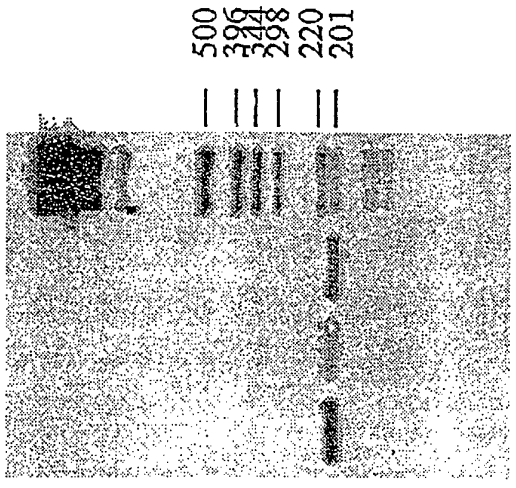


Fig 7


2036
1636
1018
500
396
α-lactalbumin
cDNA

B.


500
396
344
298
220
201

A.

Fig 8

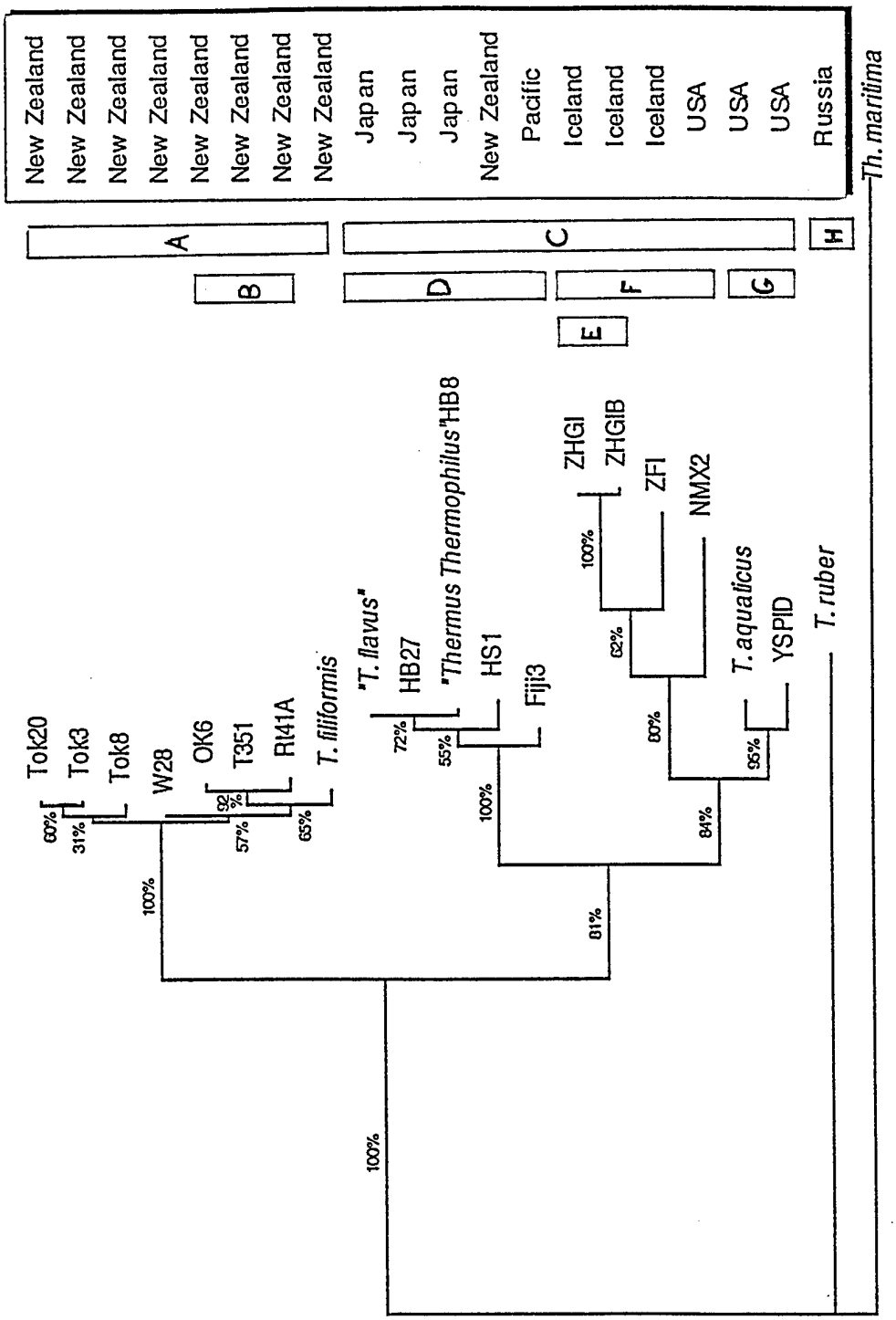

New Zealand
New Zealand
New Zealand
New Zealand
New Zealand
New Zealand
New Zealand
New Zealand
Japan
Japan
Japan
New Zealand
Pacific
Iceland
Iceland
Iceland
USA
USA
USA
Russia
Th. maritima
A
B
C
D
E
F
G
H
Tok20
Tok3
Tok8
W28
OK6
T351
Rt41A
T. filiformis
"T. flavus"
HB27
"Thermus Thermophilus"HB8
HS1
Fiji3
ZHGI
ZHGIB
ZFI
NMX2
T. aquaticus
YSPID
T. ruber
60%
31%
100%
92%
57%
65%
72%
55%
100%
81%
100%
100%
82%
80%
95%
84%
100%

Fig 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NZ 94/00135

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.[6] C12N 9/12, 15/54

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
ELECTRONIC DATABASES AS BELOW

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
AU : IPC : C12N 9/12, 15/54

Electronic data base consulted during the international search (name of data base, and where practicable, search terms used)
DERWENT : WPAT, BIOT, CA : KEYWORDS - DNA( )POLYMERASE#, DNA( )NUCLEOTIDYLTRANSFERASE#, REVERSE( )TRANSCRIPTASE#, THERMUS( )FILIFORMIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
X	EP 592035 (EASTMAN KODAK COMPANY) 13 April 1994 (13.04.94) See claim 4	1, 2
X	WO,A, 92/06202 (CETUS CORPORATION) 16 April 1992 (16.04.92) See whole document	1, 2
X	AU,B, 646430 (71764/91) CETUS CORPORATION 11 July 1991 (11.07.91) See example 1	1, 2

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 March 1995 (16.03.95)	22 Mar 1995 (22.03.95)

Name and mailing address of the ISA/AU

AUSTRALIAN INDUSTRIAL PROPERTY ORGANISATION
PO BOX 200
WODEN ACT 2606
AUSTRALIA

Facsimile No. 06 2853929

Authorized officer

JOHN ASHMAN

Telephone No. (06) 2832364

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
EP	5920355	JP	6209775	US	5338671		
WO	92062022	AU	86688/91	AU	89077/91	CA	2090614
		CA	2092317	EP	550687	EP	550696
		WO	9206200				
AU	71764/91	AU	63296/94	CA	1071196	CA	2071213
		EP	506825	EP	506889	WO	9109950
		WO	9109944	US	5310652		
		US	5322770				

END OF ANNEX