Cloning strategies for polymerase chain reaction products

Anjan Misra and Subrata Sinha

Department of Biochemistry, All India Institute of Medical Sciences, New Delhi 110 029, India

Cloning of the PCR products is being increasingly used by a number of laboratories working on a diverse range of problems. There are a number of methods available each one with its relative advantages and disadvantages. This review seeks to highlight some of these methods and discuss their practical applicability under different experimental conditions.

In the last 10 years, polymerase chain reaction (PCR) has emerged as a very useful and robust technique for researchers. Cloning of a PCR product is often useful for further characterization like restriction mapping, sequencing and also for expression and functional analysis. Sequencing of a cloned product is easier because many cloning vectors are commercially available, with a choice of specific primer sequences (M13 F & R, T3, T7, SP6) flanking the multiple cloning site in opposite directions. There are basically two approaches for cloning PCR products. The first is to start with any PCR-amplified fragment and insert it into the vector by ligation *in vitro* (ligation-dependent cloning). The second is to generate a PCR fragment, having suitable ends which can be inserted into the vector without ligating it enzymatically *in vitro* (ligation-independent cloning).

Ligation-dependent cloning

In ligation-dependent cloning, the DNA fragment of interest (insert) is ligated to the vector, using DNA ligase and then, this ligated product is used to transform bacteria. This is followed by selection of colony with right-sized insert. Here any DNA can be ligated into any vector. In this sense, this type of cloning offers versatile options.

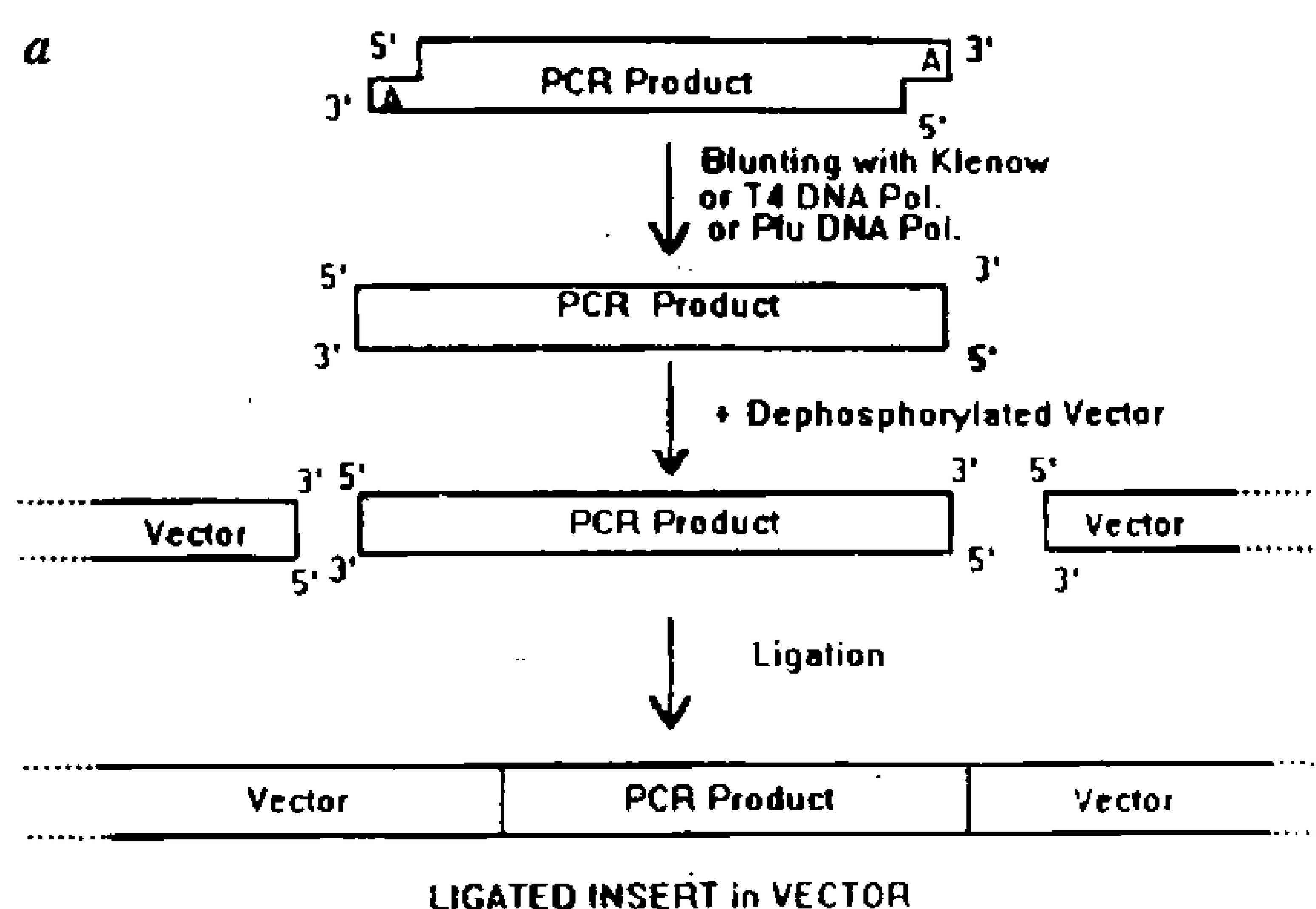
The strategies successfully employed for ligation-dependent PCR product cloning include:

Blunt end ligation

The conventional approach to clone PCR product is by blunt end ligation. Since the PCR product, amplified by *Taq* DNA polymerase has ragged ends, it is first smoothed by 3′ → 5′ exonuclease activity of either of the following enzymes: Klenow fragment of *E.coli* DNA polymerase, T4 DNA polymerase, Vent or deep Vent DNA polymerase and *Pfu* DNA polymerase. *Pfu* DNA polymerase seems to be the best of these enzymes though it is the most expensive. Also, a PCR reaction with either Vent or *Pfu* DNA polymerase gives rise to a blunt-ended product which does not need any polishing reaction.

The vector is digested with restriction enzymes (RE) (e.g. *Eco*RV, *Sma*I, *Hin*dII, etc.) that gives blunt ends. For some unknown reason/s ligation efficiency is 5–50 times more when *Eco*RV is used, compared to other enzymes[1]. The vector is dephosphorylated with alkaline phosphatase to reduce vector self-ligation. In next step, vector and insert are ligated to their blunt ends using T4 DNA ligase (Figure 1 *a*).

a



b

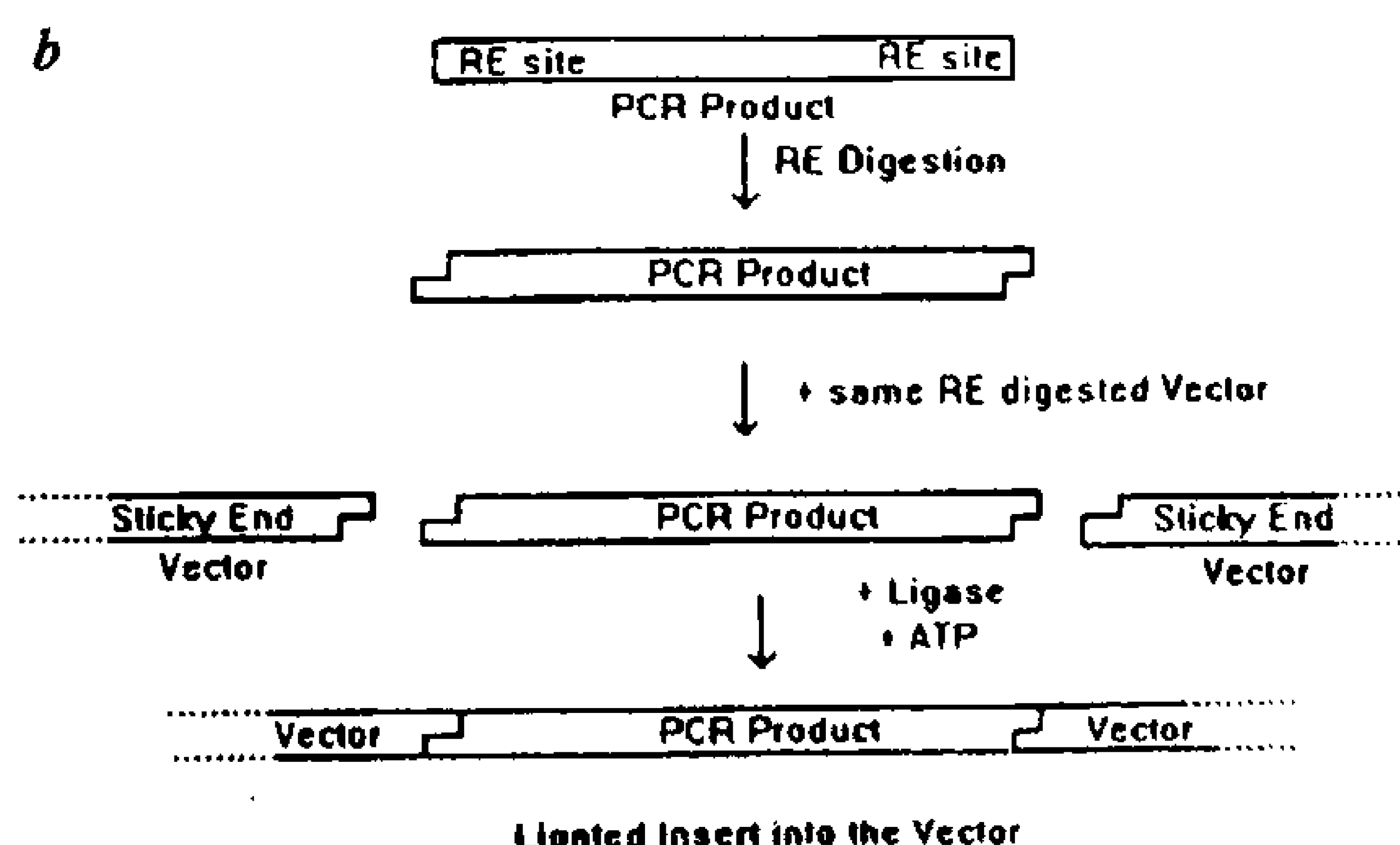


Figure 1. Ligation of PCR products: ***a***, blunt end ligation; ***b***, generation of sticky ends on an insert by primer modification.

This is a multistep process and complete removal of alkaline phosphatase (used in the dephosphorylation step) is sometimes a problem. The residual enzyme reduces ligation efficiency. Overall efficiency in blunt end ligation is far less than sticky end ligation, making this strategy less preferred. Also while cloning PCR product, the need of adding/modifying one or few bases to the vector/insert, may not allow cloning for expression.

In the commercially available cloning system (PCR script), the vector is linearized with a rare 8 base pair cutting enzyme *Srf*I, which remains active in the proprietory ligation buffer. The insert and ligase are added to it. When vector self-ligation occurs during the ligation step, *Srf*I linearizes it again so that the effective linearized vector concentration remains high[2].

Modification of primers

Regarding efficiency, the most successful approach is to incorporate sticky end restriction enzyme site at the 5′ end of PCR primer. The idea is to digest the PCR product and vector with the same restriction endonuclease followed by sticky end ligation of them (Figure 1 *b*).

This strategy has two disadvantages. Firstly, these PCR products having RE site at ends very often cannot be digested to create sticky ends as expected[3,4], defeating the purpose of the exercise. Secondly, if the same RE site is also located within the insert, the fragment gets digested. In case of long primers (18–30 mers), incorporation of the RE site to 5′ end may not change the primer-binding kinetics to template very much; but in techniques like RAPD analysis, where as short as 10 mer primers are used, incorporation of extra bases to create RE site at 5′ end of primers effectively changes the PCR reaction kinetics. So the RE site has to be within the 10 mer itself, to be used as a primer.

A strategy has been proposed by Stoker[5] to avoid the above problem, where, by carefully designing the primer and exploiting the $3' \rightarrow 5'$ exonuclease activity of T4 DNA polymearse, PCR cloning can be done directionally. In this method two long primers having the sequences (CG and CCGG respectively) at their 5′ ends are used. PCR fragments amplified by these primers are incubated with T4 DNA polymerase and only dATP and dTTP but no dCTP and dGTP (Figure 2). T4 DNA polymerase cleaves the 5′ end. A and T are replaced but not C and G. This gives *Xma*I and *Acc*I sites at the two ends of PCR fragment to be ligated. This permits directional cloning without any RE treatment. The same strategy may be used for making other RE sites.

Bhat *et al.*[6] tried a strategy where PCR fragment was incubated with T4 DNA polymerase and T4 polynucleotide kinase together in a universal buffer system with dNTPs and ATP. The insert, thus flush-ended smoothed and phosphorylated was then blunt end ligated with nonphosphorylated *Bam*HI linker (5′CCGGATCCGG3′). In the next step, the whole mixture was diluted so that linker concentration dropped to 1–2 μM; dNTPs and *Taq* polymerase were added and PCR carried out. The DNA fragment thus amplified was then *Bam*HI digested and cloned into vector by sticky end ligation. The major advantage of this technique is that a very small amount of DNA can be cloned.

Vector modification

The popular approach so far for cloning PCR products is by vector modification. Though many thermostable DNA polymerases are available commercially, *Taq* DNA polymerase is the most widely used enzyme. Products amplified with *Taq* DNA polymerase, very often have a 3′ overhang of single nucleotide, which almost always is adenine (A), due to *Taq*'s strong preference for dATP[7,8]. Many researchers have tried to capitalize the presence of this extra A at the 3′ end. The goal is to create an extra T overhang at 3′ end of the vector, now called T vector. When a *Taq* polymerase amplified DNA fragment is ligated to such a vector, the 3′ A of the insert DNA and 5′ T of vector complement each other and ligase catalyses the formation of the phosphodiester bonds. T vectors have an added advantage, due to 3′ T overhang, they do not self-ligate.

Since ligation in T vectors is at least 50 times more efficient than blunt end ligation[9–11], these can also be used for cloning PCR amplified fragments generated by Vent or *Pfu* DNA polymerases which gives blunt-ended products. Here an extra 3′A is added to the DNA fragment by incubating it with dATP and *Taq* DNA polymearse (Figure 3). Also, when a DNA fragment to be cloned has RE sites at its ends which are not present

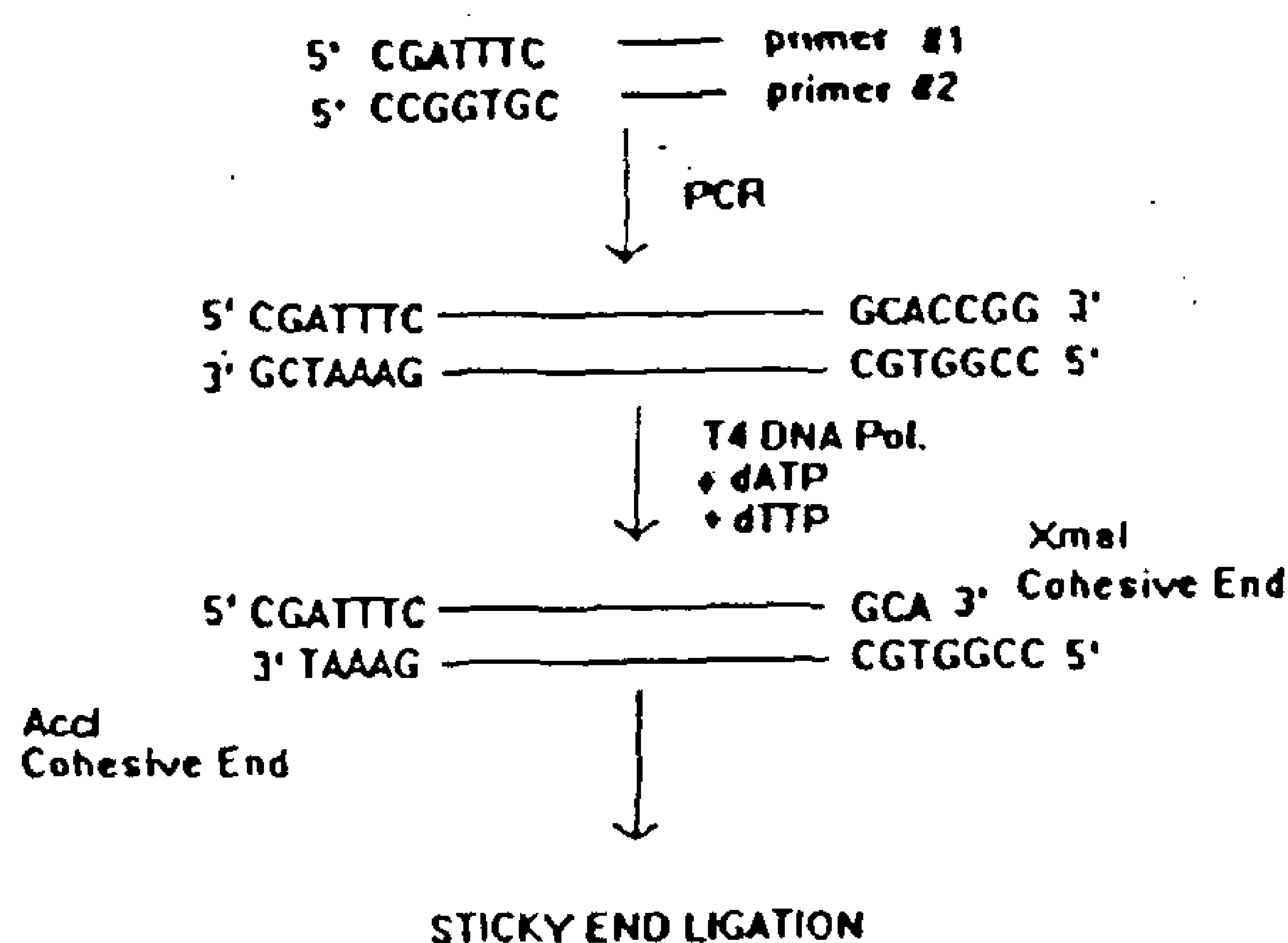


Figure 2. Directional cloning of PCR product by primer modification. Here defined cohesive termini are created by T4 DNA polymerase.

in the vector, the ends are first smoothened by exonuclease activity. An extra A is then added to the 3′ ends to make it compatible with a T vector (Figure 4).

Strategies for making T vectors

For generating a 3′ single T overhang, an *Eco*RV cut vector, having blunt ends can be incubated with *Taq* DNA polymerase and dTTP. *Taq* polymerase will add an extra T to the 3′ ends of the linearized vector because of its property of adding an extra nucleotide independent of template. The vector, with a 3′ T overhang, called T vector, now can then be ligated with PCR product (preferably kinased)[9] (Figure 5). At least two groups of such vectors, having 3′ T overhangs (TA vector, Invitrogen, CA and pGEMT, Promega, USA) are commercially available.

However, The companies supply linearized vectors having 3′ T overhangs ready for ligation. The limitation is not only the cost, but also the linearized vector has to be handled carefully and within a span of six months. Repeated freeze thawing leads to loss of the T overhang and thereby reduced efficiency.

Another method to create a 3′ T overhang is to incubate the smooth-ended vector with terminal transferase and ddTTP[10]. The use of ddTTP ensures addition of only one T residue. A ddT tailed vector can ligate directly to a PCR product with a 3′ A overhang. It is claimed that 90% of the transformants was recombinant and 80% of them had insert of correct size.

While making any type of T overhang, depending on the efficiency of the extra T addition, many vector molecules would not have any T overhang. These would self-ligate to increase the background colonies. An approach for increasing the proportions of effective T vector concentration and reducing background has been suggested by Hadjeb and Berkowitz[12]. After adding an extra T overhang using *Taq* polymerase, to the blunt-ended vector, the whole reaction mixture was subjected to ligation reaction followed by agarose gel purification. The vector molecules without the T overhang self-ligated and migrated differently on gel while the ligated molecules had a different mobility. The linearized vector with 3′ T overhang formed a single band. This band was purified and used for ligation. Increase in cloning efficiency from 5% to 80% by incorporation of this step has been reported.

An innovative approach for making a T-vector capitalizing on the property of a rare restriction enzyme *Xcm*I was first proposed by Mead *et al.*[11] and later improved by Schutte *et al.*[13]. Mead *et al.* synthesized a linker which has two *Xcm*I sites in the middle and *Eco*RI and *Hin*dIII sites at both ends. The linker was designed so that when inserted into the *Eco*RI and *Hin*dIII double-digested vector, it did not disrupt the open reading frame of the LacZ sequence and retained the blue white selection advantage. On *Xcm*I digestion, it released a very short fragment, thereby leaving the vector having two *Xcm*I digested ends with 3′ T overhangs (Figure 6 *a*). The main advantage of this strategy is that T overhang is produced by enzymatic digestion. But even in this strategy, a gel purification step to get rid of the short DNA fragment (36 bp) generated upon digestion of the engineered vector is needed to increase the cloning efficiency. This step not only reduces the yield of 'ready to ligate' vector, but also the left over agarose may reduce the efficiency of ligation.

To avoid this gel purification step, Schutte *et al.*[13] designed a much shorter linker, (Figure 6 *b*) which releases around 20 bp fragment upon digestion. This

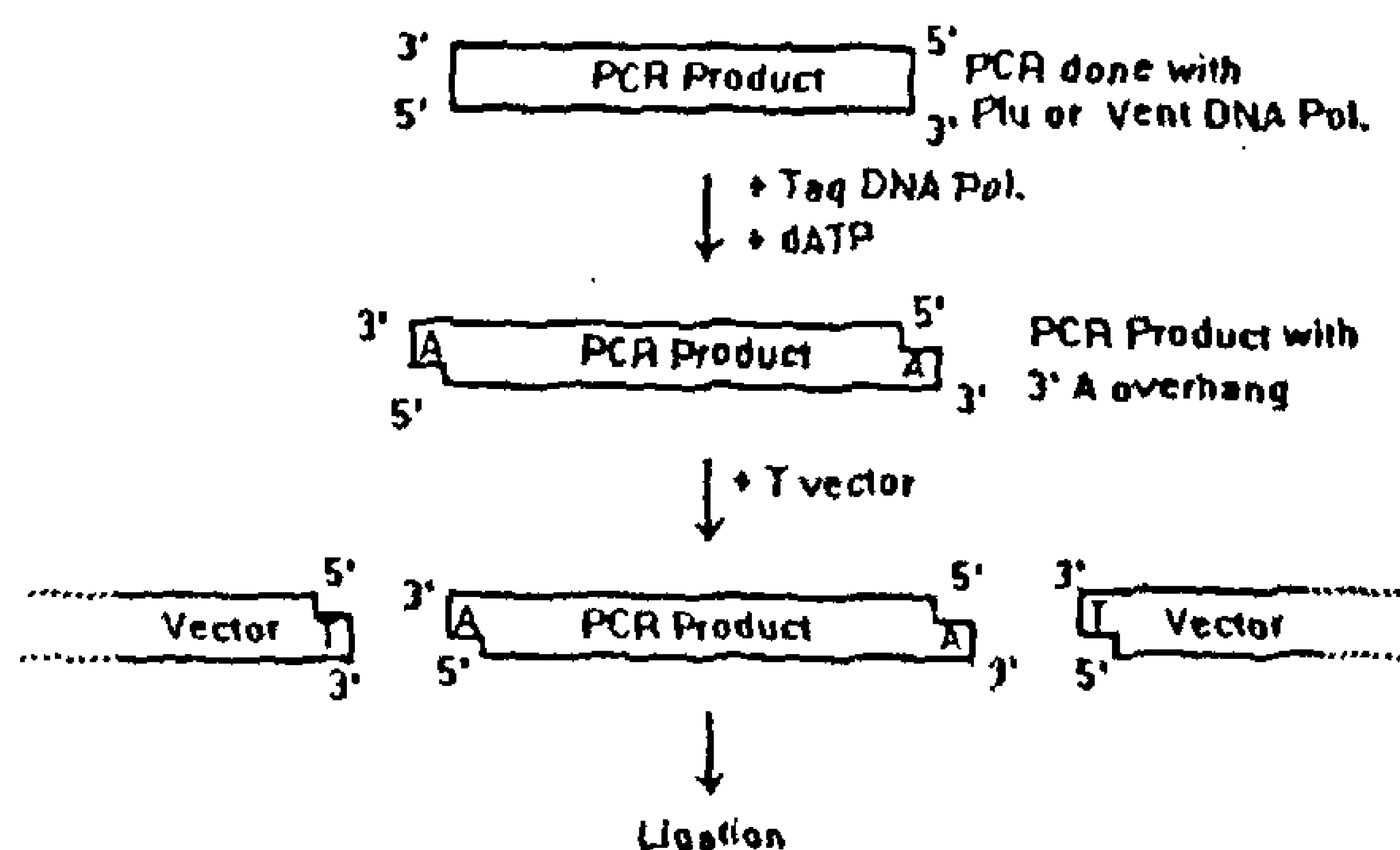


Figure 3. Generating 3′ A overhang in insert for cloning in T vector.

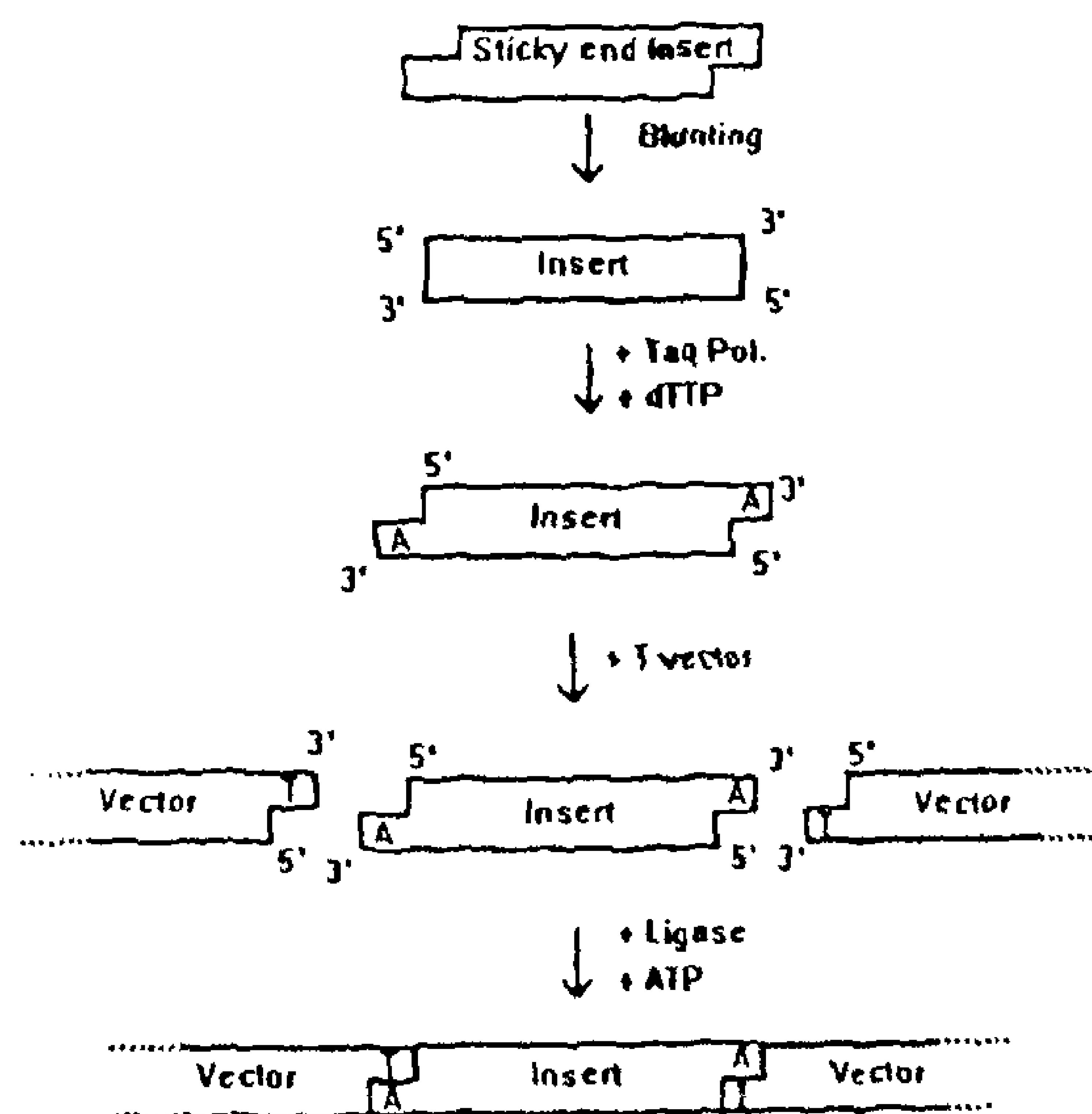


Figure 4. Strategy for making 3′ A overhang in sticky end insert to make it compatible to T vector.

fragment can be removed by using commercially available columns or glass beads. In this strategy, 100% of the vector used for ligation will have a 3′ T overhang, an advantage over other methods. These engineered vectors are available from the ATCC.

Ligase independent methods

Since in most cloning experiments, the efficiency of ligation is a major limiting step, strategies have also been developed to avoid ligase reaction during cloning.

Shuldiner *et al.*[14] first described a PCR-based recombination-dependent strategy for directional cloning. The first step was to amplify the insert DNA with a set of primers whose 5′ ends were complementary to the 3′ ends of the desired linearized plasmid vectors (primer a & b in Figure 7). The PCR product, cleaned from excess primers was divided into two and then added to an equimolar amount of vector in two separate tubes. In one tube primer a & c and in the other, the other primers b & d were added. Both were amplified by PCR. During the first cycle of PCR, some of the 5′ ends of the PCR products annealed to the 3′ ends of the plasmids[15]. Extension resulted in the ligation of the PCR insert into the complementary 3′ end of the plasmid. Subsequent cycles amplified the 'ligated' product in each tube. In next step, amplified products from both tubes were mixed together, denatured with NaOH, neutralized in boiling Tris and incubated for 3–6 h at 60°C. This resulted in complementary heterologous re-annealing followed by cyclization of the recombinant DNA, ready for transformation of host.

It is a fast, easy and reliable strategy without the need of any restriction or modifying enzyme. However, the primers have to be designed carefully so that 5′ ends of the PCR product anneal with 3′ ends of the desired linearized vector in the first PCR step. This approach has its own limitations too, e.g. it requires two plasmid specific primers in addition to the two PCR target primers (which are synthesized by addition of several extra bases at 5′ end to obtain complementary ends to the 3′ ends of the linearized plasmid intended for cloning into). Besides, several steps are necessary before the insert-containing plasmid is ready for bacterial transformation, including amplification of the complete cloning vector in a PCR reaction, which restricts the size of the plasmid to be used.

Garces and Laborda[16] have improved this strategy by making it a single step reaction. They used two 40 mer primers, the 5′ 20 bp of each corresponding to the plasmid sequence immediately adjacent to the cloning restriction site of choice on either side. PCR amplification with these primers generated products whose 5′ ends were complementary to the 3′ ends of the plasmid at both extremes of the cloning site. The product was then denatured and annealed with the same restriction enzyme-digested vector (Figure 8). The annealed product was then subjected to extension reaction. Cyclization of this product gave recombinant DNA ready for transformation of bacteria. This is a single tube reaction using only two primers. Larger size vectors can also be used as efficiently.

To make ligase independent cloning easier a UDG vector[19] has been introduced. Here, at the 5′ end of the PCR primers, instead of adenine (A), uracil (U) is incorporated[17–19]. So, all the PCR amplified products

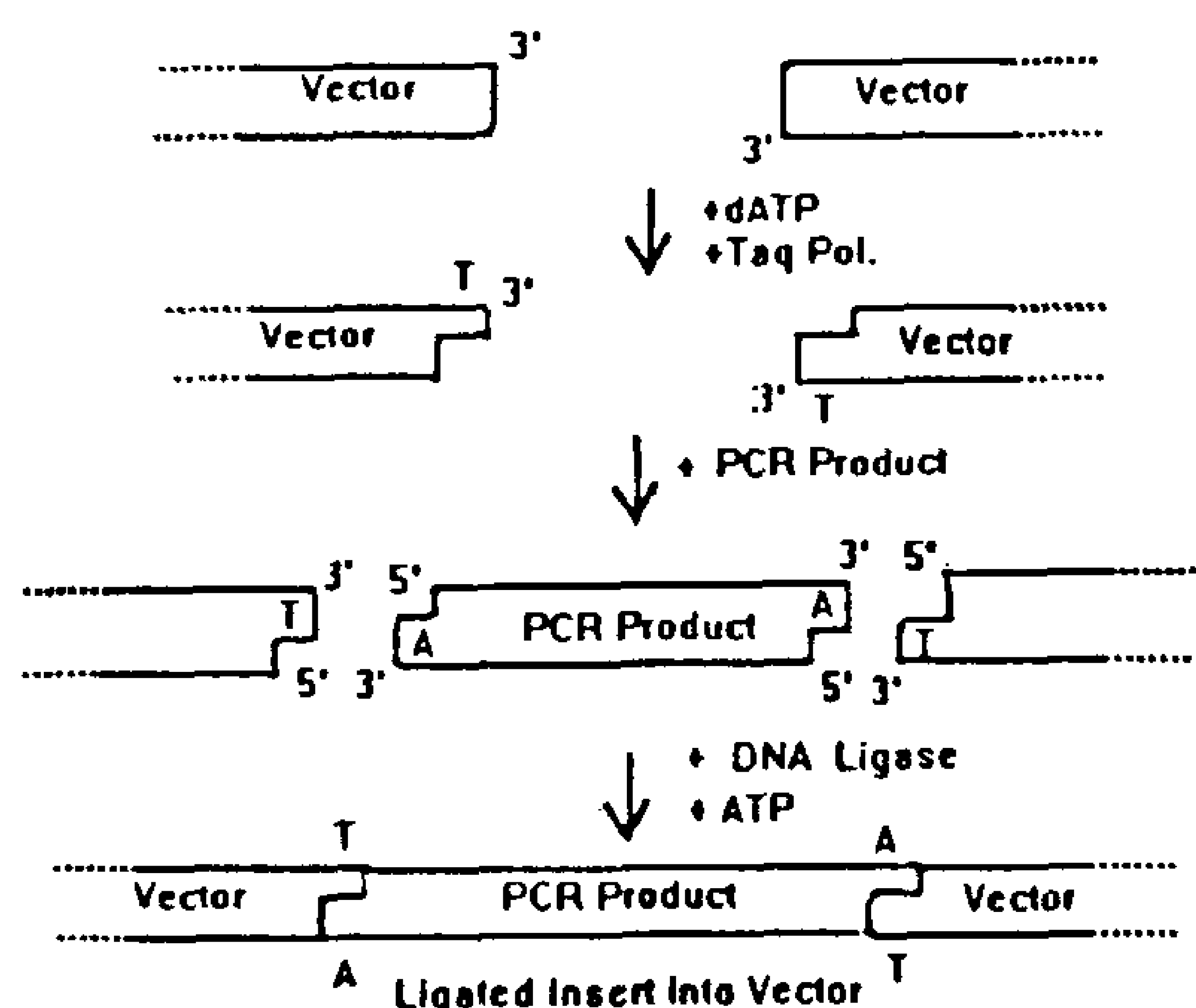


Figure 5. Strategy for making 3′ T overhang by *Taq* DNA polymerase. Adapted from ref. 9.

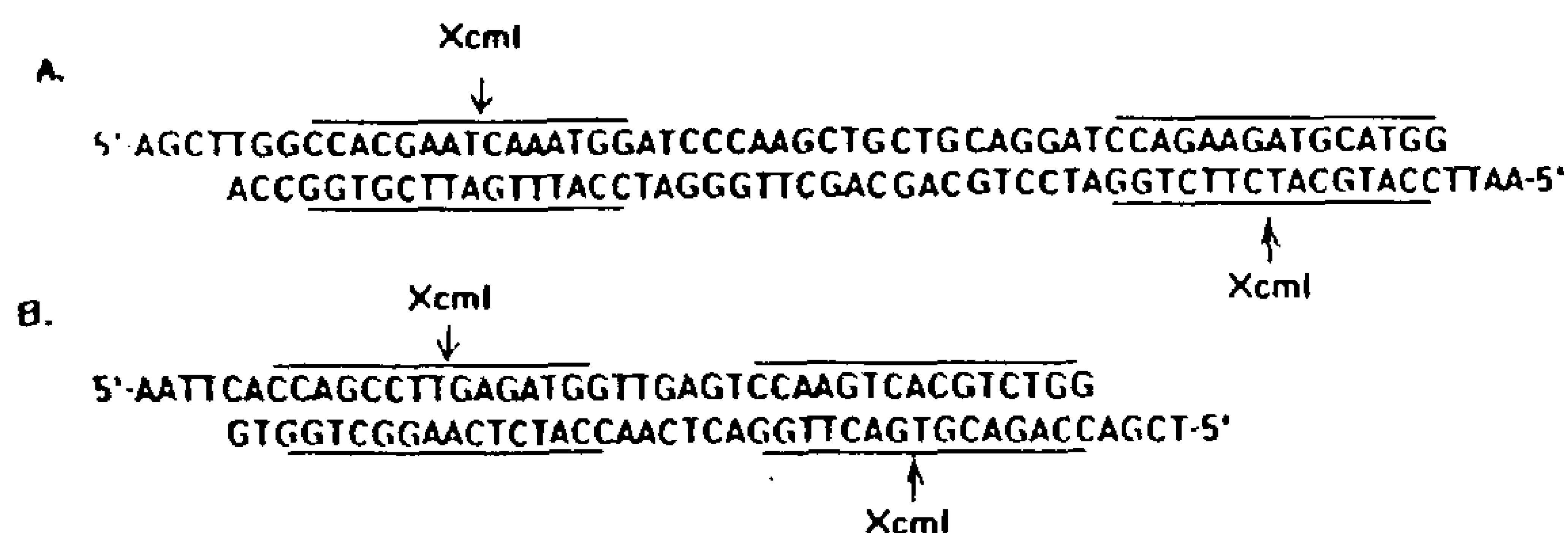


Figure 6. Linkers incorporated into the multiple cloning site of vector to generate T vector by *XcmI* digestion.

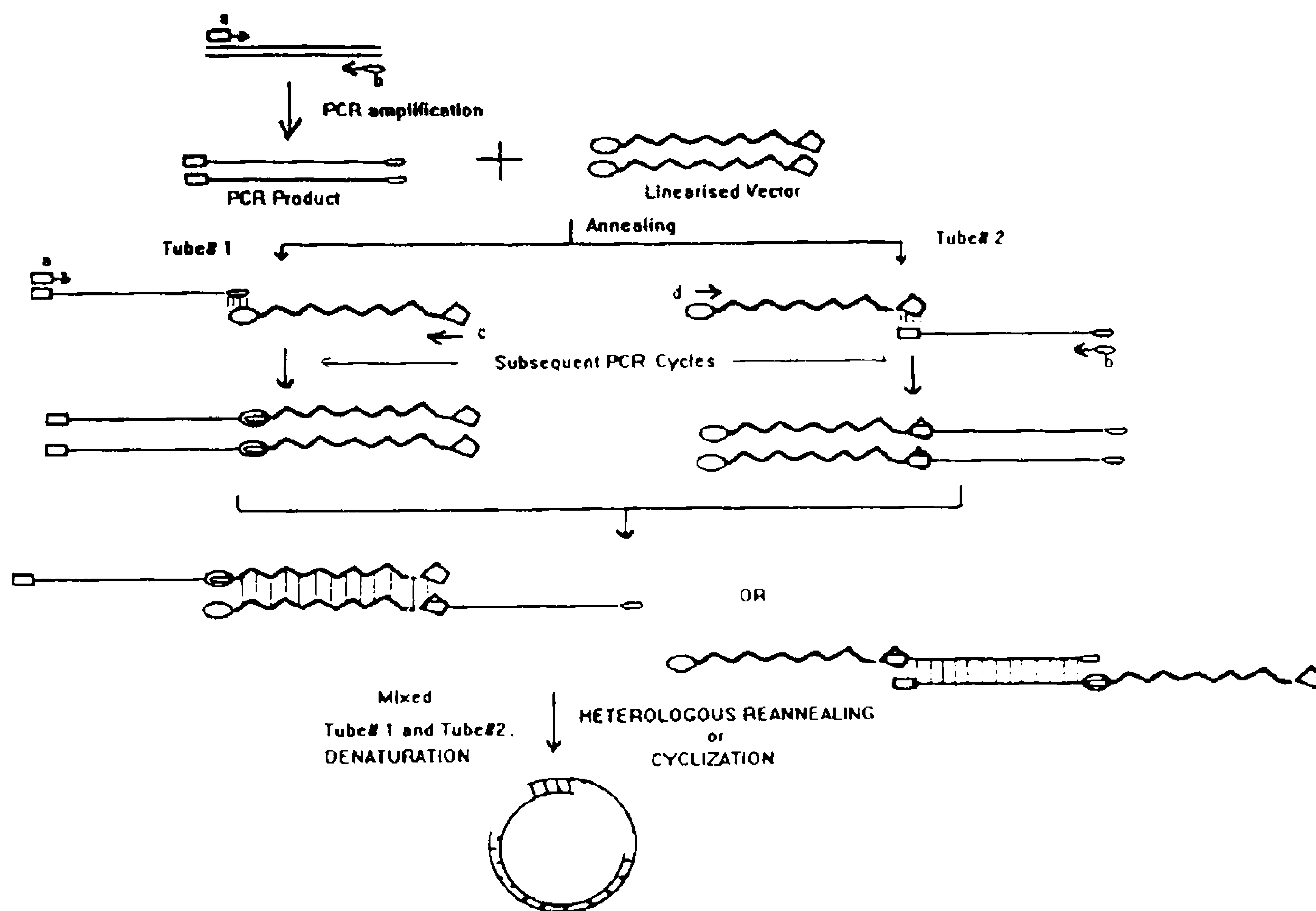


Figure 7. Strategy for ligation independent cloning. Adapted from ref. 14.

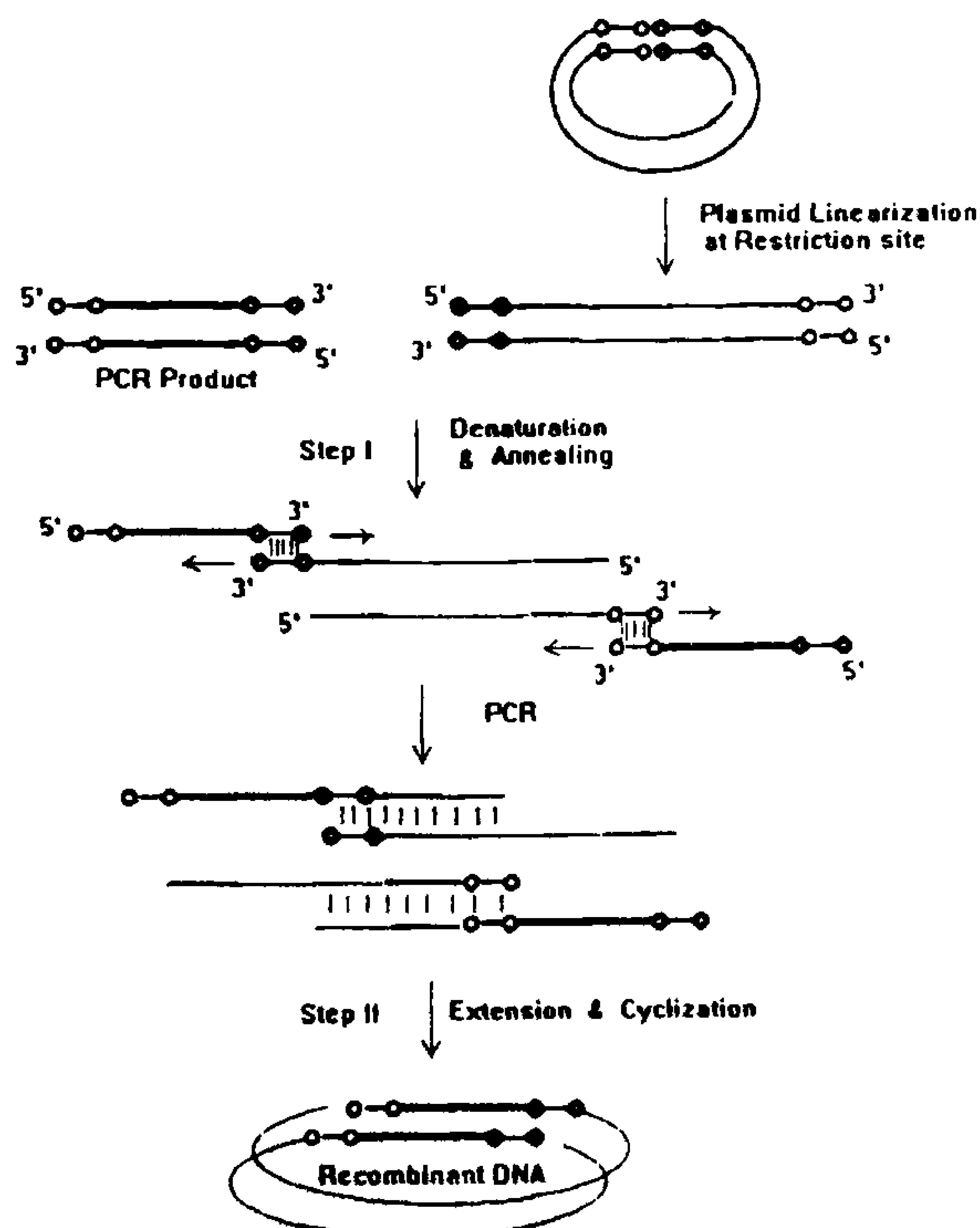


Figure 8. One-step strategy for ligase independent cloning. Adapted from ref. 17.

have U in place of A at their 5′ ends (Figure 9). The PCR product is then cleaved with enzyme uracil deoxyribonucleotide glycosylase (UDG) which cleaves all the U residues disrupting base pairing and thereby exposing the 3′ overhangs. The UDG-treated product is then allowed to anneal to the complementary 3′ ends of UDG cloning vector. This UDG vector has both ampicillin resistance and cloning sites within the *LacZ* gene sequence, making selection of proper transformants easy.

Kumar and Varshney[20] have further improved this strategy of utilizing UDG for cloning. They synthesized primers with one U residue at a key position within last few bases at the 5′ ends. The exact place of incorporating U is shown in Figure 10. When amplified products of these primers were digested with UDG, the U residues were cleaved. As a result the bases 5′ to this U residue could not form stable pair rendering a sticky end for *Sac*I. This product was then ligated to *Sac*I digested cloning vector. This approach has the advantage that (i) only one U is to be incorporated into the primer, which not only avoids a repeat sequence at the both ends of the product but also eliminates the need of synthesizing a long primer. (ii) Since all the commonly used vectors have a *Sac*I site in the multiple cloning site, no specialized vector is needed for cloning. Generation of other restriction sites may also be possible by this strategy.

These ligase independent methods have the advantage

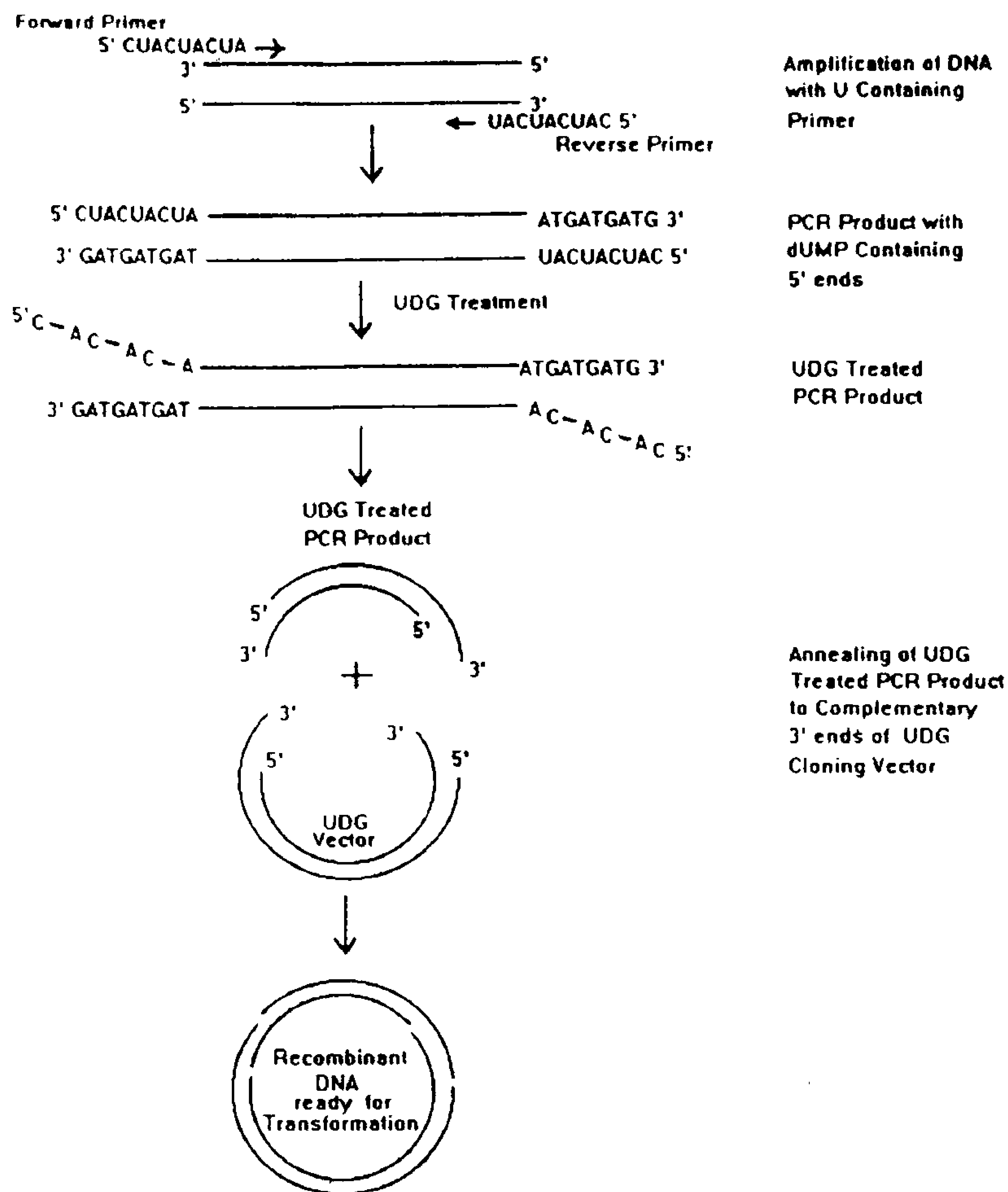


Figure 9. Cloning of PCR product in UDG vector. Adapted from Rashtchian, A., *Curr. Biol.*, 1995, **6**, 30–36.

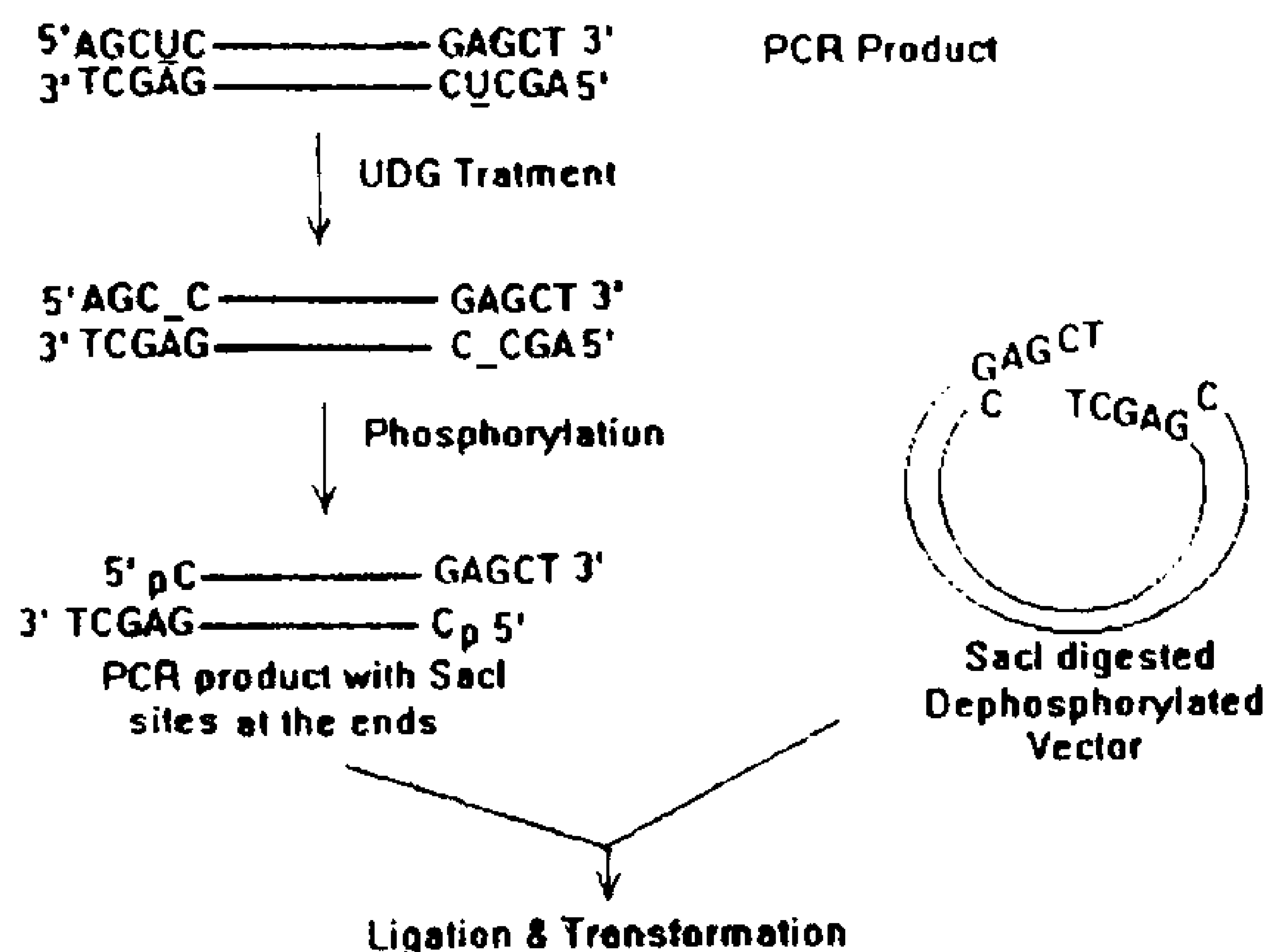


Figure 10. Cloning of PCR product using UDG, by incorporation of a single U residue in the primer. Adapted from ref. 20.

of avoiding the multiple steps like (a) preparation of vector and insert, (b) making suitable changes at the vector/insert ends, (c) ligation reaction, etc. The limitation of this approach is that the primers for every PCR product have to be specifically synthesized.

This set of strategies is by no means exhaustive and the list keeps on increasing. The final choice of strategy will depend on the requirement of the particular laboratory. However, when primers are common for a number of targets to be cloned, ligase independent methods are to be preferred, otherwise ligase-dependent methods are advantageous even after taking into account the various problems and often inefficient nature of ligation and transformation.

1. Papp, T., Kirchner, S., Diener, U., Jafari, M., Golka, A. and Schiffmann, D., *Trends Genet.*, 1995, **11**, 169.
2. *Stratagene Cloning System: Manual*, Stratagene, Hidelberg, 1994, p. 90.

3. *Manual*, New England Biolab, USA, 1995, pp. 208–209.
4. Kaufman, D. L. and Evans G. A., *Biotechniques*, 1990, **9**, 304–306.
5. Stoker, A. W., *Nucleic Acids Res.*, 1990, **18**, 4290.
6. Bhat, G. J., Lodes, M. J., Myler, P. J. and Sruart, K. D., *Nucleic Acids Res.*, 1990, **19**, 398.
7. Clark, J. M., *Nucleic Acids Res.*, 1988, **16**, 9677–9686.
8. Sanchez, A., Bullegos, M., Burgos, M., Jimener, R. and Guardia, R. D., *Trends Genet.*, 1996, **12**, 44.
9. Marchuk, D., Drumm, M., Saulino, A. and Collins, F. S., *Nucleic Acids Res.*, 1990, **19**, 1154.
10. Holton, T. A. and Graham, M. W., *Nucleic Acids Res.*, 1991, **19**, 1156.
11. Mead, D. A., Pey, N. K., Herrnstadt, C., Marcil, R. A. and Smith, L. M., *Nucleic Acids Res.*, 1991, **9**, 657–663.
12. Hadjeb, N. and Berkowitz, G. A., *Biotechniques*, 1996, **20**, 20–21.
13. Schutte, B. C., Ranade, K., Pruessner, J. and Dracopoli, N., *Biotechniques*, 1997, **22**, 40–44.
14. Shuldiner, A. R., Scott, L. A., Roth, J., *Nucleic Acids Res.*, 1990, **15**, 1920.
15. Higuchi, R., Krummel, B. and Saiki, R. K., *Nucleic Acids Res.*, 1988, **16**, 7351–7367.
16. Garces, C. and Laborda, J., *Anal Biochem.*, 1995, **230**, 178–180.
17. Aslanidis, C. and de Jong., *Nucleic Acids Res.*, 1991, **18**, 6069–6074.
18. Buchman, G. W., Schuster, D. M, Rashtchian, A., *Focus*, 1992, **14**, 41–45.
19. Rashtchian, A., Buchman G. W., Schuster, D. M. and Berninger M. S., *Anal. Biochem.*, 1992, **206**, 91–97.
20. Kumar, N. V. and Varshney, U., *Curr. Sci.*, 1994, **67**, 728–734.

Received 14 July 1997; revised accepted 6 October 1997

A three-dimensionally modulated structure in a chiral smectic-C liquid crystal

P. A. Pramod, R. Pratibha and N. V. Madhusudana

Raman Research Institute, C. V. Raman Avenue, Bangalore 560 080, India

In this article we report the discovery of a new twist grain boundary phase. This phase is characterized by a 2-dimensional undulation of the smectic-C*-like blocks in the form of a square lattice. We suggest that this three-dimensionally modulated structure, which was not anticipated by theory, owes its origin to chiral interactions.

THE formal analogy between superconductors and smectic liquid crystals was invoked by De Gennes[1] to predict the possibility of an intermediate phase with a lattice of dislocations in smectics. Goodby *et al.*[2] discovered such a structure in a highly chiral liquid crystal. This twist grain boundary (TGB_A) phase consists of a helical stack of blocks of smectic-A (S_A) liquid crystal, separated by grain boundaries made of an array of screw dislocations (Figure 1 *a*), in accordance with a structure which had been worked out by Renn and Lubensky[3]. Unlike superconductors, smectic liquid crystals can have other modifications like the smectic-C (S_C) in which the molecules are tilted with respect to the layer normal and the smectic-C* (S_{C^*}) in which the tilt direction has a helical arrangement about the layer normal. Although TGB phases with S_C-like blocks (TGB_C) (Figure 1 *b*) and S_{C^*}-like blocks (TGB_{C^*}) have been theoretically predicted[4,5], only the TGB_C phase has been experimentally characterized in some detail[6,7]. Liquid crystals are rather soft, and can exhibit novel geometrical structures. We have found in a binary mixture a new TGB phase which has a 2D undulation of the S_{C^*} blocks in the form of a square lattice.

Experimental

The new phase was found in binary mixtures of the chiral compound 4-(2′-methyl butyl phenyl 4′-*n*-octyl biphenyl-4-carboxylate (CE8)) and 2-cyano-4-heptyl-phenyl-4′-pentyl-4-biphenyl carboxylate (7(CN)5) which have very similar lengths and molecular structures. On heating, CE8 exhibits the phase sequence (with temperature in °C): crystal 67 S_{I^*} 70 S_{C^*} 85 S_A 134.6 N^* 140.5 I, where N^* stands for chiral nematic and I for isotropic phases. On the other hand, 7(CN)5 has a wide nematic range (crystal 45 N 102 I) and X-ray studies have shown that it has a strong skew cybotactic (S_C-like) short range order[8]. The phase diagram of the binary mixtures is shown in Figure 2. The TGB phases are found only in mixtures with ~5 wt% to 45 wt% of 7(CN)5. Most of the physical studies were conducted on a mixture with about 36 wt% of 7(CN)5, which