NUCLEOTIDE SEQUENCE ADJACENT TO POLYADENYLIC ACID IN GLOBIN MESSENGER RNA

N.J. PROUDFOOT and G.G. BROWNLEE

Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Received 12 October 1973

1. Introduction

The sequence adjacent to poly A in messenger RNA is of special interest. Poly A has been shown to be 3′-terminal in mRNA and HnRNA* [1] and is probably added by a post-transcriptional process [2]. The sequence adjacent to it may therefore represent a termination signal for transcription and/or a recognition site for the addition of poly A.

In this study we have used rabbit globin mRNA as a model compound to investigate this sequence. Short ^{32}P-labelled transcripts were obtained using the reverse transcriptase activity of *E. coli* DNA polymerase I in the presence of Mn^{2+} and oligo–T primer. A unique sequence (5′) A–U–U–G–C–poly A (3′) was thus deduced for the globin mRNA.

2. Methods

Rabbit globin mRNA (0.3 mg/ml) was preincubated with primer $(pT)_{12}OH$, (0.026 mM, P.L. Biochemicals, Inc., Milwaukee, Wisconsin, U.S.A.) in NaCl (0.1 M) at 37°C for 15 min. The reaction mixture comprised mRNA–$(pT)_{12}OH$ complex (30 μg/ml with respect to mRNA), three non-radioactive deoxyribonucleoside triphosphates (0.1 mM), one radioactive [α-^{32}P] deoxyribonucleoside triphosphate (specific activity 50 mCi/μmole, using 0.2 μCi per experiment), $MnCl_2$ (0.5 mM), mercaptoethanol (0.13 mM), Tris–HCl pH 7.5 (66 mM) and DNA polymerase I(0.3–3 units, Boehringher Mannheim GmbH, Germany). Incubations were carried out in a total volume of 10 μl for 25 min at 37°C and the reactions stopped by addition of EDTA to 100 mM with cooling. Incorporation was measured by precipitation of the reaction mixture with excess trichloroacetic acid (5%), tetrapotassium pyrophosphate (0.5%) in the presence of 2 mg of bovine serum albumin.

For sequence studies, the reaction mixture containing EDTA was applied directly onto the first dimension of a standard two-dimensional fractionation system [3]. Nearest neighbour analysis of the oligonucleotides was carried out by combined digestion with micrococcal nuclease and spleen phosphodiesterase (Worthington Biochemicals Corp., Freehold, New Jersey, USA [4].

3. Results

3.1. *Preliminary studies on reverse transcriptase activity of DNA polymerase*

The enzyme reaction was initially characterised by measuring the incorporation of radioactivity from α-^{32}P-labelled nucleoside triphosphates into acid precipitable counts under various conditions. The globin messenger RNA was prepared and assayed in duck reticulocyte lysates by Dr. J. Lingrel. The mRNA was active in the synthesis of both α and β globin, though it was probably contaminated with up to 30% of 9 S ribosomal RNA fragments.

Table 1 shows that the incorporation was dependent on the presence of the enzyme, the mRNA, the primer, all four deoxyribonucleoside triphosphates and Mn^{2+}. The sensitivity of the reaction to pancreatic ribonuclease (used under conditions known to

* Heterogeneous nuclear RNA.

Table 1
Characterisation of transcription.

Experiment	Conditions	Incorporation (cpm)
1	Control	80 000
2	Minus enzyme	80
3	Minus mRNA	150
4	Minus $(pT)_{12}OH$	1500
5	Minus dGTP	450
6	Mg^{2+} replacing Mn^{2+} *	1540
7	Control	8200
8	Plus RNAase**	140

Experiments 1–6 were carried out with [α-^{32}P] dCTP. Experiments 7–8 were carried out with [α-^{32}P] dTTP.
* The reaction mixture was modified to contain 8 mM Mg^{2+}, 60 mM KCl, 50 mM Tris–HCl pH 7.5 and 0.1 mM EDTA.
** mRNA = primer complex was preincubated in 0.2 mg/ml pancreatic RNAase for 5 min at 37°C.

degrade RNA to small oligonucleotides) demonstrates that the mRNA rather than a possible DNA contaminant was being copied. The 30% ribosomal RNA contaminants cannot be copied as they are devoid of poly A. The requirement for primer strongly suggests that synthesis is initiated on $(pT)_{12}OH$ hybridised with the poly A sequence known to be present in both α and β messengers [5].

3.2. *Sequence studies*

The reverse transcriptase activity of DNA polymerase described above may be used to study the sequence of bases immediately adjacent to the $(pT)_{12}OH$ primer. This was achieved by omitting one or more of the deoxyribonucleoside triphosphates from the reaction, thus restricting the elongation of the primer to a few bases [4]. The initial approach was to incubate mRNA in the presence of one labelled triphosphate. The labelled product was then purified by two-dimensional fractionation and characterised by its position on the fingerprint and by its nearest neighbour analysis [6]. Theoretically, in order to obtain a single product on the fingerprint, it is necessary to omit TTP from the reaction in all experiments. Thus only $(pT)_{12}OH$ primers which are hybridised with the twelve A residues on the 5′-end of the poly A sequence are extended.

The first residue to be added to the primer was determined by performing incubations with either [α-^{32}P] dCTP, [α-^{32}P] dGTP or [α-^{32}P] dATP in turn as the only triphosphate present. Of these only [α-^{32}P] dGTP gave labelled products on two-dimensional fractionation (fig. 1A). Surprisingly, four products (A1–A4) were present, the largest (A1) predominating. On analysis (table 2) all four spots gave Tp as their nearest neighbour, thus showing that all four products have the dinucleotide sequence $T\overset{*}{p}G_{OH}$ † at their 3′-end. The main product (A1) moves in a position on the fingerprint compatible with the primer extended by a single G, viz $(pT)_{12}\overset{*}{p}G_{OH}$. The smaller products which, from their analysis, are identical at their 3′-end and can therefore only differ from A1 at their 5′-end, probably derive both from the exonuclease activity of DNA polymerase degrading $(pT)_{12}\overset{*}{p}G_{OH}$ from its 5′-end [7] and from elongation of impurities in the primer preparation. Indeed, 5′-terminal labelling of the purchased $(pT)_{12}OH$ primer with [γ-^{32}P] ATP and T_4-induced phosphokinase (a gift from Dr. T.M. Harrison) followed by fractionation using homochromatography [3] revealed the presence of significant amounts of OH $T(pT)_{12}OH$ and smaller amounts of OH $T(pT)_{11}OH$ and $(pT)_{13}OH$.

The second residue to be added to the primer was determined by incubation with [α-^{32}P] dGTP plus either unlabelled dATP or dCTP. The former incubation gave the same products as in fig. 1A while the latter gave the result shown in fig. 1B. Comparison of figs. 1A and 1B clearly shows that two new products (B1 and B2) were obtained. Nearest neighbour analysis (table 2) again shows only $T\overset{*}{p}G_{OH}$ sequences in B1–B6. B1 and B2 can therefore represent only C additions with no possibility of a C addition followed by a G. The sequence $T\overset{*}{p}GpC\overset{*}{p}G_{OH}$ is therefore discounted. To establish the nature of B1 and B2 a further experiment with [α-^{32}P] dCTP and unlabelled dGTP was carried out, as shown in fig. 1C. Nearest neighbour analyses of C1–C5 (table 2) show that only $G\overset{*}{p}C_{OH}$ sequences were present. so that C2, the strongest component, is likely to be $(pT)_{12}pGpC_{OH}$. We assume that this product (C2) corresponds to B2 in fig. 1B and that B1 and C1 are present because of the extension of the $(pT)_{13}OH$ contaminant in the primer (see above). The nearest neighbour information (table 2) rules out the possibility that B1 and C1 have two C residues, i.e. $(pT)_{12}pG\overset{*}{p}C\overset{*}{p}C_{OH}$. The smaller

† The asterisk indicates the labelled [^{32}P] phosphate residue.

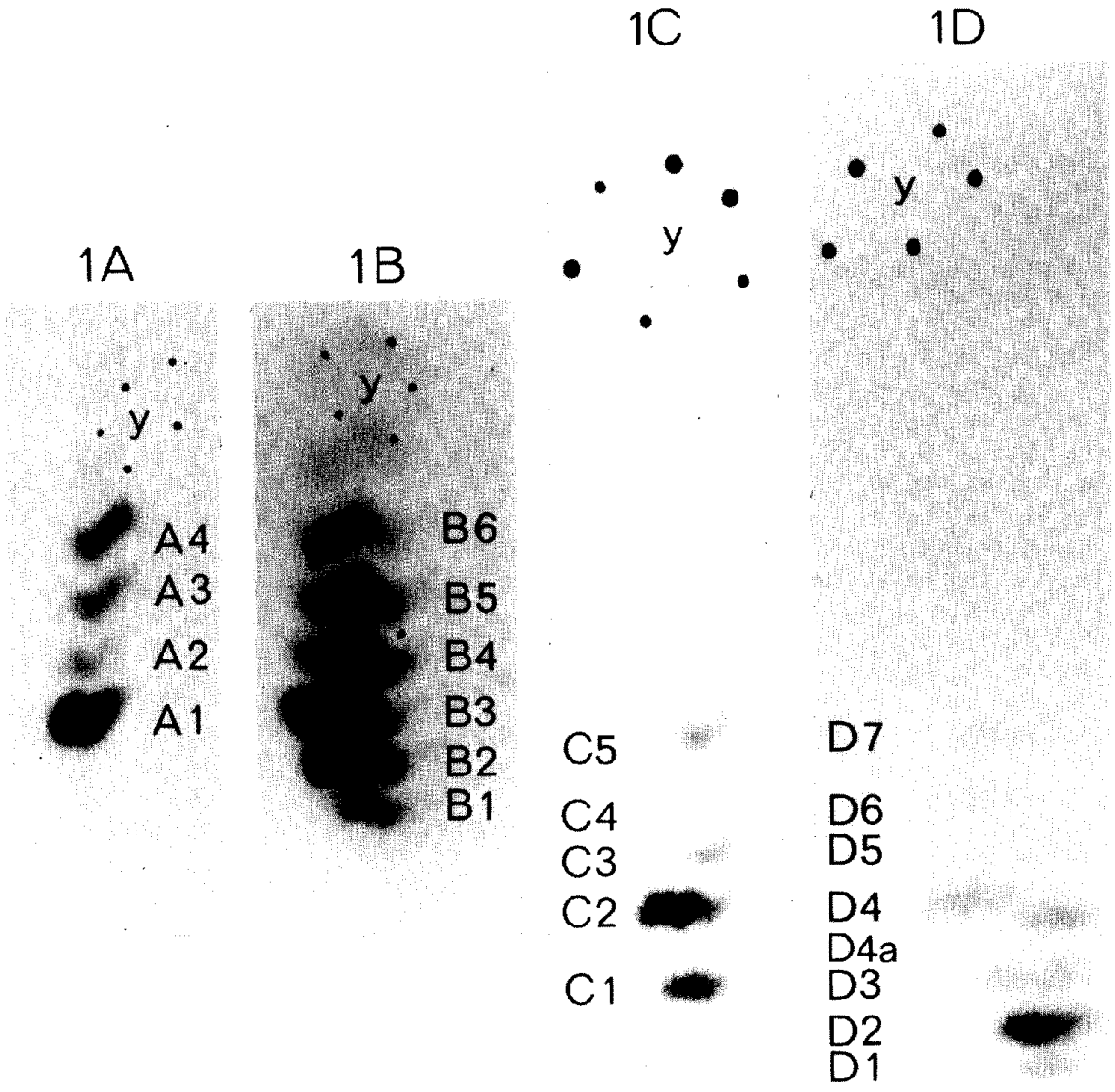

Fig. 1. Two-dimensional fractionations of partial transcripts of globin mRNA, using ionophoresis on Cellogel at pH 3.5 in the first dimension (right to left) and in the second dimension homochromatography on DEAE thin layers (bottom to top) (Y is the yellow, methyl orange marker). Incubations were carried out as in the Methods except that the reaction was limited to 1 min at 25°C. and the following triphosphates were used: 1A, [^{32}P] dGTP; 1B, [^{32}P] dGTP, dCTP; 1C, [^{32}P] dCTP ; 1D, [^{32}P] dCTP, dGTP, dATP. 1A and 1B were run in parallel, but in a separate experiment from 1C and 1D which were also run in parallel. The origin of each fingerprint is not visible as only the product-containing part is shown.

products B4–B6 and C3–C5 presumably derive from exonuclease activity of DNA polymerase or from further primer heterogeneity.

The third residue was established by incubation with [α-^{32}P] dCTP and unlabelled dGTP and dATP. The result (fig. 1D) shows the presence of three pro-

Table 2
Nearest neighbour information.

Spot no.†	Input label	Analysis	Conclusion
A1–A4	[α-^{32}P] dGTP	Tp	—TpG$_{OH}$
B1–B6	[α-^{32}P] dGTP	Tp	—TpG$_{OH}$
C1–C5	[α-^{32}P] dCTP	Gp	—GpC$_{OH}$
D1–D7	[α-^{32}P] dCTP	Gp	—GpC$_{OH}$
E	[α-^{32}P] dATP	CP and Ap††	—CpApAp

† Refers to figs. 1 and 2.
†† Yields estimated as equimolar by visual inspection of radioautograph.

ducts (D1, D2 and D4a) on a fingerprint otherwise similar to fig. 1C. From their positions, D3, D4, D5, D6 and D7 are assumed to be the same as C1–C5. Nearest neighbour analysis of D1–D7 again shows only G$\overset{*}{p}$C sequences (table 2). This suggests that D1 and D2 represent the addition of at least one A residue. D4a presumably represents the extension of OH T(pT)$_{11}$OH, a contaminant in the primer, to OH T(pT)$_{11}$pG$\overset{*}{p}$CpA$_{OH}$. A further incubation with [α-^{32}P] dATP, dGTP and dCTP at higher temperature and longer time is shown in fig. 2. One major product is present which gives both Cp and Ap on nearest neighbour analysis (table 2) implying the sequence C$\overset{*}{p}$A$\overset{*}{p}$A$_{OH}$ at its 3′-end. Its overall sequence is thus likely to be (pT)$_{12}$pGpC$\overset{*}{p}$A$\overset{*}{p}$A$_{OH}$. As this structure represents the end product of synthesis in the absence of TTP we can assume the next residue is likely to be T.

Although (pT)$_{12}$pGpCpApA$_{OH}$ is likely to be the sequence which results from the extension of the primer hybridised to poly A in the mRNA, the following sequences, viz: (pT)$_{12}$pGpT$_{OH}$, (pT)$_{12}$pGpCpT$_{OH}$ and (pT)$_{12}$pGpCpApT$_{OH}$, cannot be discounted as in the absence of TTP, (pT)$_{12}$pG$_{OH}$, (pT)$_{12}$pGpC$_{OH}$ and (pT)$_{12}$pGpCpA$_{OH}$ respectively would accumulate. These products, being identical to the intermediates of (pT)$_{12}$pGpCpApA$_{OH}$ are not distinguishable.

In summary, both the nearest neighbour (table 2) and 'positional' evidence (figs. 1 and 2) support the sequence (pT)$_{12}$pGpCpApA$_{OH}$. The next residue is likely to be a T, although this is inferred rather than experimentally demonstrated. Therefore, assuming correct base-pairing, this represents a sequence of (5′)ApUpUpGpCp(Ap)$_{12}$(3′) likely to be present in all transcribed poly A regions of the globin mRNA preparation.

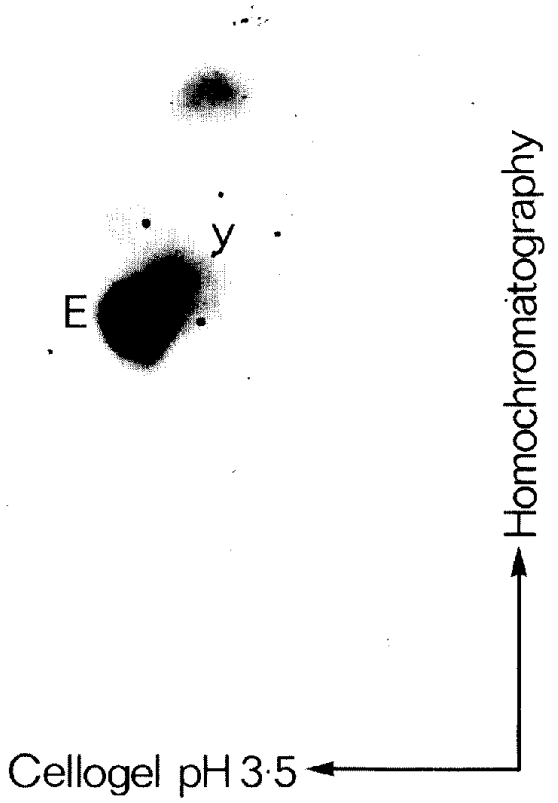

Fig. 2. Two-dimensional fractionation obtained as in fig. 1, with the following alterations: incubations were carried out at 37°C for 5 min and the triphosphates used were [^{32}P] dATP, dGTP and dCTP.

4. Discussion

The sequence studies described here are of interest for several reasons. A unique sequence A–U–U–G–C-poly A has been deduced for residues adjacent to poly A in rabbit globin and there is no evidence of any sequence heterogeneity in these experiments. Thus, as both α and β mRNA species were shown to be present in our preparation, it would appear likely that both

messengers have this sequence in common. The possibility of one sequence being selected, however, cannot be discounted until the experiments can be repeated with pure α and β mRNA's.

The nucleotide sequence is consistent with the preliminary studies of Hunt [8] showing the presence of a sequence G–Py–A_6 and G–Py–A_7 at the 3′-termini of rabbit globin mRNA. Similar but not identical results were obtained previously by Burr and Lingrel [9]. However, it is unclear whether our results are directly comparable with these studies (which report the presence of a short poly A tract) as recent reports show that much longer (larger than 50 residues) poly A sequences are present in rabbit globin mRNA [10] and at the 3′-terminus of mouse globin mRNA [11]. It may be that globin mRNA has a variable length of poly A at its 3′-terminus.

There is already good evidence that poly A is added to many mRNA's by a post-transcriptional process catalysed by a nuclear poly A polymerase [12]. The sequence adjacent to the poly A should therefore be the last transcription product of the DNA-dependent RNA polymerase and could well represent part of a recognition site for such poly A adding enzymes [13]. The evidence that α and β rabbit globin mRNA's may contain the same sequence is consistent with this hypothesis. Alternatively this sequence may reflect the common evolutionary origin of α and β-globin mRNA rather than a general sequence present in all messenger RNA's.

The reverse transcriptase activity of DNA polymerase reported in this paper will be described in more detail elsewhere. However, the fact that a unique sequence can be obtained using this activity, in combination with the control experiments described in the first part of this paper, fully substantiates this activity and shows that faithful DNA copies are probably made. It would therefore seem possible that all replicating cells, and not just those infected with oncogenic viruses, contain potential reverse transcriptase activity, though this has not yet been demonstrated in vivo.

References

[1] Adesnki, M., Salditt, M., Thomas, W. and Darnell, J.E. (1972) J. Mol. Biol. 71, 21–30.
[2] Jelinek, W., Adesnik, M., Salditt, M., Sheiness, D., Wall, R., Molloy, G., Philipson, L. and Darnell, J.E. (1973) J. Mol. Biol. 75, 515–532.
[3] Brownlee, G.G. and Sanger, F. (1969) Eur. J. Biochem. 11, 395–399.
[4] Wu, R. and Taylor, E. (1971) J. Mol. Biol. 57 491–511.
[5] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. U.S. 69, 1408–1412.
[6] Sanger F., Donelson, J.E., Coulson, A.R., Kossel, H. and Fischer, D. (1973) Proc. Natl. Acad. Sci. U.S 70. 1201–1213.
[7] Kelly, R.B., Atkinson, M.R., Huberman, J.A. and Kornberg, A. (1969) Nature 224, 495–501.
[8] Hunt, J.H. (1973) Biochem. J. 131, 315–325.
[9] Burr H. and Lingrel, J.B. (1971) Nature New Biol. 233, 41–43.
[10] Proudfoot, N.J. and Brownlee G.G. (1973) 9th Int. Congr. Biochem., Stockholm, p. 178.
[11] Mansbridge, J.N., Crossley, J.A., Lanyon, W.G. and Williamson, R. (1973) Biochem. Soc. Trans. 1, 588–589.
[12] Winters, M.A. and Edmonds, M. (1973) J. Biol. Chem. 248, 4756–4762.
[13] Winters, M.A and Edmonds, M. (1973) J. Biol. Chem. 248. 4763–4768.