Volume 6 Number 4 April 1979 Nucleic Acids Research

Influence of nucleotide sequence adjacent to duplex DNA termini on 3'terminal labeling by terminal transferase

Ranajit Roychoudhury, Chen-Pei D.Tu and Ray Wu§

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Wing Hall, Ithaca, NY 14853, USA

Received 6 February 1979

ABSTRACT

We have analyzed the effect of base-pairing at the exposed ends of Hinc II fragments of SV40 DNA on the efficiency of ribonucleotide incorporation catalyzed by terminal transferase. Wide variations in the labeling efficiency of individual DNA fragments have been observed. To elucidate the nature of this variation at the molecular level, we have correlated this effect with nucleotide sequence adjacent to the cleavage site. We found that a G:C base pair right at the exposed end drastically reduces the incorporation of ribonucleotides. Furthermore, the higher the number of G:C base pairs adjacent to the exposed end, the greater the reduction in labeling efficiency. From these results, we conclude that the labeling efficiency is determined by the degree of 'terminal breathing' of the DNA molecule at the exposed end of the duplex.

INTRODUCTION

The enzyme terminal deoxynucleotidyl transferase requires a single-stranded DNA or oligonucleotide as primer for addition of either ribonucleotides[1,2,3] or deoxynucleotides[4]. A base-paired structure near the 3'-ends inhibits the transferase reaction[4-6]. This effect is more pronounced when the 3' terminal nucleotide is base-paired within a protruding 5'-end such as that found at the cohesive ends of λ DNA[7]. We have shown that this inhibition due to a base-paired structure can be substantially reduced by using Co^{+2} ion in place of Mg^{+2} ion whereby any given DNA duplex can be made to incorporate nucleotides[8]. However, we observed a difference in the labeling efficiency of different duplex DNA fragments produced by digestion with different restriction endonucleases.

The endonuclease Hind III makes a staggered cut and gives an A:T base pair at the 3'-end. SV40 DNA fragments produced by Hind III endonuclease were found to be efficiently labeled by terminal transferase[8]. However, we were intrigued to find that certain DNA fragments produced by other restriction endonucleases which also produce A:T base pairs at the 3'-end

were not efficiently labeled. This observation raised the possibility that perhaps the sequence of several nucleotides in the vicinity of duplex termini exert some influence on the addition of nucleotides at the 3' termini. In this communication we show that the nucleotide sequence adjacent to the exposed end considerably influences the incorporation of ribonucleotides at the 3'-end of duplex DNA. The labeling efficiency of different duplex DNA fragments can be correlated with the number of G:C base pairs and is probably determined by the degree of 'terminal breathing' at the exposed ends.

MATERIALS AND METHODS

[α-^{32}P]rCTP (100-300 Ci/mmole) was obtained from New England Nuclear Corporation. Covalently closed circular SV40 DNA was isolated as described earlier[9]. Terminal transferase (specific activity 20,000 units/mg) used in these experiments was isolated in the laboratory of Dr. Hans Kössel[3]. *E. coli* alkaline phosphatase (BAPF, 30 units/mg) and pancreatic DNase (2300 units/mg, RNase free, DPFF-grade) were purchased from Worthington. Restriction endonucleases *Hinc*II, *Hind*III, *Hae*III and *Bam*H1 were obtained from New England Biolabs.

Agarose gel and polyacrylamide gel electrophoresis were performed according to Wu *et al.*[9]. Electrophoresis on cellulose acetate strips, homochromatography[10,16] and calculation of mobility shift of oligonucleotides were carried out according to Tu *et al.*[11]. Separation of labeled DNA fragments by gel electrophoresis and their elution in purified state were achieved as described recently[9,12]. Ribonucleotide terminated DNA fragments were treated with alkali followed by phosphatase in order to obtain products containing single ribonucleotides[2]. Acid-insoluble radioactivity was determined by using filter paper disc method of Bollum[13]. Partial pancreatic DNase digestion was carried out as described earlier[9].

Digestion of DNA with restriction endonuclease - The reaction mixture (300 µl) contained 10 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 50 mM NaCl, 3 A_{260} unit (150 µg) of covalently closed circular SV40 DNA and 160 units of *Hinc*II endonuclease. After 12 hours at 37°, an aliquot (5 µl) was monitored by agarose gel electrophoresis to check the completeness of digestion.

Digestion of isolated labeled DNA fragments - The labeled DNA fragments with 3' terminal label on both strands were eluted from the gel and ethanol precipitated. In order to obtain DNA ends with label in only one strand, the isolated *Hinc*II fragments A, B and C (15 pmoles each) were digested

with 30 units of HindIII endonuclease in 60 µl of reaction mixture as described above at 37° for 7 hours. The fragment D (15 pmoles) and E (15 pmoles) were similarly digested (without NaCl) with 30 units of BamHl and 5 units of HaeIII respectively.

RESULTS

Incorporation of rCMP at the duplex blunt ends produced by HincII endonuclease. Among the different restriction endonuclease-digested DNA fragments studied, we have noticed a wide variation in the labeling efficiency of fragments produced by HincII endonuclease. Therefore, a detailed analysis of the labeling efficiency of these fragments is described. We have previously shown that the kinetics of ribonucleotide incorporation into single-stranded primer, $(dT)_{\overline{30}}$, were 5-6-fold higher with Co^{+2} ion as compared to that with Mg^{+2} ion at the start of the reaction, which gradually declined to a 2-3-fold difference at the plateau[8]. After investigating the same kinetics with duplex DNA (HincII fragments) as primer, we observed a 10-fold difference in the rate of incorporation of rCMP in the presence of Co^{+2} ion. Furthermore, such a difference is maintained throughout the incubation period including the stage when a plateau is reached (Fig. 1).

Nature of incorporation into individual HincII fragments. In order to determine the nature of incorporation into individual fragments we incubated 1 A_{260} unit (210 pmoles of 3'-OH ends) of SV40 HincII fragments with 40 units of terminal transferase for four hours and isolated the labeled DNA fragments by gel filtration. Of the total radioactivity applied to the column, 220 pmoles of ^{32}P activity were eluted with the excluded DNA fragments. This accounts for an average of one rCMP residue added per DNA strand. The DNA concentration in this experiment was relatively high (5 A_{260}/ml). At a DNA concentration of 1.5 A_{260}/ml an average of two rCMP residues were added (Fig. 1).

When the labeled DNA fragments were subjected to electrophoresis on polyacrylamide gel, a striking result was obtained. It is evident from the intensity of bands (Fig. 2) that all the bands were not labeled with equal efficiency. Due to a limitation in space, two other small fragments (Hind L, M) are not shown. Since each of these fragments has 3' terminal label on both the strands, it is possible that the two 3' terminal ends of the same fragment may be labeled with different efficiency.

In order to answer this question, we eluted the fragments A, B, C, D and E and digested these individually with other restriction endonucleases

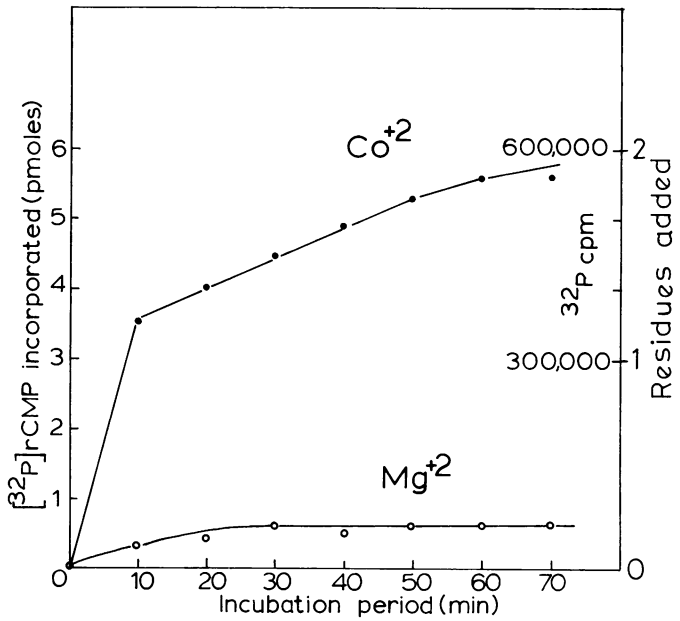

Fig. 1 - Incorporation of [^{32}P]rCMP into HincII digested DNA fragments in the presence of either of two metal ions.
The reaction mixture (100 µl) contained 0.15 A_{260} unit (30 pmoles of 3'-OH ends) of HincII digest of SV40 DNA, 700 pmoles of [α-^{32}P]rCTP (100,000 cpm/pmole), 100 mM K-cacodylate pH 7.0, 200 µM dithiothreitol, 16 units of terminal transferase and either 1 mM $CoCl_2$ or 10 mM $MgCl_2$. At intervals indicated, aliquots (10 µl) were monitored for acid-insoluble radioactivity.

in order to separate the two labeled ends. The bands A, B and C were digested with HindIII endonuclease. The bands D and E were digested with BamHl and HaeIII respectively. The digested fragments were subjected to electrophoresis. As can be seen from Figure 3, we found that the two ends of fragment A (A_1 and A_2) were labeled with equal efficiency. The radioactivity in the fragment B (B_1 and B_2) was about half of that observed in fragment A. The fragment C was labeled with high efficiency and the separated ends C_1 and C_2 contained equal amounts of radioactivity. A remarkable difference in the labeling efficiency of the two ends in the same duplex DNA fragment was noticed with fragments D and E. One end of fragment D (D_2) contained approximately 4-fold higher radioactivity as compared to that in the other end (D_1). Similarly, the radioactivity of fragment E_2 was about 4-fold higher than that of fragment E_1.

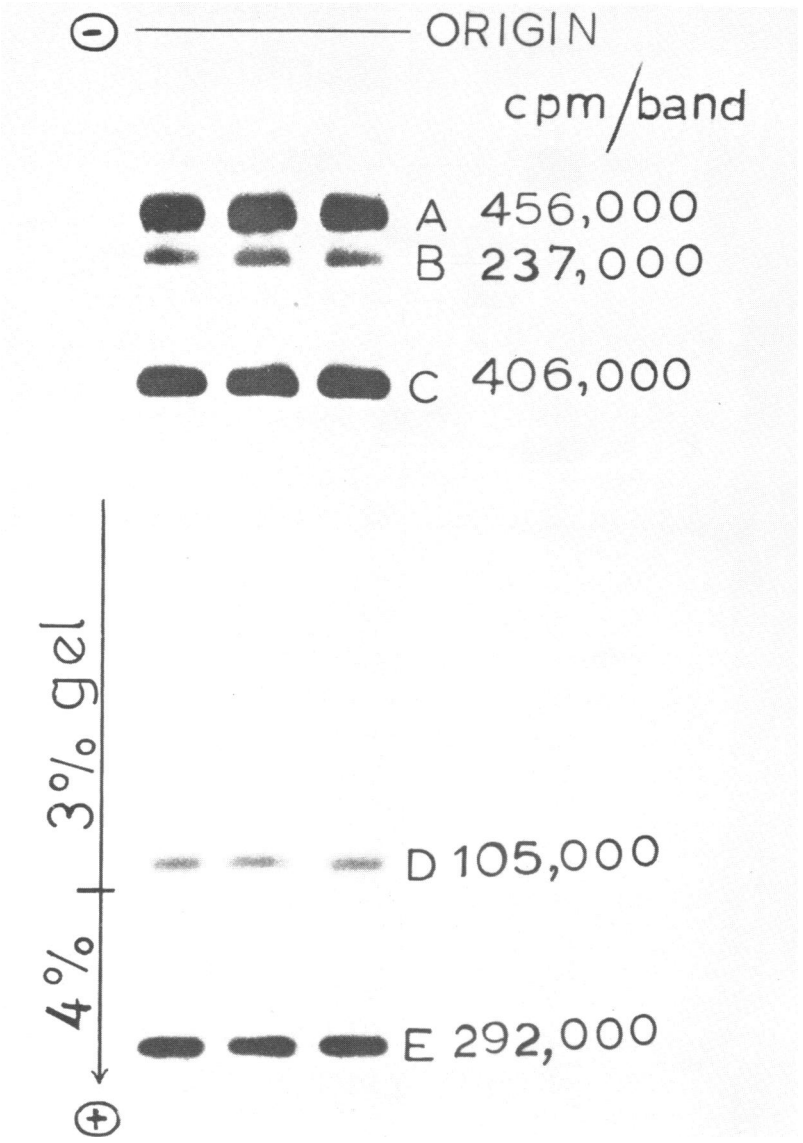

Fig. 2 - Nature of incorporation into individual duplex DNA fragments. The reaction mixture (200 µl) contained 1 A_{260} unit (210 pmoles of 3'-OH ends) of HincII digested SV40 DNA, 1980 pmoles of [α-^{32}P]rCTP, 1 mM $CoCl_2$ and 40 units of terminal transferase. Other components were those described in Fig. 1. After 4 hours at 37° the labeled DNA was isolated by gel filtration. An aliquot was monitored by polyacrylamide gel electrophoresis followed by autoradiography. The five HincII bands were cut out and counted for cerenkov radiation.

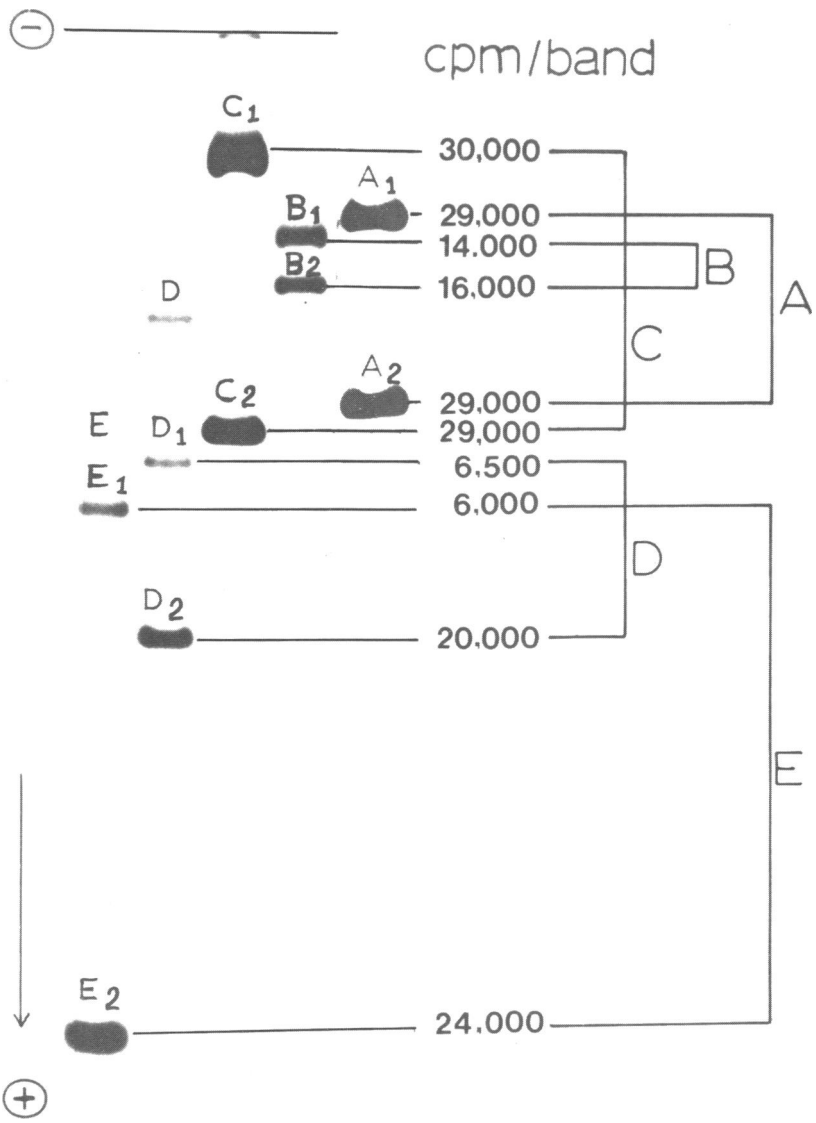

Fig. 3 - Separation of unique labeled ends by electrophoresis after restriction endonuclease digestion of individually isolated labeled DNA fragments.
HincII fragments A, B, C, D and E as shown in Fig. 2 were individually isolated and digested with restriction endonucleases as described in the "methods". The radioactive DNA fragments A, B, C, D and E contained 65,000, 30,000, 66,000, 26,000 and 32,000 counts per minute, respectively. The digestion mixtures were subjected to electrophoresis in 4% polyacrylamide gel and the labeled DNA bands were detected by autoradiography.

In order to understand the nature of these differences in the labeling efficiency of the duplex ends at a molecular level, we considered the possibility of correlating this effect with the nucleotide sequence adjacent to the cleavage site.

Correlation of nucleotide sequence adjacent to the cleavage site with labeling efficiency of the individual duplex ends.

DNA sequence analysis[10,11] of end labeled fragments shows that (Figure 4) the labeled A_1-end consists of the sequence 3' T-T-G-T-C-A-T-A-G-A-A..5'. Therefore, in its base-paired state, this duplex end consists of the following structure:

5'pA-A-C-A-G-T-A-T-C-T-T...
3' T-T-G-T-C-A-T-A-G-A-A...

Similarly, the other end (A_2) consists of the following structure:

5'pA-A-C-A-T-T-T-G-T-T-C-T-C-T...
3' T-T-G-T-A-A-A-C-A-A-G-A-G-A...

We note that within the first eight nucleotides from the exposed end, there are only two G:C base pairs (underlined). Consequently, both the exposed ends are expected to exhibit the same degree of 'terminal breathing'. Since terminal transferase prefers a single-stranded DNA as primer[4], then a greater 'terminal breathing' will favor a greater priming activity of these ends. Indeed, both 3' ends of fragment A_1 and A_2 were labeled equally and efficiently.

Using similar analysis (maps not shown), the duplex structures at the exposed ends of B_1 and B_2 were found to be:

5'pA-A-C-T-G-G-T-A-A-G-T-T... (B_1)
3' T-T-G-A-C-C-A-T-T-C-A-A...

and

5'pA-A-C-T-G-T-A-G-C.... (B_2)
3' T-T-G-A-C-A-T-C-G....

In this case, within a span of eight nucleotides from the exposed end, both the structures consist of three G:C base pairs. Thus, these ends were more strongly H-bonded as compared to the structures observed with A_1 and A_2. Consequently these ends were less efficiently labeled.

When we consider the nucleotide sequence adjacent to the exposed ends of the fragment C, we notice that the C_1-end and the C_2-end reveal the sequences (maps not shown):

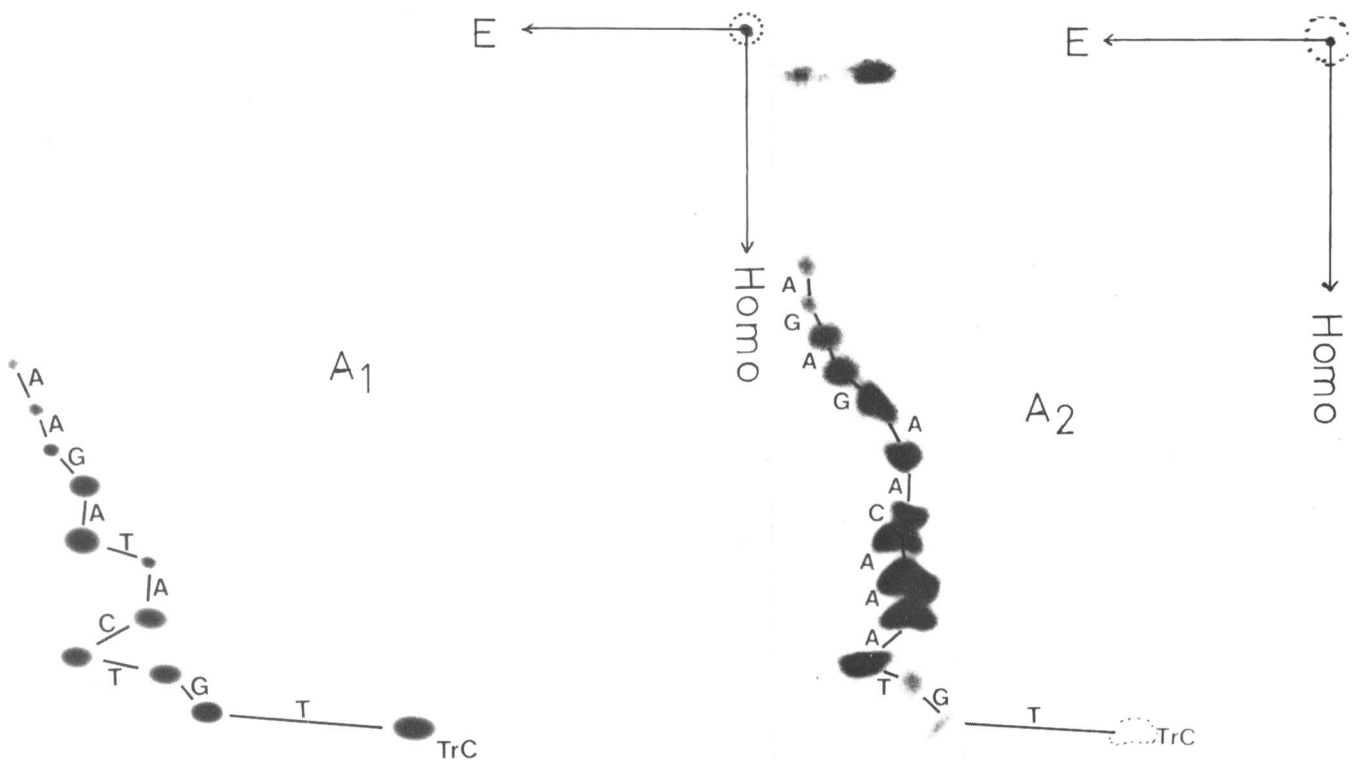

Fig. 4 - The 3' terminal nucleotide sequence at the two ends of fragment A.
The two ends of fragment A (A_1 and A_2) as shown in Fig. 3 were eluted, treated with alkali and phosphatase and isolated as single ribonucleotide addition product as described in "methods". After partial pancreatic DNase digestion, these were subjected to two-dimensional mapping[10,16]. The dotted circles, E and Homo indicate origin of application, electrophoresis and homochromatography, respectively.

5'pA-A-C-A-A-C-A-A-C-A-A-T-T... (C_1)
3' T-T-G-T-T-G-T-T-G-T-T-A-A...

5'pA-A-C-A-A-A-C-A-G-T-T-T... (C_2)
3' T-T-G-T-T-T-G-T-C-A-A-A...

Similar to fragment A_1 and A_2, both C_1 and C_2 contain only two G:C base pairs within a span of eight nucleotides from the exposed end of the duplex. Consequently, the C_1 and C_2 ends are labeled with an efficiency similar to that observed with the duplex ends A_1 and A_2.

The nucleotide sequence adjacent to the D_1-end was found to be:

5'pG-A-C-A-T-T-T-G-T-G-G-G-C... (D_1)
3' C-T-G-T-A-A-A-C-A-C-C-C-G...

Thus, within a stretch of eight nucleotides adjacent to the exposed end, there are three G:C base pairs. Furthermore, there is a G:C base pair right at the exposed end and two G:C base pairs within the first three nucleotides from the end. Therefore, the D_1-end was labeled with extremely poor efficiency.

The sequence at the D_2-end was obtained as:

5'pA-A-C-T-T-G-T-T-T-A-T-T... (D_2)
3' T-T-G-A-A-C-A-A-A-T-A-A...

Within a stretch of eight nucleotides, there are only two G:C base pairs. Furthermore, the exposed end in this case consists of an A:T base pair. Therefore, this end was labeled with high efficiency.

The sequence at the E_1-end was found to be:

5'pG-A-C-A-G-T-C-A-G-C... (E_1)
3' C-T-G-T-C-A-G-T-C-G...

Again, a G:C base pair is right at the exposed end and there are two G:C pairs within the first three nucleotides from the exposed end. Furthermore, there are four G:C base pairs within the stretch of eight nucleotides from the exposed end. This high G:C content may explain the very low efficiency of labeling by terminal transferase. In contrast, the E_2-end reveals the sequence:

5'pA-A-C-A-G-C-A-G-A... (E_2)
3' T-T-G-T-C-G-T-C-T...

Although there are also four G:C base pairs within a span of eight nucleotides, there is only one G:C base pair within the first three nucleotides

from the exposed end. Therefore, the E_2-end was labeled with a 4-fold higher efficiency as compared to E_1-end.

DISCUSSION

We have previously shown that the primer-specificity of terminal transferase is substantially altered in the presence of Co^{+2} ion whereby all forms of duplex DNA could be labeled with either ribonucleotides or deoxynucleotides[8]. Although, the inhibition of labeling at the 3'-end due to base pairing is reduced, there still remained a considerable variation in the labeling efficiency of DNA fragments generated by different restriction endonucleases. Our initial attempt to explain this variation on the basis of base pairing at the recognition sequence of restriction endonucleases provided only a partial answer. When we examined the nucleotide sequence more thoroughly, we noticed that one, two or even three identical base pairs at the 3'-end do not necessarily ensure equal labeling of different duplex DNA fragments. For example, all four exposed duplex ends in HincII fragments A and B consist of identical three nucleotide base pairs 5'pA-A-C-.
3' T-T-G-

However, the two ends in fragment B are labeled with about half the efficiency of those observed with fragment A. When we consider longer nucleotide sequence up to eight base pairs from the exposed end, we notice that the strands in the B fragment contain more G:C base pairs as compared to those in fragment A.

We have further noticed a drastic reduction in labeling efficiency, when a G:C base pair exists right at the exposed end (e.g. D_1-end, E_1-end). Moreover, the greater the number of G:C base pairs closer to the exposed end, the greater the reduction in labeling efficiency (D_1-end and E_1-end).

Similar reduction in labeling efficiency due to G:C rich sequence near the exposed end was noticed using polynucleotide kinase but the extent of reduction was somewhat less dramatic (Roychoudhury, Tu and Wu, unpublished). Consistent with the above result was the observation that the exposed end produced by SmaI enzyme (which gave three consecutive G:C pairs at the exposed end) was also poorly labeled by polynucleotide kinase (Rothstein and Wu, unpublished). Therefore, it appears that the nucleotide sequence adjacent to the exposed end of the duplex DNA fragments considerably influences the labeling efficiency of these enzymes. Our results with

HincII fragments as examples strongly suggest that the degree of 'terminal breathing' of double-stranded DNA determines the labeling efficiency.

Analysis of the DNA sequence presented here was completed in early 1977 at a time when only some of these sequences were published by others. Since then, all of the sequences have been confirmed by the reports of the complete sequence analysis of SV40 DNA by Reddy et al.[14] and Fiers et al[15].

ACKNOWLEDGEMENTS

This work was supported by research grants CA14989 and GM 24904 from the National Institute of Health, U.S. Public Health Service.

§To whom correspondence should be addressed.

REFERENCES

1. Roychoudhury, R. and Kössel, H. (1971) Eur. J. Biochem. 22, 310-320.
2. Kössel, H. and Roychoudhury, R. (1971) Eur. J. Biochem. 22, 271-276.
3. Kössel, H., Roychoudhury, R., Fischer, D. and Otto, A. (1974) Methods Enzymol. 29E, 322-341.
4. Bollum, F. J. (1974) The Enzymes 10, 145-171.
5. Jensen, R. H., Wodzinski, R. J. and Rogoff, M. H. (1971) Biochem. Biophys. Res. commun. 43, 384-392.
6. Lobban, P. E. and Kaiser, A. D. (1973) J. Mol. Biol. 78, 453-471.
7. Wu, R. and Taylor, E. (1971) J. Mol. Biol. 57, 491-511.
8. Roychoudhury, R., Jay, E. and Wu, R. (1976) Nucleic Acids Res. 3, 863-877.
9. Wu, R., Jay, E. and Roychoudhury, R. (1976) Methods Cancer Res. 12, 87-176.
10. Brownlee, G. G. and Sanger, F. (1969) Eur. J. Biochem. 11, 395-399.
11. Tu, C. D., Jay, E., Bahl, C. P. and Wu, R. (1976) Anal. Biochem. 74, 73-96.
12. Roychoudhury, R. and Wu, R. (1979) Methods Enzymol. (in press).
13. Bollum, F. J. (1959) J. Biol. Chem. 234, 2733-2734.
14. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L. and Weissman, S. M. (1978) Science 200, 494-502.
15. Fiers, W., Contreras, R., Haegeman, R., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. and Ysebaert, M. (1978) Nature 273, 113-120.
16. Jay, E., Bambara, R., Padmanabhan, R. and Wu, R. (1974) Nucleic Acids Res. 1, 331-353.