



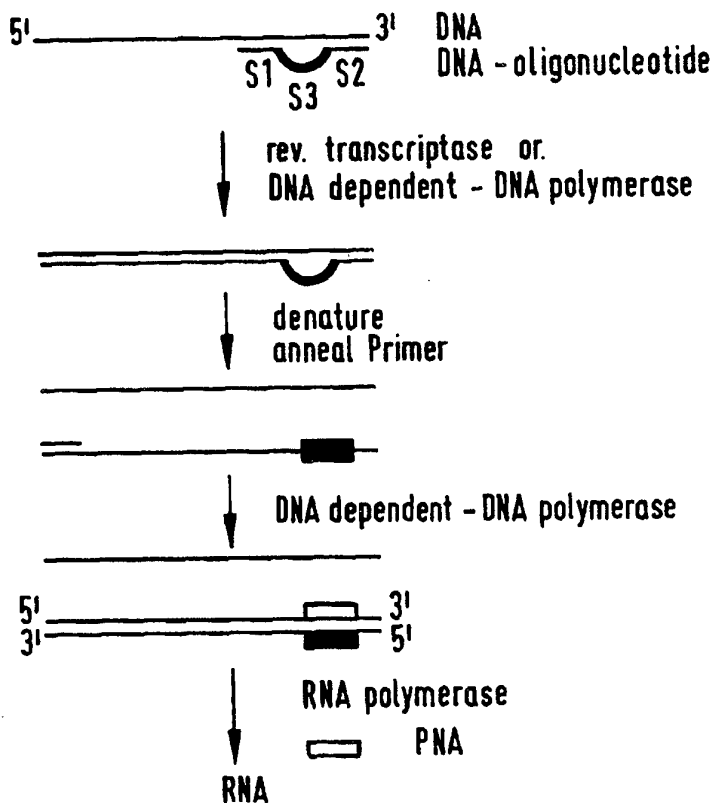
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : C12Q 1/68, C12N 15/10</p>	<p>A1</p>	<p>(11) International Publication Number: <b>WO 95/14789</b> (43) International Publication Date: 1 June 1995 (01.06.95)</p>
<p>(21) International Application Number: PCT/EP94/03858 (22) International Filing Date: 22 November 1994 (22.11.94) (30) Priority Data: 9324245.1 25 November 1993 (25.11.93) GB (71) Applicant: BUCHARDT, Dorte (heir of the deceased inventor) [DK/DK]; Sondergardsvej 73, DK-3500 Varlose (DK). (71)(72) Applicants and Inventors: BERG, Rolf, Henrik [DK/DK]; Langelandsvej 20 B 3 th., DK-2000 Fredericksberg (DK). EGHOLM, Michael [US/US]; 1231 Lexington Ridge Drive, Lexington, MA 02173 (US). NIELSEN, Peter, Eigil [DK/DK]; Hjortevænget 509, DK-2980 Kokkedal (DK). (72) Inventor: BUCHARDT, Ole (deceased). (74) Agents: KNAUER, Martin et al.; Boehringer Mannheim GmbH, D-68298 Mannheim (DE).</p>	<p>(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  Published <i>With international search report.</i></p>	

(54) Title: NUCLEIC ACID ANALOGUE INDUCED TRANSCRIPTION OF DOUBLE STRANDED DNA

(57) Abstract

RNA is transcribed from a double stranded DNA template by forming a complex by hybridising to the template at a desired transcription initiation site one or more oligonucleic acid analogues of the PNA type capable of forming a transcription initiation site with the DNA and exposing the complex to the action of a DNA dependant RNA polymerase in the presence of nucleoside triphosphates. Equal length transcripts may be obtained by placing a block to transcription downstream from the initiation site or by cutting the template at such a selected location. The initiation site is formed by displacement of one strand of the DNA locally by the PNA hybridisation.



**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	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

Nucleic acid analogue induced transcription of double stranded DNA

The present invention relates to the use of analogues of naturally occurring nucleic acids to produce sites for the in vitro initiation of transcription of double stranded DNA, to the production of RNA transcripts thereby, the amplification and/or detection of such transcripts and in vitro diagnostics techniques based on the above.

Analogues of nucleic acids having a peptide or similar backbone bearing pendant ligands such as nucleic acid bases described in WO 92/20703 (PNA's) have been shown to have a number of unusual properties. These include the ability to form complexes with double stranded DNA in which two strands of PNA complementary in sequence to one of the DNA strands hybridise to the DNA displacing the other DNA strand. A high level of sequence specificity has been shown.

Transcription of DNA to form a strand of RNA of corresponding sequence is initiated in nature by the sequence specific recognition of a promotor region of the double stranded DNA either by RNA polymerase or by auxiliary transcription factors. Subsequently, a transcription initiation open complex is formed in which about 12 base pairs of the DNA helix is melted so as to expose the bases of the template strand for base pairing with the RNA strand being synthesised. It has been shown that E. coli and phage T7 RNA polymerase can utilise synthetic "RNA/DNA bubble duplex" complexes containing an RNA/DNA duplex and a single stranded DNA D-loop for transcription initiation purposes.

We have now discovered that initiation can similarly be initiated from a strand displacement complex formed between PNA and double stranded DNA. This presents the prospect of having a ready and simple way of preparing single stranded transcripts from a double stranded template.

Generally, techniques for identifying DNA sequences depend upon having the DNA in single stranded form. Once single stranded, the DNA can be hybridised to a probe of complementary sequence and such hybridisation can be detected in various ways. Transcription of RNA from a double stranded template DNA presents an alternative form of

- 2 -

method for obtaining a single stranded product for detection and unlike processes of denaturation of the original DNA, it avoids the presence of corresponding amounts of the complementary single stranded product which can compete in the detection process.

Furthermore, the production of an RNA transcript opens the way for amplification of the original DNA sequence without the use of the polymerase chain reaction and without much of the difficulty normally associated with the 3SR amplification technique. In 3SR, a starting RNA is amplified by first hybridising to it a DNA primer constructed to include a T7 polymerase promoter sequence. The primer-DNA/template-RNA is extended in its DNA strand by reverse transcriptase, the RNA strand is digested by RNase H, and the resulting single stranded DNA transcript is made double stranded with reverse transcriptase to provide a template for transcription by T7 RNA polymerase to make large number of RNA copies. This process is dependent on the construction of a DNA primer of the correct sequence downstream from the T7 promoter sequence. It is further dependent on obtaining the nucleic acid sequence of interest in the form of RNA.

According to the present invention however, a nucleic acid sequence of interest obtained in the form of double stranded DNA can be amplified as multiple single stranded RNA copies by synthesising a primer or multiple primer sequences of the nucleic acid analogue, which will generally be much more straightforward than the prior art methods. The RNA transcripts produced can be converted to DNA if desired.

Accordingly, the present invention provides a method of transcribing RNA from a double stranded DNA template comprising forming a complex by hybridising to said template at a desired transcription initiation site one or more oligo-nucleic acid analogues capable of forming a transcription initiation site with said DNA and exposing said complex to the action of a DNA dependent RNA polymerase in the presence of nucleoside triphosphates.

Optionally, a pair of said oligo-nucleic acid analogues are hybridised to said DNA at spaced locations thereon, on the same or different strands thereof.

Preferably, said pair of oligo-nucleic acids are spaced by from 0 to 10, more preferably 0 to 5 base pairs of said DNA.

Preferably also, the or each said oligo-nucleic acid analogue has a length of from 5 to 60 nucleic acid analogue units.

- 3 -

Optionally, a block to transcription is placed at a location downstream from said desired initiation site so as to produce equal length transcripts in said transcription. A suitable way of producing a said block is by hybridising to said DNA an oligo-nucleic acid analogue capable of blocking transcription. Otherwise, individual transcription events may terminate randomly downstreams from the initiation site leading to long transcription products of varying length.

The length of the transcripts can also be controlled by cutting the DNA template with a restriction enzyme at a specific downstream location prior to transcription.

The nucleic acid analogue capable of forming a transcription initiation site is preferably a compound that has nucleobases attached to an aminoethylglycine backbone or other like backbone including polyamides, polythioamides, polysulfinamides and polysulfonamides, which compounds we call peptide nucleic acids or PNA. Compounds of this kind surprisingly bind strongly and sequence selectively to both RNA and DNA.

The synthesis of this type of compound is fully described in WO 92/20703.

The recognition by PNA of RNA, ssDNA or dsDNA can take place in sequences at least 5 bases long. A more preferred recognition sequence length is 5 - 60 base pairs long. Sequences between 10 and 20 bases are of particular interest since this is the range within which unique DNA sequences of prokaryotes and eukaryotes are found. Sequences of 17 - 18 bases are of special interest since this is the length of unique sequences in the human genome.

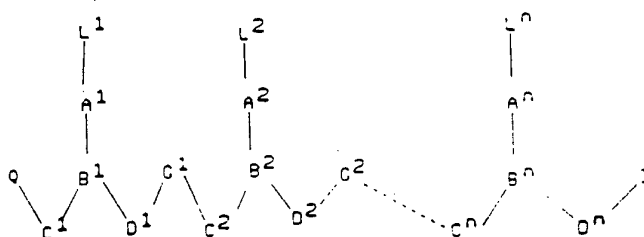
Preferably, the or a nucleic acid analogue used is capable of hybridising to a nucleic acid of complementary sequence to form a hybrid which is more stable against denaturation by heat than a hybrid between the conventional deoxyribonucleotide corresponding in sequence to said analogue and said nucleic acid.

Preferably, also the or a nucleic acid analogue used is a peptide nucleic acid in which said backbone is a polyamide backbone, each said ligand being bonded directly or indirectly to an aza nitrogen atom in said backbone, and said ligand bearing nitrogen atoms mainly being separated from one another in said backbone by from 4 to 8 intervening atoms.

- 4 -

Also, it is preferred that the or a nucleic acid analogue used is capable of hybridising to a double stranded nucleic acid in which one strand has a sequence complementary to said analogue, in such a way as to displace the other strand from said one strand.

More preferred PNA compounds for use in the invention have the formula:



Formula 1

wherein:

$n$  is at least 2,

each of  $L^1-L^n$  is independently selected from the group consisting of hydrogen, hydroxy,  $(C_1-C_4)$ alkanoyl, naturally occurring nucleobases, non-naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding groups, heterocyclic moieties, and reporter ligands;

each of  $C^1-C^n$  is  $(CR^6R^7)_y$  (preferably  $CR^6R^7$ ,  $CHR^6CHR^7$  or  $CR^6R^7CH_2$ ) where  $R^6$  is hydrogen and  $R^7$  is selected from the group consisting of the side chains of naturally occurring alpha amino acids, or  $R^6$  and  $R^7$  are independently selected from the group consisting of hydrogen,  $(C_2-C_6)$ alkyl, aryl, aralkyl, heteroaryl, hydroxy,  $(C_1-C_6)$ alkoxy,  $(C_1-C_6)$ alkylthio,  $NR^3R^4$  and  $SR^5$ , where  $R^3$  and  $R^4$  are as defined below, and  $R^5$  is hydrogen,  $(C_1-C_6)$ alkyl, hydroxy, alkoxy, or alkylthio-substituted  $(C_1-C_6)$ alkyl or  $R^6$  and  $R^7$  taken together complete an alicyclic or heterocyclic system;

each of  $D^1-D^n$  is  $(CR^6R^7)_z$  (preferably  $CR^6R^7$ ,  $CH_2CR^6R^7$ , or  $CHR^6CHR^7$ ) where  $R^6$  and  $R^7$  are as defined above;

- 5 -

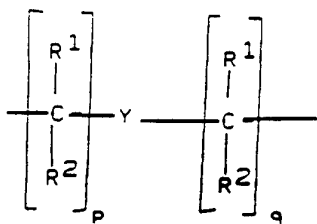
each of  $y$  and  $z$  is zero or an integer from 1 to 10, the sum  $y + z$  being at least 2, preferably greater than 2 but not more than 10, e.g. 3;

each of  $G^1-G^{n-1}$  is  $-NR^3CO-$ ,  $-NR^3C^5-$ ,  $-NR^3SO-$  or  $-NR^3SO_2-$ ,

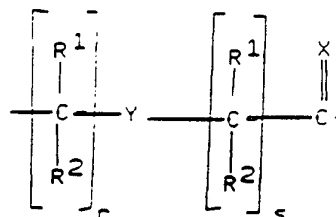
in either orientation, where  $R^3$  is as defined below;

each of  $A^1-A^n$  and  $B^1-B^n$  are selected such that:

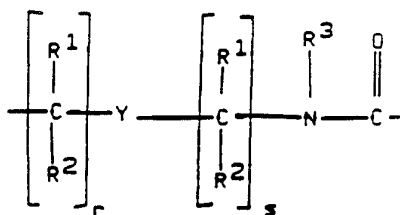
- (a) A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or  $R^3N^+$ ; or  
 (b) A is a group of formula (IId) and B is CH;



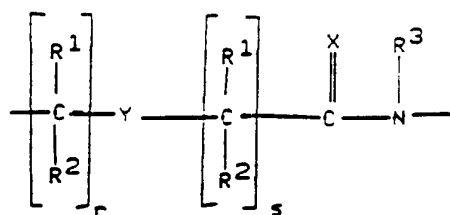
Formula 2



Formula IIb



Formula IIc



Formula IId

wherein:

X is O, S, Se,  $NR^3$ ,  $CH_2$  or  $C(CH_3)_2$ ;  
 Y is a single bond, O, S or  $NR^4$ ;

each of  $p$  and  $q$  is zero or an integer from 1 to 5, the sum  $p + q$  being not more than 10;

- 6 -

each of  $r$  and  $s$  is zero or an integer from 1 to 5, the sum  $r + s$  being not more than 10;

each  $R^1$  and  $R^2$  is independently selected from the group consisting of hydrogen,  $(C_1-C_4)$ alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; and

each  $R^3$  and  $R^4$  is independently selected from the group consisting of hydrogen,  $(C_1-C_4)$ alkyl, hydroxy- or alkoxy- or alkylthio-substituted  $(C_1-C_4)$ alkyl, hydroxy, alkoxy, alkylthio and amino;

$Q$  is  $-CO_2H$ ,  $-CONR'R''$ ,  $-SO_3H$  or  $-SO_2NR'R''$  or an activated derivative of  $-CO_2H$  or  $-SO_3H$ ; and

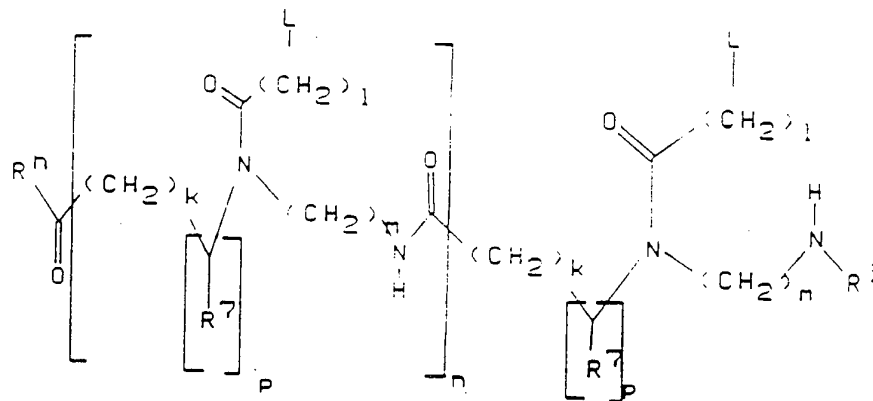
$I$  is  $NR'''R''''$  or  $-NR'''C(O)R''''$ , where  $R'$ ,  $R''$ ,  $R'''$  and  $R''''$  are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands, intercalators, chelators, peptides, proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphosphates, nucleotide triphosphates, oligonucleotides, including both oligoribonucleotides and oligodeoxyribonucleotides, oligonucleosides and soluble and non-soluble polymers. "Oligonucleosides" includes nucleobases bonded to ribose and connected via a backbone other than the normal phosphate backbone of nucleic acids.

In the above structures wherein  $R'$ ,  $R''$ ,  $R'''$  and  $R''''$  are oligonucleotides or oligonucleosides, such structures can be considered chimeric structures between PNA compounds and the oligonucleotide or oligonucleoside.

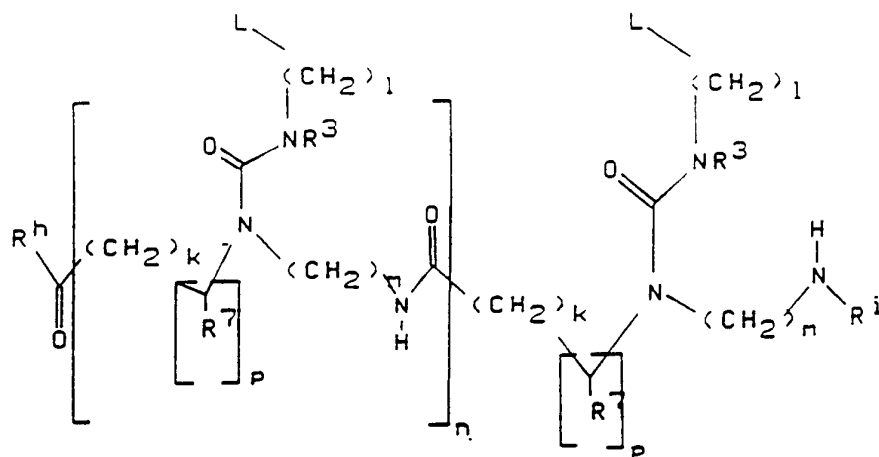
Generally, at least one of  $L^1-L^n$  will be naturally occurring nucleobase, a non-naturally occurring nucleobase, a DNA intercalator, or a nucleobase binding group.

Preferred PNA-containing compounds are compounds of the formula III, IV or V:

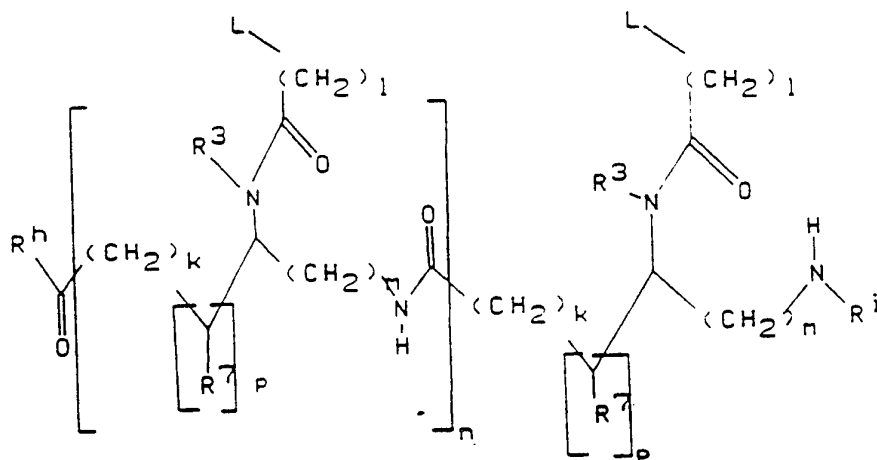




Formula III



Formula IV



Formula V

- 8 -

wherein:

each L is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties, naturally occurring nucleobases, and non-naturally occurring nucleobases;

each R<sup>7</sup> is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

n is an integer greater than 1,

each k, l, and m is, independently, zero or an integer from 1 to 5;

each p is zero or 1;

R<sup>h</sup> is OH, NH<sub>2</sub> or -NHLysNH<sub>2</sub>; and

R<sup>l</sup> is H or COCH<sub>3</sub>.

The invention includes a diagnostics method comprising carrying out a transcription to produce RNA in accordance with the methods of the invention as described above and detecting the production of said RNA. Such a method may be used to test for the presence or absence in sample DNA of a sequence matching that of one or more PNA's employed.

Suitable, said RNA is captured to a nucleic acid probe of complementary sequence and is also bound to a nucleic acid probe bearing a detectable label.

The invention includes a method of nucleic acid amplification comprising preparing an RNA transcript from DNA by a method in accordance with the above description and in such a way as to produce multiple RNA transcript copies from each molecule of DNA template.

If desired RNA transcribed according to the invention can be further amplified using the 3SR technique.

A further subject of the invention is a method of converting a (starting) nucleic acid into a substrate for transcribing RNA including forming a counterstrand of said nucleic acid incorporating a specific binding sequence, hybridizing the nucleic acid with a nucleic acid analogue, preferably a peptide nucleic acid, complementary to the specific binding sequence,

- 9 -

and treating the mixture under conditions adapted for transcription under the control of the newly created promotor side.

In a specific embodiment (described with reference to FIG 6) an RNA to be transcribed is hybridized to a DNA-oligonucleotide containing at least one nucleic acid sequence S1 hybridizable with said RNA in the region of the 3'-end of said RNA. This sequence may have a length of at least 15 nucleotides, preferably of about 20 nucleotides. Especially preferred the oligonucleotide contains an additional nucleotide sequence S2 which can hybridize to a site which is located even more at the 3'-end of the RNA. This sequence may be more than 15 nucleotides long, but will preferably be about 40 nucleotides long. In this preferred embodiment the 2 sequences of the oligonucleotide are connected to each other by a third sequence S3 which will be the specific binding sequence, adapted in sequence for specific hybridization to a nucleic acid analogue used as the promotor initiation site.

Therefore the third sequence is preferably more than 8 nucleotides long preferably about 10 nucleotides. These nucleotides preferably contain only pyrimidine bases as nucleobases. Especially preferred the third sequence consists of a homopyrimidine stretch.

The sequences of the starting RNA which correspond to the first and second sequence of the oligonucleotide are separated from each other by a sequence of about the same length as the third sequence, but not being able to hybridize to the third sequence.

For the incorporation of the specific binding sequence (third sequence) the oligonucleotide is hybridized to the nucleic acid and extended by additional mononucleotides using the starting RNA as a template. This is made by the reverse transcriptase reaction which is known in the art. This will lead to a partially double stranded nucleic acid, which contains a stretch being able to hybridize with a sequence complementary to the specific binding sequence. In a further step, this hybrid is brought into contact with the nucleic acid analogue to hybridize with said specific binding sequence. In case of a homopyrimidine stretch the nucleic acid analogue will contain the corresponding homopurine stretch.

According to the invention this construct will act as a substrate for an RNA polymerase for the production of RNA in a transcription reaction. This simple hybridization and elongation reaction provides simple access to RNAs, especially when introduced into an amplification cycle like NASBA or 3SR.

- 10 -

In another embodiment (described with reference to FIG 7) a single DNA is used to produce the above mentioned double stranded nucleic acid containing the specific stretch. In this embodiment the single stranded DNA is hybridized to a DNA-oligonucleotide as described with reference to FIG 1. Then reverse transcriptase or DNA dependent DNA polymerase is added to extend this oligonucleotide using the DNA as a matrix. After denaturation a primer complementary to the extension product is added and elongated to produce an extension product complementary to the first extension product. Then the nucleic acid analogue is added and transcription is initiated.

The invention will be further illustrated by the following examples which make reference to the appended drawings in which:

Figure 1 is an autoradiograph of a gel produced in Example 1.

Figure 2 is an autoradiograph of a gel produced in Example 2.

Figure 3 is an autoradiograph of a gel produced in Example 3.

Figure 4 is a model figure showing construction of the different PNA promoters used in the experiments of figure 5.

Figure 5 is an autoradiograph of an experiment showing competition between PNA promoters and the lacUV5 promoters.

Figure 6 and 7 show schemes for preferred embodiments of the invention.

#### Example 1

Transcription initiation (1) by single oligo-PNA, (2) by two oligo-PNA's arranged trans and (3) by two oligo-PNA's arranged cis.

Restriction fragments of three plasmids pT9C, pT9CT9C (pUC19 derivatives containing respectively the sequences T9C and T9CT9C) and pT9CA9GKS (Bluescript KS+ derivative containing a T9CA9G sequence) were isolated by digestion with PvuII and purification on polyacrylamide gels resulting in fragments of 338 base pairs (pT9C), 354 base pairs (pT9CT9C) and 477 base pairs (pT9CA9GKS). PNA-DNA complexes were formed by incubating PNA with the DNA fragments in 10 mM Tris-HCl pH 8.0 and 0.1 mM EDTA in total volume of 15  $\mu$ l for 1 hour at 37°C. The reaction mixture was adjusted to contain a final concentration of 40 mM Tris-HCl pH 7.9, 120 mM KCl, 5 mM MgCl<sub>2</sub>, 9.1 mM DTT, and 1 mM of ATP, CTP, GTP and 0.1 mM of UTP and 5  $\mu$ Ci <sup>32</sup>P UTP. The PNA used was T9C-lysNH<sub>2</sub> in each case.

- 11 -

The transcriptions were initiated by addition of 100 nM E. Coli RNA polymerase holoenzyme (Boehringer Mannheim GmbH). The mixtures (total volume of 30  $\mu$ l) were incubated at 37°C for 20 minutes and the RNA produced by transcription was subsequently recovered by ethanol precipitation. The RNA transcripts were analysed on 8 % denaturing polyacrylamide gels, and visualised by autoradiography to produce the gel shown in Figure 1.

As shown in the schematics in Figure 1, the three plasmids used provide respectively a single binding site for the PNA (mono), a pair of binding sites on the same DNA strand (cis), and a pair of binding sites on opposite strands of the DNA (trans).

The lanes of the gel show the effect of varying concentrations of PNA as follows:

Lanes 1, 6 and 11:	0 M
Lanes 2, 7 and 12:	3 nM
Lanes 3, 8 and 13	10 nM
Lanes 4, 9 and 14:	3 $\mu$ m
Lanes 5, 10 and 15:	10 $\mu$ m

The plasmids used in the lanes were as follows:

Lanes 1-5:	pT9C
Lanes 6-10:	pT9CT9C
Lanes 7-15:	pT9CA9G

Lane 5 shows the production of a single RNA product having the size expected if transcription proceeds from the PNA binding site in the direction shown in the corresponding schematic.

Lane 10 similarly shows the production of one RNA transcript but transcription is shown to be more efficiently promoted by the presence of two oligo PNA's at the binding site arranged in cis.

Lanes 13 to 15 show the production of two transcripts of the sizes expected if transcription is initiated on each of the two DNA strands and proceeds from the respective binding site to the end of the DNA fragment as illustrated in the schematic.

- 12 -

It is estimated that in those lanes where transcript RNA is seen, from 1 to 5 RNA molecules are being produced per DNA template molecule during the 20 minute incubation with RNA polymerase.

#### Example 2

##### Transcription initiation by single PNA oligomers of varying base sequence

Restriction fragments of plasmids containing the sequences T9C, T9A, and T9G were isolated. PNA-DNA complexes were formed with PNA oligomers of corresponding sequence as described in Example 1 and transcription was initiated also as described in Example 1 using *E. coli* polymerase. The resulting transcripts were visualized by autoradiography to produce the autoradiography shown in Figure 2, demonstrating that transcription is obtainable whichever of the bases A, C and G is present. Lanes 1, 3 and 5 are control lanes run without PNA present during the attempted transcription.

#### Example 3

##### Transcription initiation by single PNA oligomers using T7 and T3 polymerase

Using the restriction fragment from the plasmid pT9C and the PNA oligomer T9C described in Example 1, transcription was initiated generally as described in Example 1 but using separately T3 and T7 polymerase to produce the autoradiograph shown in Figure 3. Lanes 1 and 2 are controls run in the absence of PNA during attempted transcription with T7 (lane 1) and T3 (lane 2) and lanes 5 and 6 show the effect of the presence of PNA T9C on transcription mediated by T7 (lane 5) and T3 (lane 6).

#### Example 4

##### Transcription initiation using a second PNA oligomer hybridizing downstream from the first transcription initiation site

Experiments were undertaken to estimate the strength of the PNA-dependent transcription initiation. This was done by having both the strong lacUV5 *E. coli* promoter and one or more PNA targets on the same DNA fragment. These constructs and the results of the transcription experiments are presented in Figs. 4 and 5. With the constructs containing one

- 13 -

or two PNA T<sub>10</sub> targets on the template strand downstream from the lacUV5 promoter (Fig. 4 A and B) two new transcripts are observed as the PNA concentration is increased. One transcript is assigned to be initiated at the PNA target, while the other is assigned to be initiated from the UV5 promoter and arrested at the PNA site. A construct having two PNA T<sub>4</sub>CT<sub>5</sub> targets on the template strand and one on the non-template strand only produced transcripts assigned to be initiated at the two PNA loops and proceeding in opposite directions. The intensity of the RNA bands decreased at higher PNA concentrations, most likely as a consequence of full occupancy of all PNA sites, since an occupied PNA site downstream from the PNA loop would arrest synthesis from this loop. It is also observed that the intensity of the band corresponding to transcription from the larger loop is significantly more intensive than that of the transcription from the smaller loop, indicating a more efficient transcription from the larger loop. This difference is more pronounced than is apparent from the autoradiograph taking into account the different sizes of the transcripts.

The relative positions of hybridization are shown in Fig. 5. The sequence to the left indicates the PNA targets corresponding to the upper strand from 5' to 3' in the models. PNA binding is shown by thick bars. The direction of transcription is shown with the approximated length of the RNA product indicated.

The fragment including a single, a double, or a triple PNA binding site together with the lacUV5 promoter was incubated with the desired amount of PNA for 1 hr at 37°C. The transcription reactions were performed as described above except that 20 nM *E. coli* RNA polymerase was used.

The plasmids pT9C, pT9CT9C (pUC19 derivatives), pT9CA9GKS, pT10KS, and pA10KS (Bluescript KS + derivatives) were constructed as described (Nielsen, P.E., Egholm, M., Berg, R.H. & Buchardt, O. (1993) *Anti-Cancer Drug Des.* 8, 53-63; Nielson, P.E. Egholm, M., Berg, R.H. & Buchardt, O. (1993) *Nucleic Acids Res.* 21, 197-200. The plasmid pUV5 was constructed by cloning a 203-bp Eco RI fragment containing the lacUV5 promoter (position -150 to +93) into the Eco RI site of pUC18. pT9C-UV5 contains three PNA T<sub>4</sub>CT<sub>5</sub> targets each separated by 6bp. pA10UV5 contains a single PNA T<sub>10</sub> target cloned in the BamHI site of pUV5 (A<sub>10</sub> on the template strand), while pA10A10UV5 contains two T<sub>10</sub> PNA targets separated by a 6-bp linker cloned in the BamHI site of pUV5 (both A<sub>10</sub> on the template strand). The PNA was synthesized as described (Egholm, M., Buchardt, O., Nielsen, P.E. & Berg, R.H. (1992) *J. Am. Chem.*

- 14 -

Soc. 114, 1895-1897; Egholm, M., Buchardt, O., Nielsen, P.E. & Berg, R.H. (1992) J. Am. Chem. Soc. 114, 9677-9678).

Restriction fragments of the plasmids were isolated by digestion with Pvu II and purification on low-melting agarose gels resulting in fragments of 338 bp (pT9C), 354 bp (pT9CT9C), 461 bp (pA10KS and pT10KS), 477 bp (pT9CA9GKS), 354 bp (pA10UV5), 370 bp (pA10A10UV5), or 386 bp (pT9CUV5).

In vitro transcription from purified DNA fragments containing a single PNA T<sub>10</sub> target (pA10UV5) (A), two PNA T<sub>10</sub> targets in cis (pA10A10UV5) (B) or a triple PNA T<sub>10</sub>CT<sub>4</sub> target (one in trans and two in cis) (pT9CUV5) (C) as well as a lacUV5 promoter. The concentrations of PNA were as follows: lanes 1, 0 nM, lanes 2, 1 nM, lanes 3, 3 nM, lanes 4, 10 nM, lanes 5, 30 nM, lanes 6, 0.1 μM, lanes 7, 0.3 μM. Seventy nanomolar DNA was used in all experiments. Transcription from the UV5 and the PNA promoters are marked with arrows as is transcription arrest.

The experiments clearly show that the transcription initiated at the first initiation site (UV5) is stopped at the second site, allowing the production of RNA of defined length. In addition to that, the second site acts as an initiator for transcription starting at the second site.



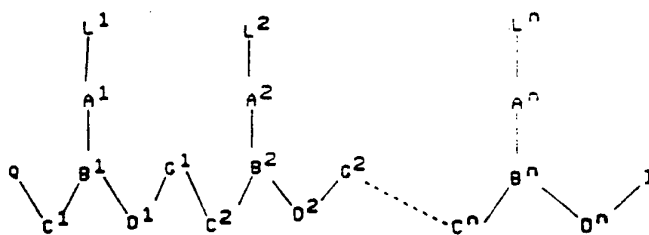
- 15 -

**Claims**

1. A method of transcribing RNA from a double stranded DNA template comprising forming a complex by hybridising to said template at a desired transcription initiation site one or more oligo-nucleic acid analogues capable of forming a transcription initiation site with said DNA and exposing said complex to the action of a DNA dependent RNA polymerase in the presence of nucleoside triphosphates.
2. A method as claimed in claim 1, wherein a pair of said oligo-nucleic acid analogues are hybridised to said DNA at adjacent or spaced locations thereon, on the same or different strands thereof.
3. A method as claimed in claim 2, wherein said pair of oligo-nucleic acids are spaced by from 0 to 10 base pairs of said DNA.
4. A method as claimed in any preceding claims, wherein the or each said oligo-nucleic acid analogue has a length of from 10 to 20 nucleic acid analogue units.
5. A method of converting a single stranded nucleic acid into a substrate for the transcription by an RNA polymerase containing the steps
  - enzymatic conversion of a nucleic acid into a double stranded DNA containing a specific binding sequence,
  - hybridizing said hybrid with a nucleic acid analogue containing a nucleobase sequence complementary to the specific binding sequence and
  - submitting the hybrid to the action of a RNA polymerase.
6. A method as claimed in any preceding claim, wherein a block to transcription is placed at a location downstream from said desired initiation site so as to produce equal length transcripts in said transcription or wherein transcript length is controlled by cutting the template at a selected location downstream from said initiation site.
7. A method as claimed in claim 5, wherein said block is produced by hybridising to said DNA an oligo-nucleic acid analogue capable of blocking transcription.

- 16 -

8. A method as claimed in any preceding claim, wherein the nucleic acid analogue has an amide, thioamide, sulphinamide, or sulphonamide backbone.
9. A method as claimed in any one of claims 1 to 7, wherein the or a said nucleic acid analogue used is a peptide nucleic acid in which said backbone is a polyamide backbone, each said ligand being bonded directly or indirectly to an aza nitrogen atom in said backbone, and said ligand bearing nitrogen atoms mainly being separated from one another in said backbone by from 4 to 8 intervening atoms.
10. A method as claimed in any preceding claim, wherein the or a nucleic acid analogue used is capable of hybridising to a nucleic acid of complementary sequence to form a hybrid which is more stable against denaturation by heat than a hybrid between the conventional deoxyribonucleotide corresponding in sequence to said analogue and said nucleic acid.
11. A method as claimed in any preceding claim, wherein the or a nucleic acid analogue used is able of hybridising to a double stranded nucleic acid in which one strand has a sequence complementary to said analogue, in such a way as to displace the other strand from said one strand.
12. A method as claimed in claim 11, wherein said nucleic acid analogue comprises a compound of the general formula 1:



Formula 1

wherein:

n is at least 2,

- 17 -

each of  $L^1-L^n$  is independently selected from the group consisting of hydrogen, hydroxy,  $(C_1-C_4)$ alkanoyl, naturally occurring nucleobases, non naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding groups, heterocyclic moieties, and reporter ligands;

each of  $C^1-C^n$  is  $(CR^6R^7)_y$  where  $R^6$  is hydrogen and  $R^7$  is selected from the group consisting of the side chains of naturally occurring alpha amino acids, or  $R^6$  and  $R^7$  are independently selected from the group consisting of hydrogen,  $(C_2-C_6)$ alkyl, aryl, aralkyl, heteroaryl, hydroxy,  $(C_1-C_6)$ alkoxy,  $(C_1-C_6)$ alkylthio,  $NR^3R^4$  and  $SR^5$ , where  $R^3$  and  $R^4$  are as defined below, and  $R^5$  is hydrogen,  $(C_1-C_6)$ alkyl, hydroxy, alkoxy, or alkylthio-substituted  $(C_1-C_6)$ alkyl or  $R^6$  and  $R^7$  taken together complete an alicyclic or heterocyclic system;

each of  $D^1-D^n$  is  $(CR^6R^7)_z$  where  $R^6$  and  $R^7$  are as defined above;

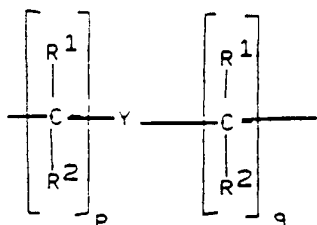
each of  $y$  and  $z$  is zero or an integer from 1 to 10, the sum  $y + z$  being from 2 to 10;

each of  $G^1-G^{n-1}$  is  $-NR^3CO-$ ,  $-NR^3C^5-$ ,  $-NR^3SO-$  or  $-NR^3SO_2-$ , in either orientation, where  $R^3$  is as defined below;

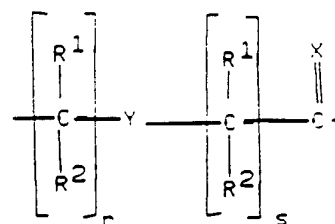
each of  $A^1-A^n$  and  $B^2-B^n$  are selected such that:

- (a) A is a group of formula (IIa), (IIb), (IIc) or (IIId), and B is N or  $R^3N^+$ ; or
- (b) A is a group of formula (IIId) and B is CH;

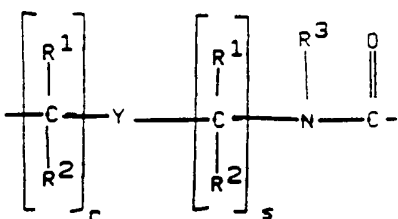
- 18 -



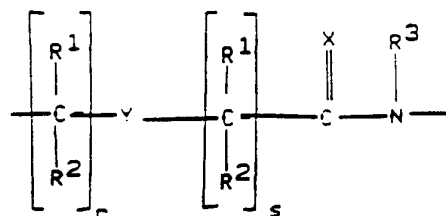
Formula 10



Formula IIb



Formula IIc



Formula II d

wherein:

X is O, S, Se, NR<sup>3</sup>, CH<sub>2</sub> or C(CH<sub>3</sub>)<sub>2</sub>;Y is a single bond, O, S or NR<sup>4</sup>;

each of p and q is zero or an integer from 1 to 5, the sum p + q being not more than 10;

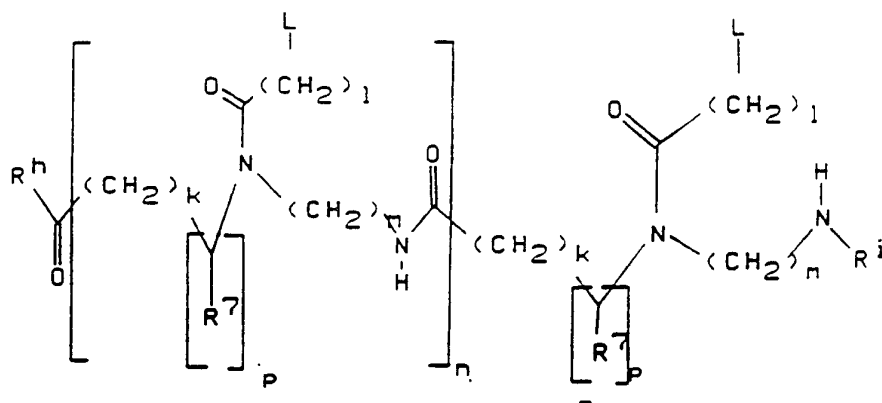
each of r and s is zero or an integer from 1 to 5, the sum r + s being not more than 10;

each R<sup>1</sup> and R<sup>2</sup> is independently selected from the group consisting of hydrogen, (C<sub>1</sub>-C<sub>4</sub>)alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; andeach R<sup>3</sup> and R<sup>4</sup> is independently selected from the group consisting of hydrogen, (C<sub>1</sub>-C<sub>4</sub>)alkyl, hydroxy- or alkoxy- or alkylthio-substituted (C<sub>1</sub>-C<sub>4</sub>)alkyl, hydroxy, alkoxy, alkylthio and amino;

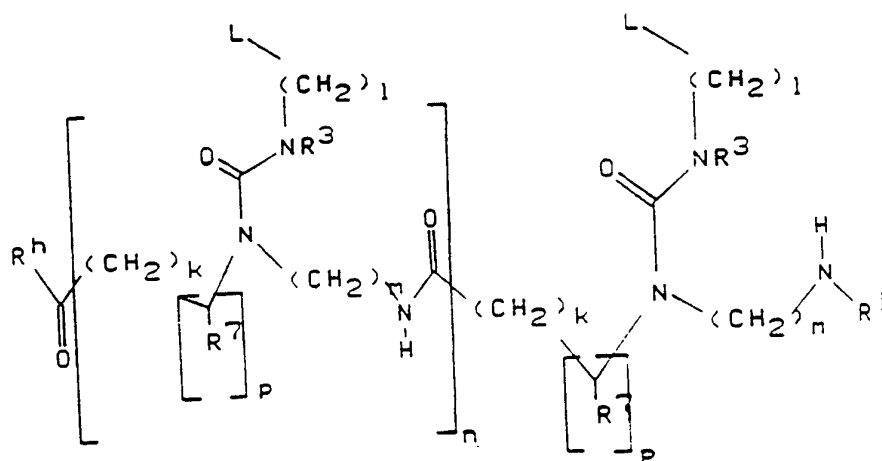
Q is  $-\text{CO}_2\text{H}$ ,  $-\text{CONR}'\text{R}''$ ,  $-\text{SO}_3\text{H}$  or  $-\text{SO}_2\text{NR}'\text{R}''$  or an activated derivative of  $-\text{CO}_2\text{H}$  or  $-\text{SO}_3\text{H}$ ; and

I is  $-\text{NR}'''\text{R}''''$  or  $-\text{NR}'''\text{C}(\text{O})\text{R}''''$ , where  $\text{R}'$ ,  $\text{R}''$ ,  $\text{R}'''$  and  $\text{R}''''$  are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands, intercalators, chelators, peptides, proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphosphates, nucleotide triphosphates, oligonucleotides, oligonucleosides and soluble and non-soluble polymers.

13. A method as claimed in claim 12, wherein said nucleic acid analogue comprises a compound of the general formula III, IV or V:

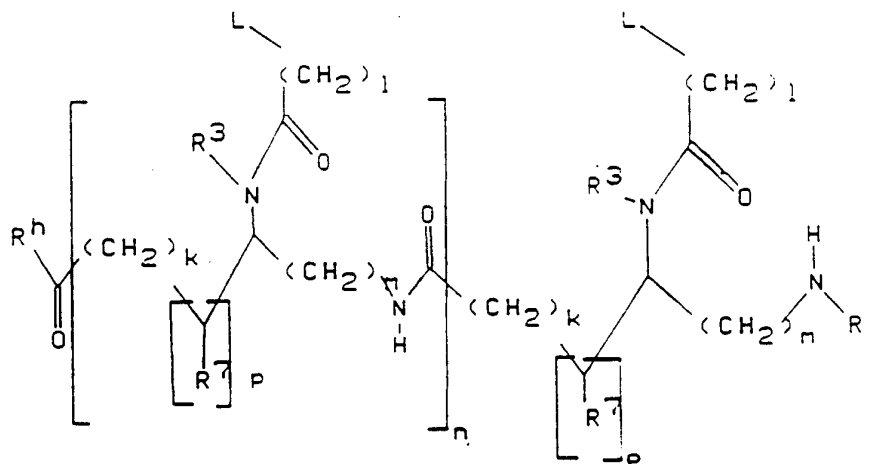


Formula III



Formula IV

- 20 -



Formula V

wherein:

each L is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties, naturally occurring nucleobases, and non-naturally occurring nucleobases; each  $R^3$  is as hereinbefore defined;

each  $R^7$  is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

n is an integer greater than 1,

each k, 1, and m is, independently, zero or an integer from 1 to 5;

each p is zero or 1;

$R^h$  is OH,  $NH_2$  or  $-NHLysNH_2$ ; and

$R'$  is H or  $COCH_3$ .

14. A diagnostics method as claimed in any preceding claim, comprising carrying out a transcription to produce RNA in accordance with any one of the preceding claims and detecting the production of said RNA.
15. A diagnostics method as claimed in claim 11, wherein said RNA is captured to a nucleic acid probe of complementary sequence and is also bound to a nucleic acid probe bearing a detectable label.

- 21 -

16. A method of nucleic acid amplification comprising preparing an RNA transcript from DNA by a method in accordance with any one of claims 1 to 13, in such a way as to produce multiple RNA copies from each DNA template molecule.

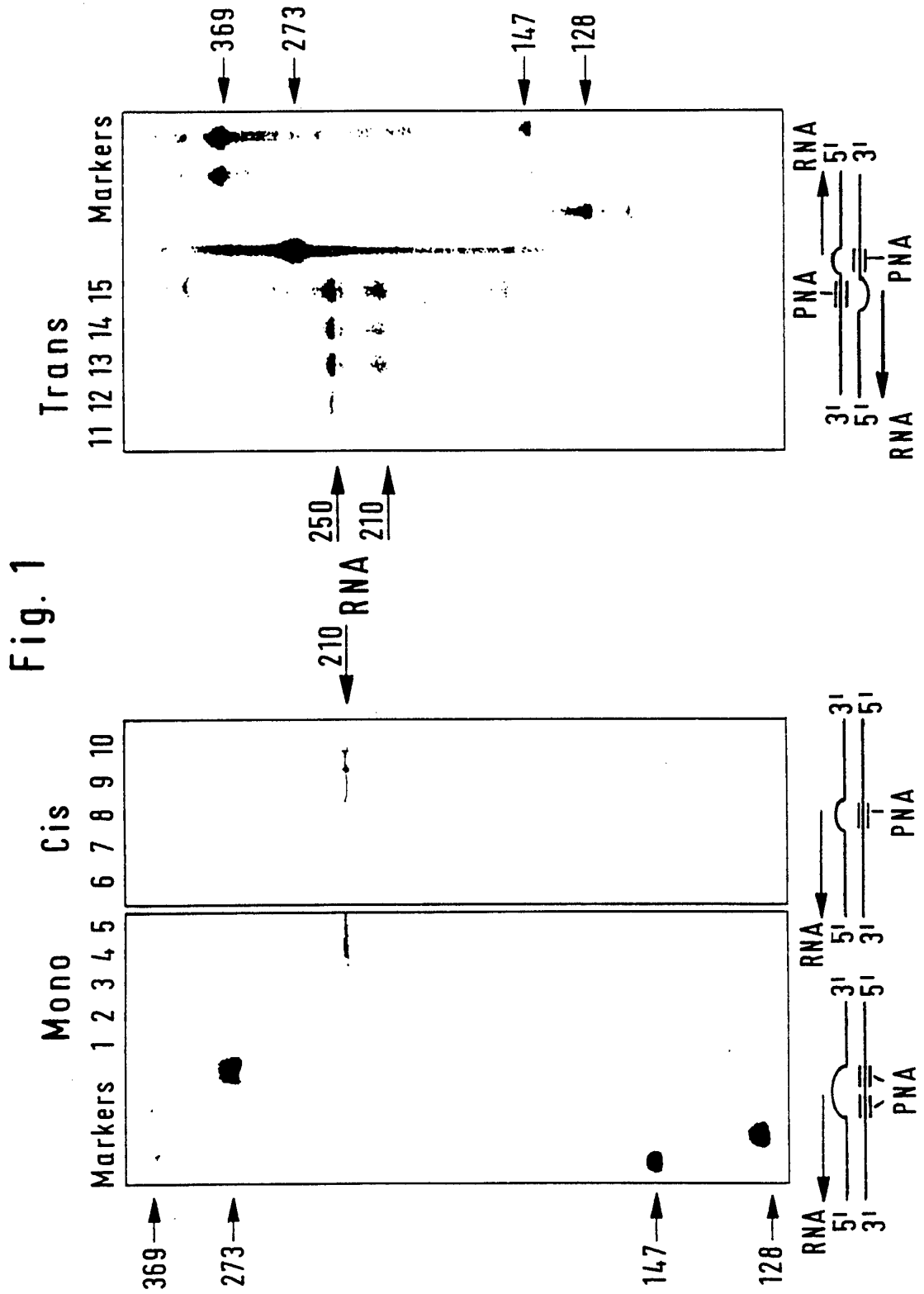
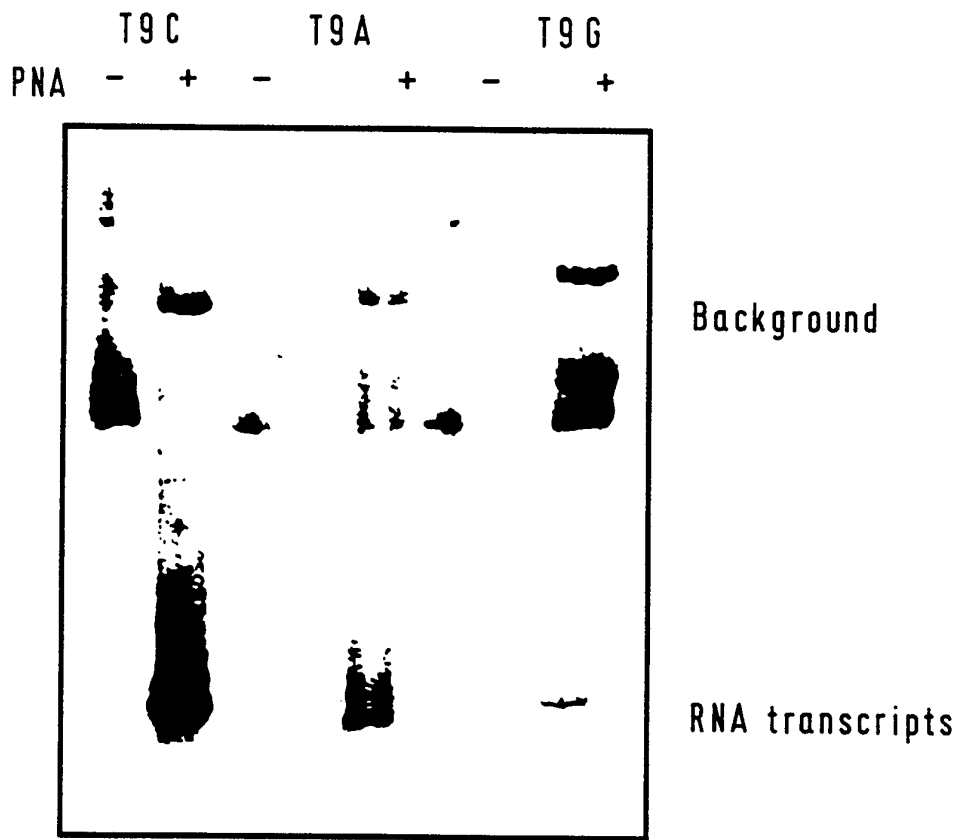


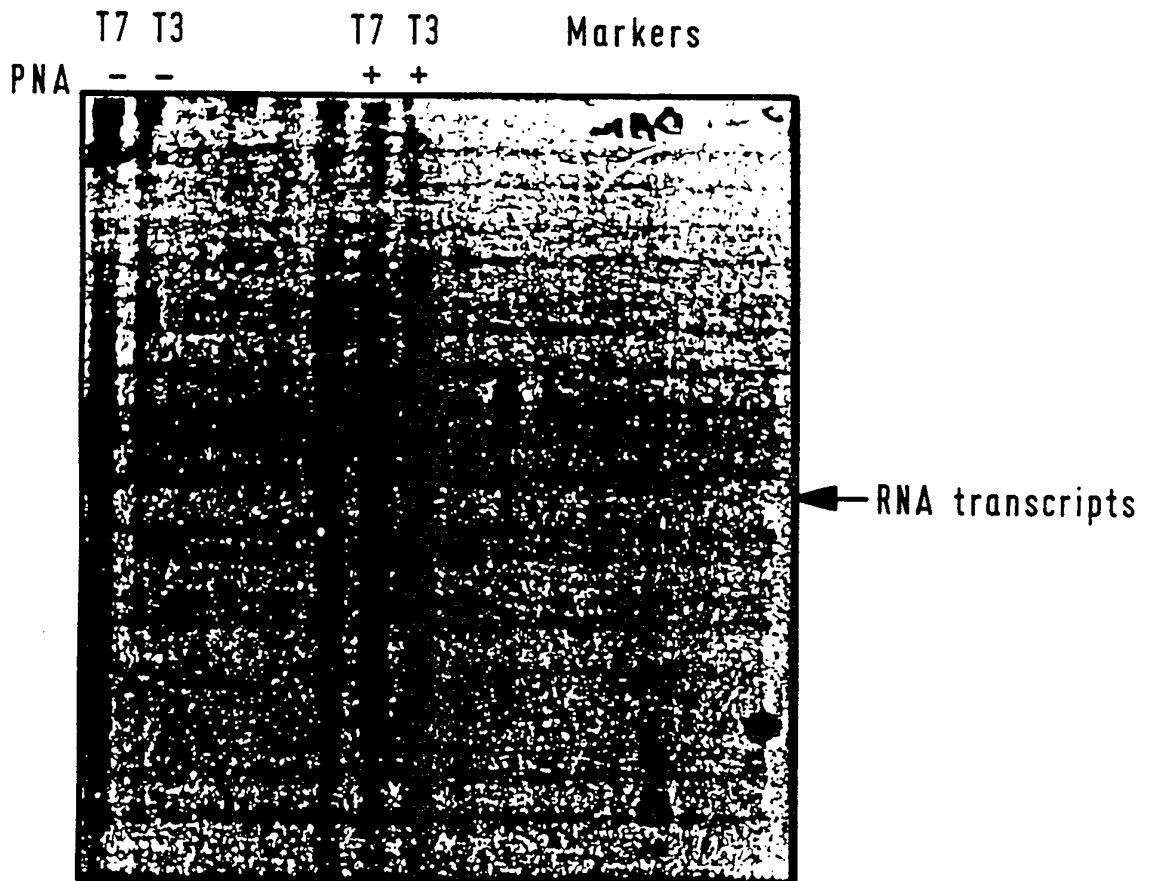


Fig. 2



E. coli RNA polymerase Run-off transcripts from PNA loops containing T, G, C and A.

Fig. 3



T3 and T7 RNA polymerase transcriptions  
from a T9C PNA loop

Fig. 4

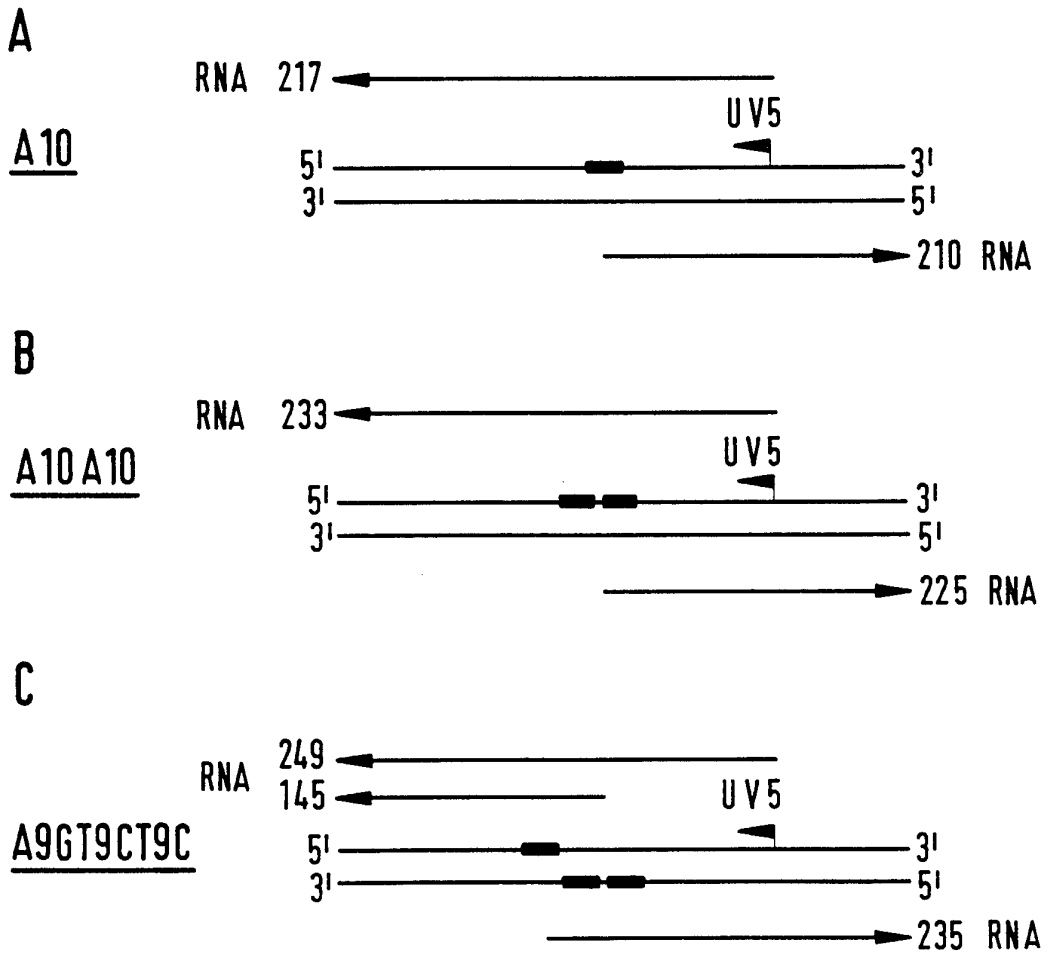


Fig. 5

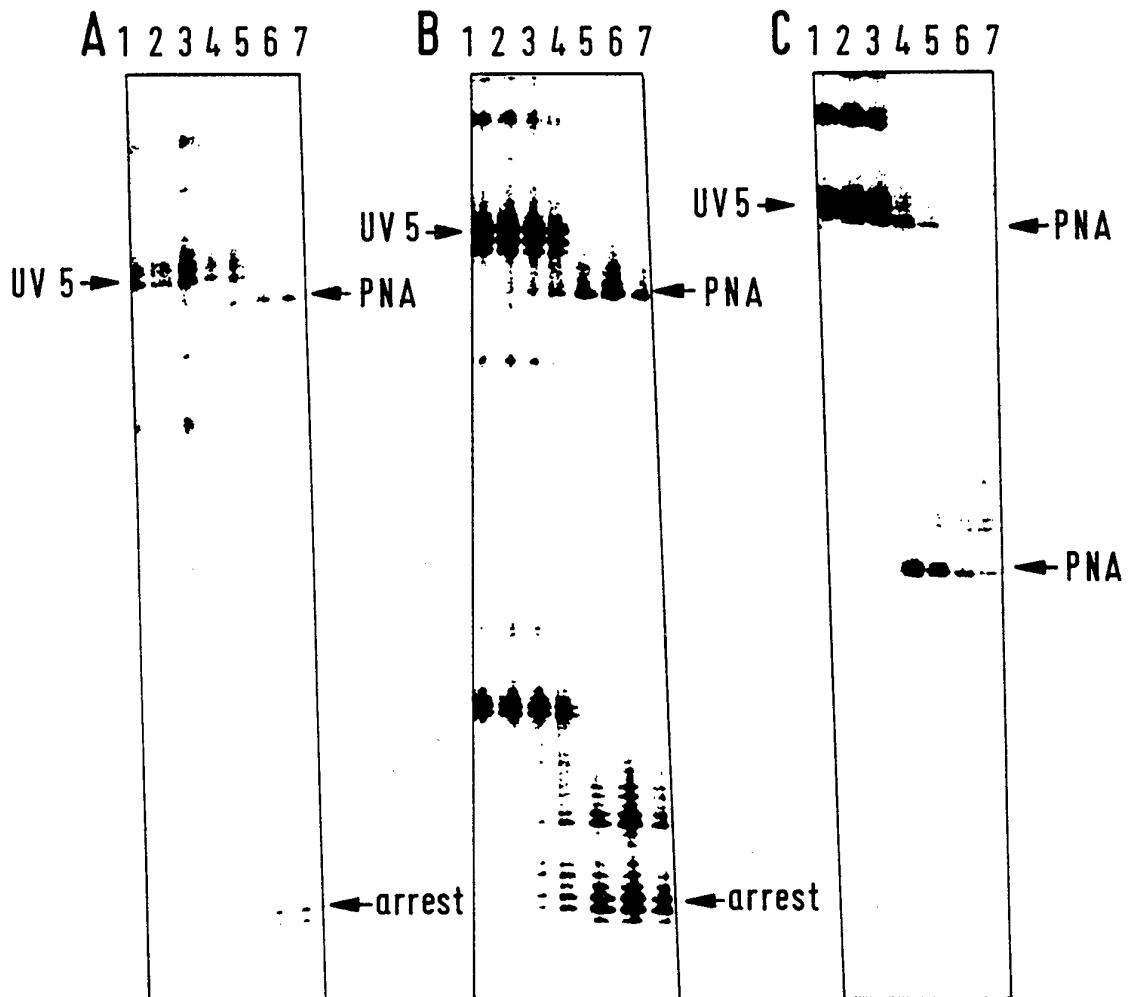


Fig. 6

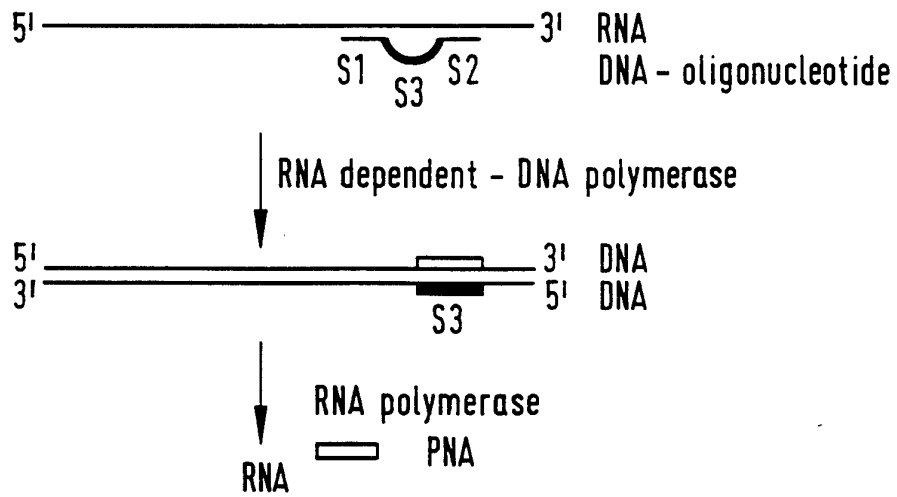
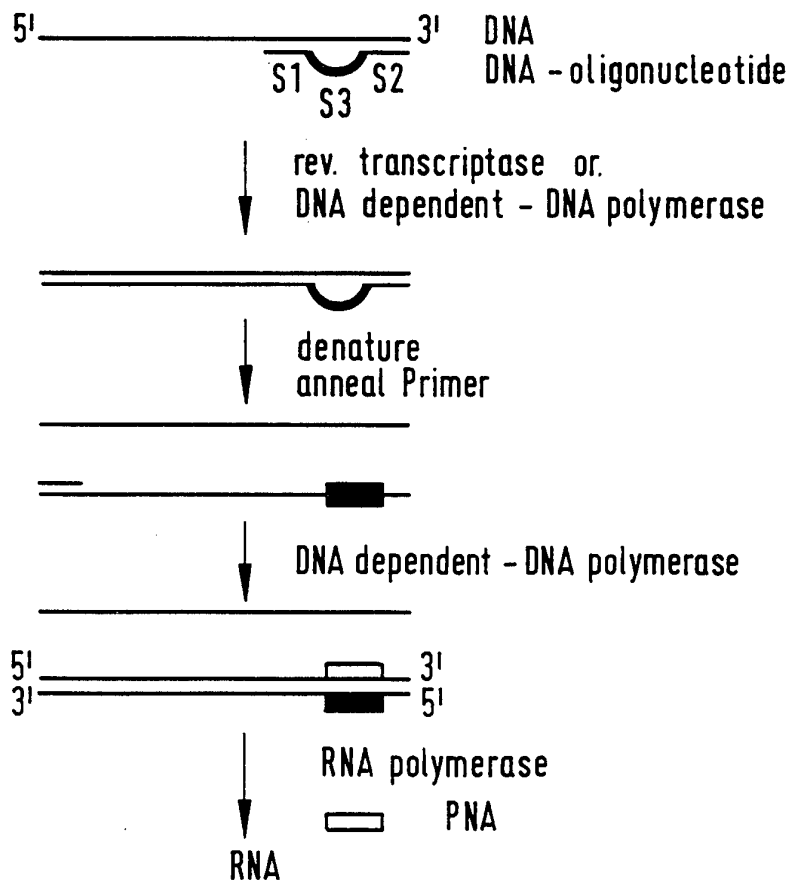


Fig. 7



**INTERNATIONAL SEARCH REPORT**

International Application No

PCT/EP 94/03858

**A. CLASSIFICATION OF SUBJECT MATTER**  
 IPC 6 C12Q1/68 C12N15/10

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)  
 IPC 6 C12Q C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category <sup>o</sup>	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SCIENCE, vol.258, 20 November 1992, AAAS, WASHINGTON, DC, US; pages 1320 - 2324 S.S. DAUBE AND P.H. VON HIPPEL 'Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes' see page 1320, left column, line 1 - page 1324, left column, line 14; figures 1-5 ---	1,4,5, 8-12,16
Y	EP,A,0 439 330 (IMPERIAL CHEMICAL INDUSTRIES PLC) 31 July 1991 see page 5, line 37 - line 45; claims 1-18; figures 2,3 --- -/--	1,4,5, 8-12,16

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

<sup>o</sup> 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

2 March 1995

Date of mailing of the international search report

08. 03. 95

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+ 31-70) 340-3016

Authorized officer

Hornig, H

## INTERNATIONAL SEARCH REPORT

Intern al Application No  
PCT/EP 94/03858

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>TRENDS IN BIOTECHNOLOGY, vol.11, no.9, September 1993, ELSEVIER SCIENCE PUBLISHERS,LTD.,CAMBRIDGE,UK; pages 384 - 386 O. BUCHARDT ET AL. 'Peptide nucleic acids and their potential applications in biotechnology' see page 385, left column, line 1 - page 386, left column, line 5; figures 2,3 ---</p>	1,4,5, 8-12,16
Y	<p>BIOCHEMISTRY, vol.28, no.19, 19 September 1989, AM. CHEM. SOC.,WASHINGTON,DC,US; pages 7829 - 7842 B. KRUMMEL AND M.J. CHAMBERLIN 'RNA chain initiation by Escherichia coli RNA polymerase. Structural transition of the enzyme in early ternary complexes' see page 7839, right column, line 50 - page 7840, left column, line 10; figure 11 see page 7841, left column, line 8 - line 20 ---</p>	1,4,5, 8-12,16
Y	<p>NUCL. ACIDS RES., vol.21, no.9, 11 May 1993, IRL PRESS, OXFORD, ENGLAND; pages 2103 - 2107 V. DEMIDOV ET AL. 'Sequence selective double strand DNA cleavage by peptide nucleic acid (PNA) targeting using nuclease S1' see page 104, right column, line 8 - page 2107, left column, line 9; figure 2 ---</p>	1,4,5, 8-12,16
A	<p>WO,A,92 20703 (BUCHARDT, EGHOLM, NIELSEN, BERG) 26 November 1992 cited in the application the whole document ---</p>	1-16
P,A	<p>WO,A,93 25706 (BUCHARDT, EGHOLM, NIELSEN, BERG) 23 December 1993 the whole document ---</p>	1-16
P,X	<p>GENE, vol.149, no.1, 4 November 1994, ELSEVIER PUBLISHERS, N.Y., U.S.; pages 139 - 145 P.E. NIELSEN ET AL. 'Sequence-specific transcription arrest by peptide nucleic acid bound to the DNA template strand' see page 143, right column, paragraph 2 ---</p>	1,4,5,8
	<p>---</p> <p style="text-align: center;">-/--</p>	

1

## INTERNATIONAL SEARCH REPORT

Intern al Application No

PCT/EP 94/03858

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	PROC. NATL. ACAD SCI., vol.91, no.9, 26 April 1994, NATL. ACAD SCI., WASHINGTON, DC, US; pages 3892 - 3895 N.E. MOLLEGAARD ET AL. 'Peptide nucleic acid x DNA strand displacement loops as artificial transcription promoters' see page 3893, left column, line 1 - page 3895, right column, line 11 -----	1-13



**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

Intern. Application No  
**PCT/EP 94/03858**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0439330	31-07-91	AU-B- 652548	01-09-94
		AU-A- 6938491	01-08-91
		GB-A, B 2240339	31-07-91
		JP-A- 6133800	17-05-94
-----			
WO-A-9220703	26-11-92	AU-A- 1880692	30-12-92
		AU-A- 1884392	30-12-92
		BR-A- 9206049	27-12-94
		CA-A- 2109320	25-11-92
		WO-A- 9220702	26-11-92
		EP-A- 0586474	16-03-94
		EP-A- 0586618	16-03-94
		HU-A- 66597	28-12-94
		JP-T- 6509063	13-10-94
		JP-T- 6506945	04-08-94
		NO-A- 934122	11-01-94
NO-A- 934235	20-01-94		
-----			
WO-A-9325706	23-12-93	AU-B- 4323593	04-01-94
		CA-A- 2136831	23-12-93
-----			