Volume 12 Number 4 1984 Nucleic Acids Research

Different restriction enzyme-generated sticky DNA ends can be joined *in vitro*

Mien-Chie Hung and Pieter C.Wensink

Department of Biochemistry and Rosenstiel Center, Brandeis University, Waltham, MA 02254, USA

Received 9 December 1983; Revised 12 January 1984; Accepted 18 January 1984

ABSTRACT

We describe a simple method for joining the 5'-protruding, single-stranded DNA ends generated by restriction enzymes. The method allows ends with different sequences to be joined and prevents identical ends from being joined. This is accomplished by partially filling the single strands in a controlled reverse transcriptase reaction. Partial filling can create new single-stranded ends that can be ligated to different, partially filled ends. In almost all useful cases, partial filling simultaneously eliminates the self-complementarity of identical ends and thus prevents them from being joined by DNA ligase. Although all possible combinations of partially filled ends were not tested, the tests performed indicate that the method is fairly general. We demonstrate that ends of the same length can be ligated with useful efficiency if they are: 1) one nucleotide long and complementary; 2) two nucleotides long and complementary or have a mismatch (dA:dC) at one position; or 3) three nucleotides long and, in our test, have a dT:dC mismatch at the middle position.

INTRODUCTION

Many methods are now used for *in vitro* construction of recombinant DNA molecules. These methods use steps taken from one of two general approaches. In one approach, single-stranded, self-complementary (sticky) ends are formed by restriction enzyme digestion and then joined by a DNA ligase reaction (1). This sticky-end approach is rapid and reliable, but is limited because it only joins identical sticky ends. Also, unless a dephosphorylation step is added to the procedure (2), this approach usually is inefficient because it constructs many unwanted joints, for example, those between the ends of a vector molecule. In some cases this inefficiency can also be eliminated by restriction digestion after ligation. For example, when two enzymes (e.g. Bgl II and Bam HI) recognize different sequences but generate identical sticky ends, these ends can be ligated to form joints which, unlike the ligated vector, are insensitive to either restriction enzyme. In the other general approach, blunt double-stranded ends are formed by one of several techniques, for example, by shear force (3) or S1 nuclease digestion (4).

These blunt ends are then recombined by ligating directly (5) or by adding homopolymer (6) or restriction sticky ends and then ligating (7,8). The blunt-end approach is laborious, less reliable and less efficient than the sticky-end approach.

In this paper we describe an alternative that is as rapid and, in most applications, nearly as efficient as the sticky-end approach. It also has some advantages. It increases the number of different in vitro recombinations that can be done efficiently because it allows most sticky ends to be joined to most other different sticky ends. Furthermore, it blocks joining of identical sticky ends. These two features allow in vitro recombination to be directed to occur between particular ends and therefore between particular DNA molecules.

The approach requires partial and controlled filling of restriction enzyme-generated sticky ends. Although the method applies to other sticky ends, in this paper we will only discuss its use in joining four nucleotide long, 5'-protruding sticky ends. Most of these ends have an internal, symmetrical complementarity. That is, the first and fourth nucleotides are complementary and so are the second and third. For these ends partial filling of identical ends yields single strands that are no longer self-complementary and, therefore, almost always unable to be annealed and ligated to each other. However, a partially filled end can be ligated to a different, partially filled end. An example of this strategy is shown in Figure 1a. A Xba I generated end, 5'-CTAG-3', and a Hind III generated end, 5'-AGCT-3', cannot be joined by ligase. However, if each end is filled with two nucleotides, then the 5'-CT-3' terminus of the Xba I end can be ligated to the complementary 5'-AG-3' terminus of the Hind III end. These two altered ends can be ligated together without any significant competing reaction joining the identical, modified ends to each other.

The method is fairly general because, as we demonstrate in this paper, ligation occurs between perfectly matched single-stranded ends that are one or two nucleotides long and, in the cases tested, between imperfectly matched ends that are two or three nucleotides long. Using these observations we predict that joining can be accomplished between approximately 70% of all possible combinations of the different four nucleotide, 5'-protruding sticky ends. As an illustration of the use of this method we have constructed a table (Fig. 1b) of the number of nucleotides that need to be filled to allow ligation between ends generated by seven commonly used restriction enzymes. In the course of our work we have used the filling plan of this table to

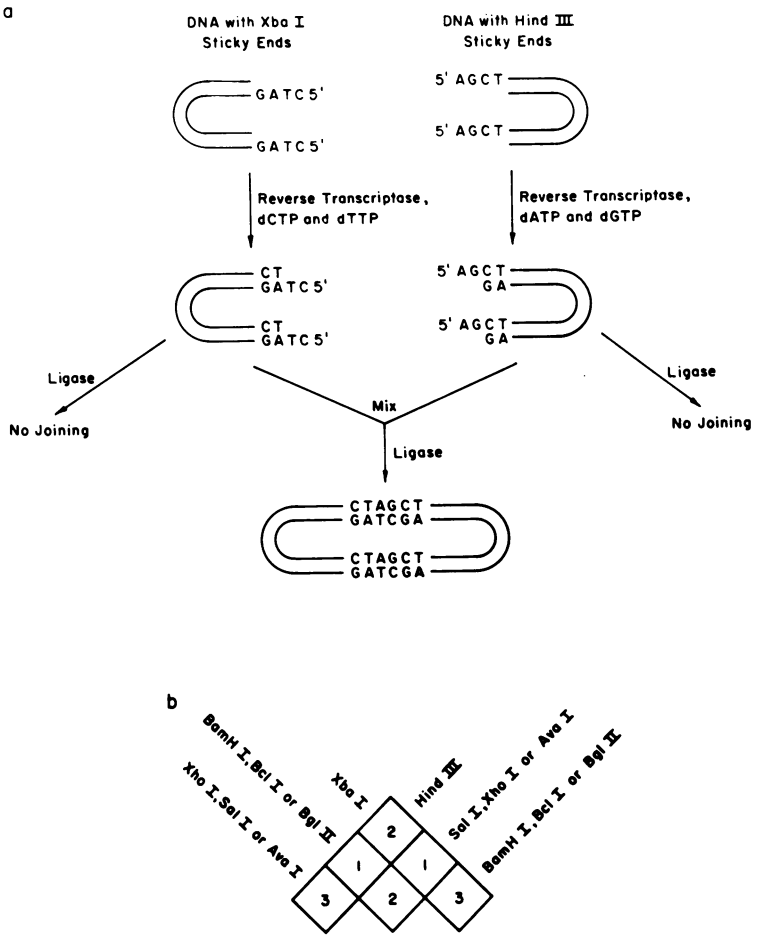

Figure 1--Strategy for joining different sticky ends. a) A plan for joining Xba I and Hind III generated sticky ends. Thin and thick lines represent two different DNA molecules. b) The number of nucleotides that must be filled in order to join ends generated by seven commonly used restriction enzymes.

construct all of these joints.

MATERIALS AND METHODS

DNAs. The plasmid pUC12 (a gift from J. Messing and J. Vieira) is derived from the pUC8 plasmid (9). It has a Xba I restriction site inserted into the poly linker region. The ampicillin sensitive clone, pYP3c1, contains a cDNA copy of the 3'-terminal 500 nucleotides of yp3 mRNA. This clone was constructed and generously provided by M. Garabedian. The yp3 mRNA codes for

yolk protein 3 of Drosophila melanogaster (10).
Enzyme reactions. The sticky ends were filled by the reverse transcriptase of Avian Myeloblastosis Virus (Life Sciences) using previously described reaction conditions (11). We assume that the large proteolytic fragment of E. coli DNA polymerase I can also be used for this filling reaction. The partial filling reactions were limited by controlling which nucleoside triphosphate precursors (PL laboratories) were included in the reactions. The ligation and restriction enzyme reactions were done as previously described (11). In most cases ligation was carried out at 15° C with 4 units/µl of Bacteriophage T4 DNA ligase (New England Biolabs) and 5 µg/ml vector DNA. The ligation of mixtures of identical and non-identical ends with perfect complementarity were done under these conditions except that after 4 hr the reaction was diluted 5-fold and ligation continued for another 20 hr. In these reactions the molar ratio of insert to vector molecules was 5:1 unless specified otherwise in the text. Restriction enzymes were purchased from New England Biolabs and Boehringer-Mannheim.
Transformation assays. The DNAs were transformed (11) into E. coli K12 strain HB101 (12) and assayed on X-gal plates (13) that contained 100 µg/ml ampicillin.

RESULTS

Ligation between identical, partially filled ends is inefficient

We tested the efficiency of covalent joining between identical sticky ends filled with either one, two, three or four nucleotides. The plasmid pUC12 was digested with the Xba I restriction enzyme, thereby cleaving the plasmid at a single site. The 5'-CTAG-3' sticky end was filled in with either one (dC), two (dCdT), three (dCdTdA) or four (dCdTdAdG) nucleotides by using the appropriate nucleoside triphosphates and the reverse transcriptase from Avian Myeloblastosis Virus. Each of these reaction products as well as Xba I digested pUC12 with unfilled ends was separately treated with bacteriophage T4 DNA ligase. Agarose gel electrophoresis of the products (Fig. 2) demonstrated that, as expected from earlier observations (1,5), T4 DNA ligase joins completely filled (blunt) ends much less efficiently than unfilled sticky ends. For the purposes of this paper the more important observation is that the gel does not reveal any joining of partially filled ends.

The efficiency of ligation was also assayed by bacterial transformation in order to gain a quantitative measure of the viable self-closed molecules that would be background in the directed joining scheme presented in this

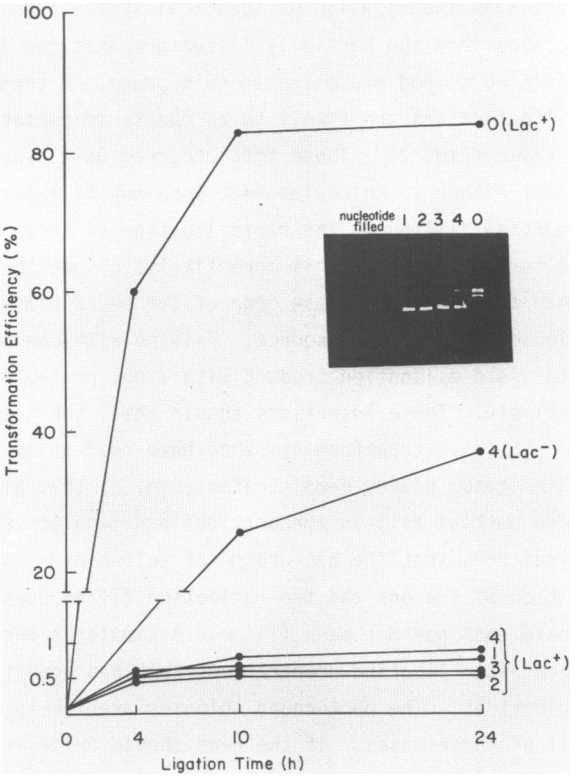

Figure 2--Efficiency of ligating partially filled, identical DNA ends. The sticky ends of Xba I digested pUC12 were filled with one (indicated by 1), two (?), three (3) or four (4) nucleotides or left unfilled (0). They were separately ligated for the times indicated and their transformation efficiency was determined relative to undigested pUC12 (4.5×10^3/ng DNA = 100%). (Lac^+) indicates the colonies that had β-galactosidase activity according to the color indicator in the plates. (Lac^-) indicates colonies that do not have this activity. The inset shows an agarose gel electrophoresis pattern of DNA ligated for 24 hr.

paper. Since the pUC12 plasmid contains a gene for ampicillin resistance, the frequency of cell transformation by closed DNA circles can be detected as the number of ampicillin resistant, transformed cells. The transformation assay indicates (Fig. 2) that the unfilled sticky ends were rapidly ligated. At the ligation plateau approximately 85% of the molecules had been rejoined. The blunt-end ligation was much slower and perhaps would eventually have reached the same plateau as sticky-end ligation. The partially filled ends ligated with the same kinetics as sticky ends, but the plateau value was approximately 100 fold lower. We conclude that partial filling causes a

substantial reduction in the ligation of identical sticky ends.

Viable molecules from the partially filled preparations are of most concern for the joining method presented in this paper. A very small portion occurred without ligation and are likely to be due to incomplete digestion by the restriction enzyme (Fig. 2). Those that occurred as a result of ligation are likely to be due either to molecules that were not filled or that were ligated despite partial filling. The rapid ligation of these molecules suggests that the former possibility is more likely. Since the Xba I cleavage was within the β-galactosidase gene of the pUC12 plasmid, we could obtain a better indication of their source. Filling with one or two nucleotides is likely to yield a ligation product with a one or two base pair insertion, respectively. These insertions should shift the translation reading frame and thus give transformants that have no β-galactosidase activity. Color indicator plates demonstrated (Fig. 2) that all transformants from these partial fill-in preparations had β-galactosidase activity. This indicates that the background of self-closed molecules generated by ligation of the one and two nucleotide filled preparations were from molecules whose ends had not been filled. A similar transformation result was obtained after ligating preparations that had been filled with three or four nucleotides. The background colonies are likely to have the same source in all of these cases. If the need should arise, this background could be reduced by dephosphorylating the vector ends (2) or by digesting the ligated molecules with the restriction enzyme used to generate sticky ends in the vector (see below). However, as will be shown below, the decrease achieved by partial filling is adequate for most DNA constructions.

Ligation is efficient between complementary one or two nucleotide sticky ends made by partial filling

As outlined in the Introduction, many different sticky ends can be made complementary by filling part of their sticky ends. We first tested the efficiency of ligating different sticky ends that had been filled to yield two nucleotide long single-stranded ends. The ampicillin sensitive pYP3c1 plasmid was linearized with the Hind III restriction enzyme. The ends were filled to generate a two nucleotide 5'-AG-3' end. A Xba I digest of ampicillin resistant pUC12 was filled to generate a 5'-CT-3' end. As diagrammed in Figure 1a, these two preparations were either separately treated with ligase or were mixed and ligated. Because the mixture could lead to the formation of an oligomer with several alternating pUC12-yp3 subunits, the ligation reaction was first done at high DNA concentration to

favor the bimolecular reaction between pUC12 and yp3. Then the reaction was diluted to favor the unimolecular DNA ligation that leads to circle formation (see Materials and Methods). The molar ratio between the two DNAs was also varied to determine the ratio that would optimize for insertion of the yp3 gene into the pUC12 vector.

The ligation products were examined by a transformation assay that detected only cells transformed with the ampicillin resistant pUC12. Color indicator plates were used to determine the presence of an insert within the β-galactosidase gene of pUC12. The results shown in Table I indicate that a five to one molar ratio of insert fragment to vector is optimal for these molecules. At this ratio there are three- to four-fold more transformants with an insert than without an insert. Table I also indicates that joints between the perfectly matched two nucleotide sticky ends are efficiently formed (230 per nanogram of vector DNA). Thus sticky ends generated by different restriction enzymes can be efficiently joined with little background if limited filling can produce complementary two-nucleotide sticky ends.

One nucleotide sticky ends were formed by filling in three nucleotides of the yp3 Bam HI sticky ends (5'-GATC-3') and the pUC12 Xba I sticky ends (5'-CTAG-3'). These complementary 5'-G and 5'-C ends were ligated (Fig. 3) and yielded only slightly fewer pUC12 transformants with yp3 inserts than did the two-nucleotide sticky ends (Table II). The background of uninserted

Table I
Transformation of DNAs ligated at different molar ratios.

Molar ratio of insert* to vector† DNA	Number of colonies per ng of vector† DNA	
	(Lac⁻)	(Lac⁺)
0:1	0	70
2:1	86	65
5:1	230	71
10:1	43	67

* Insert DNA was pYP3c1 (see Materials & Methods) linearized by Hind III digestion and filled with dATP and dGTP (Fig. 1a).
† Vector DNA was pUC12, linearized by Xba I digestion and filled with dCTP and dTTP (Fig. 1a).
The Lac⁺ and Lac⁻ abbreviations indicate colonies with (+) and without (-) β-galactosidase activity as detected by color-indicator plates.

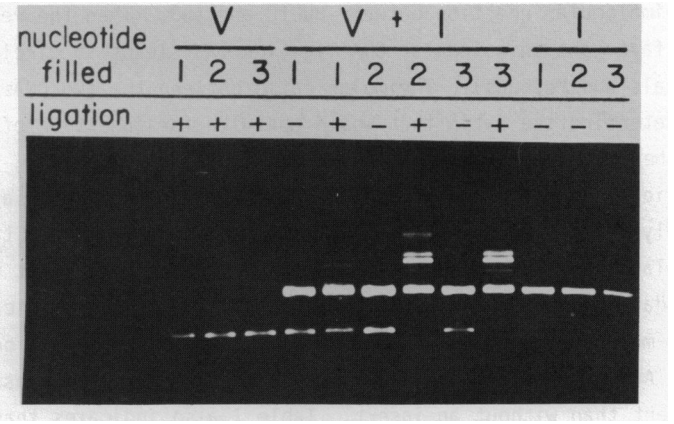

Figure 3--Ligation between different, partially filled ends. This is an agarose gel electrophoresis pattern of products from ligation reactions containing pUC12 DNA (V for vector) and/or pYP3c1 (I for insert). + indicates that ligase was included in the reaction and - indicates that it was not. The numbers indicate the number of nucleotides filled into the sticky ends. The vector DNA ends were made by Xba I and then filled as indicated in the text. In the insert DNA, a single nucleotide was filled into the Sal I-generated ends, two nucleotides were filled into the Hind III-generated ends and three nucleotides into the BamH I-generated ends.

Table II
Ligation after filling different sticky ends

Restriction enzyme digest		Sticky ends after filling	Number of colonies per ng of vector DNA	
insert(pYP3c1)	vector(pUC12)		(Lac$^-$)	(Lac$^+$)
BamH I	Xba I	5'-G and 5'-C	180	56
Hind III	Xba I	5'-AG and 5'-CT	220	55
EcoR I	Xba I	5'-AA and 5'-CT	1	75
EcoR I	Xba I	5'-AA and 5'-CT	1	10*
Sal I	Xba I	5'-TCG and 5'-CTA	6	78

The Lac$^+$ and Lac$^-$ designations are the same as in Table I. Undigested vector DNA gave 2000 colonies/ng DNA in this experiment.

* After ligation, the reaction mixture was extracted with phenol, ethanol precipitated, resuspended and digested with Xba I (1 unit/0.5µg DNA, 20µl reaction, 37°C, 1hr).

pUC12 was at the same level as in ligation of two nucleotide sticky ends. Thus one nucleotide sticky ends may also be joined efficiently with low background from self-closure of the vector, pUC12.

Ligation also occurs between two or three nucleotide sticky ends that have a single non-complementary nucleotide

Many different restriction enzyme-generated sticky ends could be joined if they could be partially filled and then ligated despite a single mismatch between the ends. Such joints can be formed. The yp3 clone was linearized by Sal I digestion and the sticky ends were filled to produce 5'-TCG-3' ends. The Xba I ends of pUC12 were filled to produce 5'-CTA-3' ends. These ends also joined (Fig. 3, lanes 4 and 5), but gave transformants with approximately 35 fold lower efficiency than the perfectly matched one- and two-nucleotide sticky ends. The background of vectors without inserts was at the same level in this experiment (Table II), so only 7% of the transformants had inserts of yp3 DNA. Thus three nucleotide sticky ends with a single internal mismatch can be joined with reasonable efficiency. A similar experiment demonstrated that a single mismatch between two nucleotide sticky ends (5'-AA-3' and 5'-TC-3') allowed ligation to occur with several fold lower efficiency and the same background (Table II). This background can be lowered by digesting the ligated molecules with Xba I, the enzyme used to generate the sticky ends of the vector (Table II). We conclude that a single nucleotide mismatch lowers the efficiency of ligating two or three nucleotide sticky ends, but still permits the filling method to be used.

An application of this method

An example of a DNA construction illustrates some of the utility of this method. We needed to insert a particular Hind III fragment into the Xba I site of a vector. The fragment to be inserted was one of two Hind III fragments of a plasmid. The other Hind III fragment had a DNA replication origin and also the same antibiotic resistance gene as the vector. Since this other fragment was nearly the same length as the sought after Hind III fragment, it could not be easily eliminated by physical methods, and would result in a large number of undesired, antibiotic resistant transformants if a conventional cloning strategy such as blunt-end ligation or ligation with linkers was used. For this reason, the method diagrammed in Figure 1a was used. Twelve randomly selected, transformed, antibiotic resistant colonies were examined. There had been no preselection for vectors with inserts. Six colonies had the vector containing the sought after Hind III fragment. Five had the other Hind III fragment inserted into this vector. The remaining

colony had a plasmid that appeared to be the second Hind III fragment whose two ends had joined, thereby forming a viable, ampicillin resistant plasmid. In this case there were fewer self-closed vector molecules than would be predicted by the model experiments reported in this paper.

DISCUSSION

Most of the possible four nucleotide, 5'-protruding sticky ends can be joined by the partial filling strategy

Most four nucleotide, 5'-protruding ends generated by restriction enzymes have their first and fourth nucleotides complementary and also their second and third complementary. The catalog of restriction enzymes available for DNA constructions will certainly increase and may eventually be a complete set that will generate all possible sticky ends of this type. For this reason we will discuss the generality of the filling approach to end joining from the viewpoint of joining all possible combinations of these ends. If ligation efficiency does not depend on G + C % base composition or on near neighbor effects of sequence, then it appears that single-stranded ends of the same length can be ligated if they are: 1) one, two or three nucleotides long and perfectly complementary; or 2) two or three nucleotides long and non-complementary in no more than one nucleotide. In the latter case it is likely that different mismatches can have different effects on ligation efficiency. For example, purine:purine mismatches are more likely to reduce ligation efficiency than are the pyrimidine:pyrimidine and the purine:pyrimidine (A:C) mismatches we have tested (14). It is also possible that a mismatch at the first or third positions of three nucleotide sticky ends causes a greater efficiency reduction than a mismatch at the middle position. The magnitude of these effects are likely to be influenced by the ionic conditions of the ligations reaction (14). Because of the complexity of these variables we make a simplifying assumption in order to estimate the

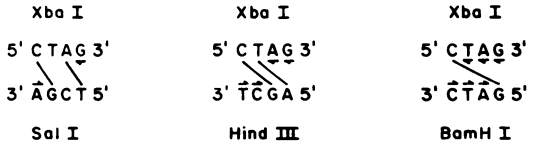

Figure 4--A diagrammatic method for determining whether two different sticky ends can be joined by the partial filling method. Base-pairing potential is indicated by solid lines. For each of the pairs of ends diagrammed, the small arrows indicate the nucleotides that must be filled to generate ligatable sticky ends.

general utility of this approach to joining sticky ends. We assume that any mismatch at any position may alter the efficiency but will still allow the method to be used. If this is the case, then 70% (168/240) of the joints between different four nucleotide ends can be formed. In all but four cases (two nucleotide filling of 5'-ATAT-3', 5'-TATA-3', 5'-GCGC-3' and 5'-CGCG-3') filling should allow these joinings to occur without a significant competing reaction joining the ends of the vector. Thus approximately 70% of the possible 240 different joining combinations are likely to occur with this method in the absence of a significant background from the vector molecule.

A simple diagrammatic method to determine whether two ends can be joined

If the sequences of two different sticky ends are known, then a simple diagrammatic method (Fig. 4) can be used to determine whether or not they can be joined by this partial filling approach. The sequence of one sticky end is diagrammed above the other in opposite polarity. If the final 5' nucleotide of each end is complementary, then three nucleotide filling can be used. If one of the final two 5' nucleotides are complementary in this orientation of the sticky ends, then two nucleotide filling can be used. If the final three are complementary with the exception of the middle nucleotide, then one nucleotide filling can be used.

ACKNOWLEDGEMENTS

We thank Robert Schleif, Howard Baum and the two anonymous referees for very useful criticism of this manuscript. This work was supported by a research grant (GM21626) from the National Institutes of Health.

REFERENCES

1. Mertz, J.E. and Davis, R.W. (1972) Proc. Natl. Acad. Sci. USA 69, 3370-3374.
2. Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H.M. (1977) Science 196, 1313-1319.
3. Hogness, D.S. and Simmon, J.R. (1964) J. Mol. Biol. 9, 411-438.
4. Seeburg, P.H., Shine, J., Martial, J.A., Baxter, J.D. and Goodman, H.M. (1977) Nature 270, 486-494.
5. Sgaramella, V., van de Sande, J.H. and Khorana, H.G. (1970) Proc. Natl. Acad. Sci. USA 67, 1468-1475.
6. Wensink, P.C., Finnegan, D.J., Donelson, J.E. and Hogness, D.S. (1974) Cell 3, 315-325.
7. Heyneker, H.L., Shine, J., Goodman, H.J., Boyer, H., Rosenberg, J., Dickson, R.E., Narang, D.A. and Riggs, A.D. (1976) Nature 263, 748-752.
8. Scheller, R.H., Dickerson, R.E., Boyer, H.W., Riggs, A.D. and Itakus, K. (1977) Science 196, 177-178.
9. Vieira, J. and Messing, J. (1982) Gene 19, 259-268.
10. Barnett, T., Pachl, C., Gergen, J.P. and Wensink, P.C. (1980) Cell 21, 729-738.

11. Schleif, R.F. and Wensink, P.C. (1981) Practical Methods in Molecular Biology (Springer Verlag, New York) Chapters 5 and 6.
12. Boyer, H.W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.
13. Miller, J.H. (1972) Experiments in Molecular Genetics, (Cold Spring Harbor Laboratory, New York), pp. 47-52.
14. Lomant, A.J., and Fresco, J.R. (1975) Prog. in Nucl. Acids Res. and Mol. Biol., 15, 185-218.