The EMBO Journal Vol.2 No.10 pp.1791 – 1794, 1983

Easy identification of cDNA clones

U. Rüther* and B. Müller-Hill

Institut für Genetik, Universität zu Köln, 5000 Köln 41, FRG

Communicated by B. Müller-Hill
Received on 27 June 1983

A set of six cloning vectors, pUR 278, 288, 289, 290, 291, 292 is presented. These vectors have the cloning sites, *Bam*HI, *Sal*I, *Pst*I, *Xba*I and *Hind*III, in all frames at the 3′ end of the *lacZ* gene. Insertion of cDNA in the proper cloning sites leads to a fusion protein of active β-galactosidase and the peptide encoded by the cDNA. A simple immunoenzymatic assay can be used to identify clones in such a cDNA library.
Key words: cDNA cloning/β-galactosidase fusions/immunoenzymatic identification

Introduction

The first step in the analysis of a eukaryotic gene is usually the synthesis and cloning of cDNA after mRNA isolation (Williams, 1981). Such cDNA may be used for the determination of the DNA sequence and for the identification of the corresponding gene. However, for genes which are expressed at a low level, the isolation of cDNA clones to be used as probes is time consuming or even not feasible.

Methods have been described to overcome this problem. Most commonly, synthetic oligonucleotides are used as hybridization probes. This approach requires knowledge of the amino acid sequence of at least part of the gene product (Suggs *et al.*, 1981). Immunoassays have been developed which permit the detection of specific clones to obviate this limitation (Broome and Gilbert, 1978; Kemp and Cowman, 1981; Koenen *et al.*, 1982). They require the expression of the particular genes in *Escherichia coli* and their subsequent identification with antibodies.

In one of the methods (Koenen *et al.*, 1982) gene fragments are cloned in *E. coli* in such a manner that their protein products are fused and stabilized as NH_2-terminal peptides to β-galactosidase. Such chimeras can be bound by antibodies (Broome and Gilbert, 1978). The immunocomplex can be visualized by using the β-galactosidase activity to hydrolyse its substrate 5-bromo-4-chloro-indolyl-β-galactoside (x-gal) to a blue insoluble product. This method has been used successfully for the identification of exons coding for chicken lysozyme (Rüther *et al.*, 1982). However, this method cannot be used for the identification of cDNA clones since there is always a stop codon at the end of the coding part of the cDNA which makes it impossible to obtain fusion proteins.

New cloning vectors have been constructed to overcome this restriction. Different cloning sites are introduced at the 3′ end of the *lacZ* gene to allow the expression of fusion proteins of active β-galactosidase and the cDNA encoded peptide in *E. coli*. These fusion proteins can be used to identify with antibodies the colonies expressing the corresponding cDNA (Koenen *et al.*, 1982). These new vectors can also be used for the expression of cloned cDNA similar to experiments described (Itakura *et al.*, 1977; Shine *et al.*, 1980; Charnay *et al.*, 1980; Davis *et al.*, 1981). The purification of fusion proteins is easy because of their β-galactosidase activity. Thus, the purification steps can be controlled by a simple colorimetric test (Miller, 1972).

*To whom reprint requests should be sent.

Results

Construction of cloning sites at the 3′ end of the lacZ gene

The *Eco*RI-*Hae*III DNA fragment coding for the C-terminal 17 amino acids of the β-galactosidase and part of the intracistronic region between *lacZ* and *Y* (Büchel *et al.*, 1980) was cloned in to the *Hinc*II-*Eco*RI sites of plasmid pUR 250 (Rüther, 1982). This plasmid, pUR 259, was digested with *Hind*III and treated with *Bal*31 as described in Materials and methods. Then the plasmid DNA was digested with *Eco*RI and the manipulated *lacZ* gene fragments were cloned into the *Sma*I-*Eco*I sites of plasmid pUC 9 (Vieira and Messing, 1982). The lengths of the inserts from 35 recombinant clones were tested after *Hind*III-*Eco*RI digestion on gels and inserts of the desired length were sequenced.

The clones (pUR 266/1 – 3) having the deletion end points

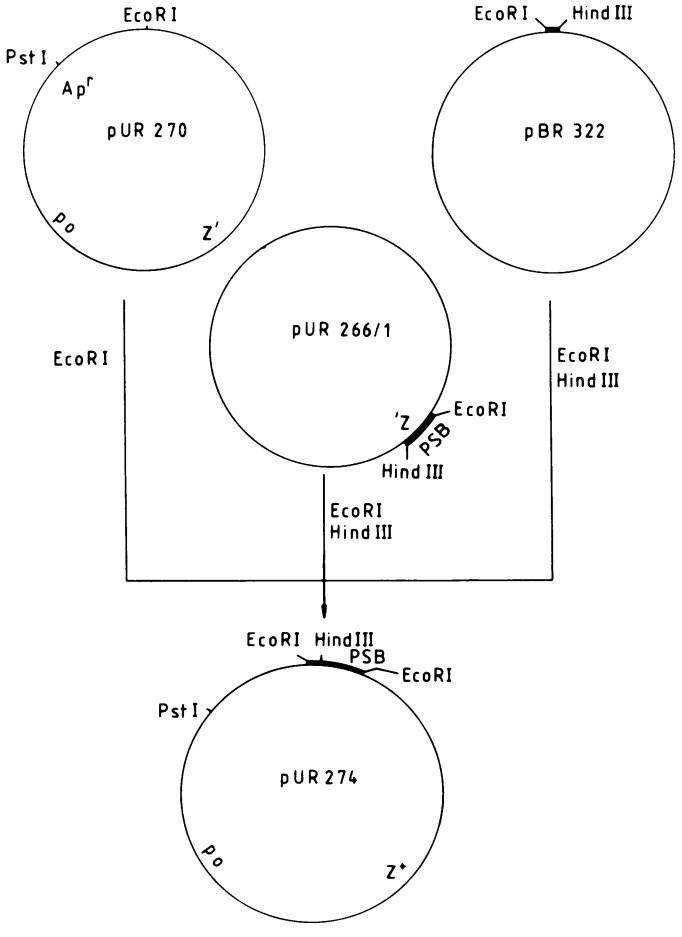

Fig. 1. Construction of plasmids pUR 274, 276, 286. As an example, the construction of pUR 274 is shown. Abbreviations: P, *Pst*I; S, *Sal*I; B, *Bam*HI.

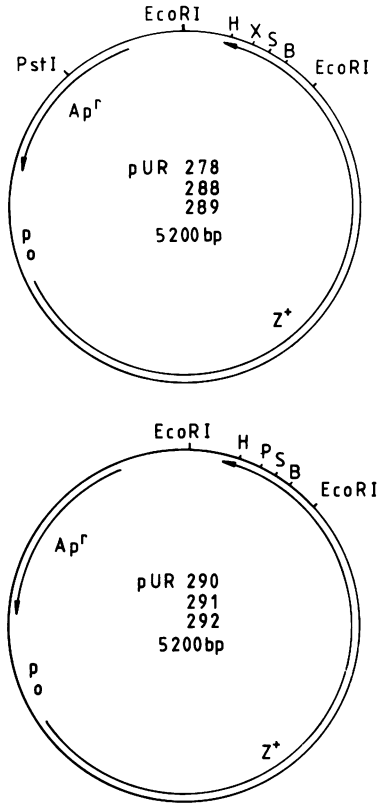

Fig. 2. Restriction map of plamsids pUR 278, 288, 289, 290, 291, 292. The differences in the 3′ end of the *lacZ* gene are shown in Figure 3. Abbreviations: H, *Hind*III; X, *Xba*I; S, *Sal*I; P, *Pst*I; B, *Bam*HI.

in different frames near the end of the *lacZ* gene were used for further constructions.

Construction of cloning vectors

Plasmid pUR 270 (see Figure 1), a variant of plasmid pUK 217 (Rüther *et al.*, 1981) carrying only the part of the *lac* operon up to the *Eco*RI site at the 3′ end of the *lacZ* gene was used for the reinsertion of the manipulated *lacZ* 3′ ends to test whether they are lac$^+$ or not. For this purpose pUR 270 was digested with *Eco*RI, while plasmids pUR 266/1−3 and pBR 322 were digested with *Hind*III and *Eco*RI. The DNAs were mixed, ligated and transformed into *F′* 11 *recA*, a *lac*$^-$ strain. On x-gal, isopropyl thiogalactoside (IPTG), ampicillin plates, blue colonies were screened. From these plasmids, DNA was isolated and analysed. All tested plasmid DNA had the expected structure as shown in Figure 1. All manipulated 3′ *lacZ* gene ends tested restore the β-galactosidase activity when introduced on plasmid pUR 270. The three new plasmids were called pUR 274, 276, 286. The *Pst*I site in the cloning region of these plasmids is not unique, since a *Pst*I site is also present in the ampicillin resistance gene (*Ap*^{r}). Therefore the cloning regions were exchanged against the cloning region of pUR 250 (Rüther, 1982) via *Bam*HI and *Hind*III sites. These three new plasmids, pUR 278, 288, 289 (Figure 2) have four unique cloning sites at the 3′ end of the *lacZ* gene in all frames (Figure 3).

Construction of cloning vectors with a unique PstI site

Because it is advantageous to use the *Pst*I site for cloning by the GC-tailing method (Roychoudhury *et al.*, 1980), cloning vectors with a unique *Pst*I site have also been constructed. The *Pst*I site in the ampicillin resistance gene (*Ap*^{r}) of plasmid

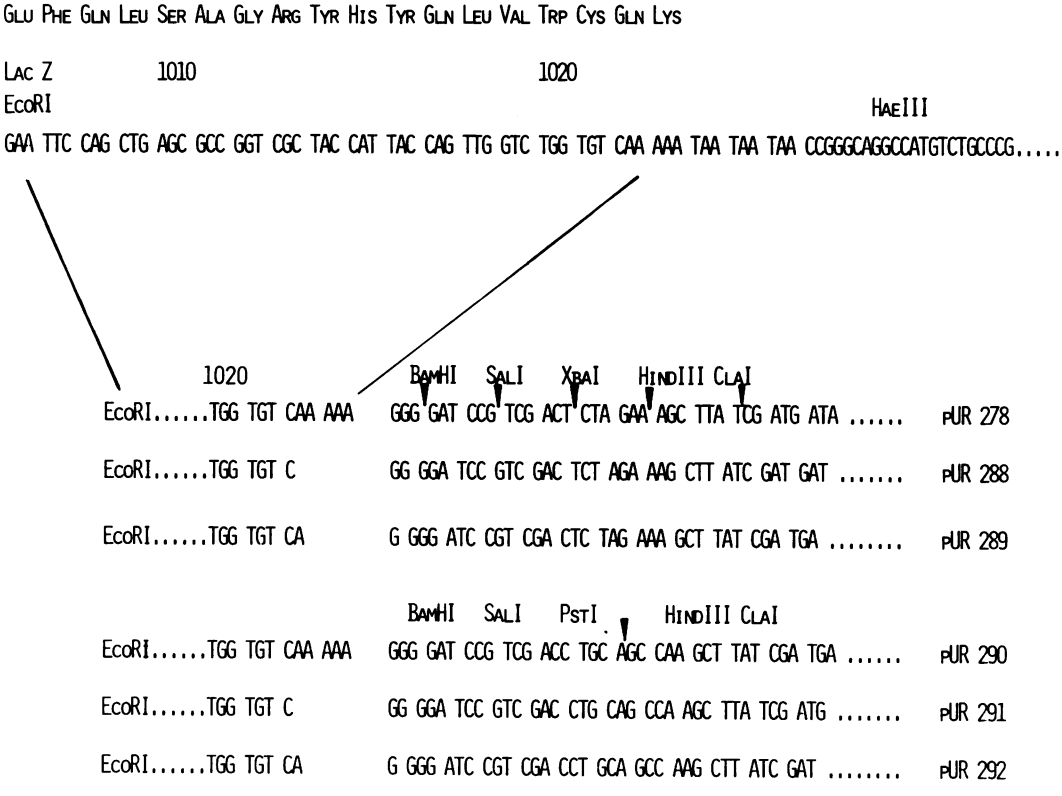

Fig. 3. The cloning region of the plasmids pUR 278, 288, 289, 290, 291, 292 at the 3′ end of the *lacZ* gene. The upper sequences are the wild-type DNA and protein sequences of the 3′ end of the *lacZ* gene. The lower sequences show the results of the subcloned *Bal*31-treated *lacZ* gene ends. Now five unique cloning sites (*Bam*HI, *Sal*I, *Xba*I or *Pst*I and *Hind*III) exist in all frames at the 3′ end of an active *lacZ* gene.

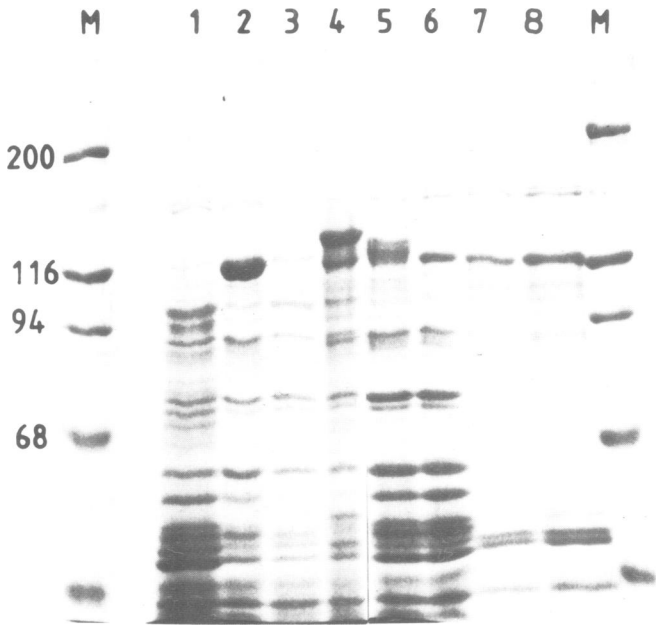

Fig. 4. Fusions of β-galactosidase with lysozyme, α- and β-casein. A Coomassie blue-stained SDS gel (7.5%) is shown. Crude extract from ~0.5 ml overnight culture is subjected to electrophoresis. Lane M shows the marker proteins myosin, β-galactosidase, phosphorylase A and bovine serum albumin (in kd). If not noted otherwise, only the cytoplasmic fraction was put on the gel and the cultures were grown in the presence of 5×10^{-4} M IPTG. **Lane 1:** *F'* 11 *recA* pUR 289 without IPTG, **lane 2:** *F'* 11 *recA* pUR 289, **lane 3:** *F'* 11 *recA* with lysozyme cDNA, **lane 4:** same as **lane 3** but membrane fraction, **lane 5:** *F'* 11 *recA* pUR 291 with α-casein cDNA, **lane 6:** *F'* 11 *recA* pUR 292 with β-casein cDNA, **lane 7:** precipitate of *F'* 11 *recA* pUR 292 (β-casein cDNA) with antibodies against milk protein, **lane 8:** precipitate of *F'* 11 *recA* pUR 292 (β-casein cDNA) with antibodies against β-galactosidase.

pUR 278 was removed by NG-mutagenesis as described by Rüther *et al.* (1981). The cloning region with the 3' end of the *lacZ* gene was then removed by *Eco*RI digestion. The resulting plasmid, pUR 280, is very similar to pUR 270, the only difference being the absence of the *Pst*I site in Ap^r. The cloning regions of pUR 274, 276 and 286 were cloned into pUR 280 and the new plasmids pUR 290, 291 and 292 were obtained (Figures 2 and 3).

Cloning and identification of different cDNA molecules

As a model, three characterized cDNA molecules were used, coding for the genes of chicken lysozyme, mouse α- and β-casein (Sippel *et al.*, 1978, Hennighausen and Sippel, 1982).

The lysozyme cDNA was cloned as a *Bam*HI fragment into pUR 289. The α- and β-casein cDNAs were cloned as *Hind*III fragments into a mixture of pUR290, 291, 292 since the frame of these cDNAs was not known. After transformation into *E. coli* BMH 71-18, the colonies were screened as described (Koenen *et al.*, 1982) with antibodies against chicken lysozyme or against mouse milk protein. Several positive clones could be detected among several hundred colonies by this method. The DNA analysis confirmed the predicted construction. To demonstrate the fusion proteins, lysates of these colonies were subjected to SDS-polyacrylamide gel electrophoresis (Laemmli, 1970). The presence of β-casein sequences in one of the fusion proteins was demonstrated by antibody precipitation (Figure 4).

Discussion

The newly developed vectors, pUR 278, 288–292, can be used for the detection of cDNA clones from a library. The method involves binding of products of the cDNA clones with specific antibodies to a solid phase. The immune complexes are identified by the enzymatic activity of the fusion protein. Therefore, a second radioactive antibody (Young and Davis, 1983) or protein A are not required. These new vectors can also be used for the cloning of cDNA with synthetic DNA linkers or dGC tailing (Roychoudhury and Wu, 1980) after classical cDNA synthesis. The combination of cloning sites allows the use of the method of Okayama and Berg (1982) in which vector DNA primes the cDNA synthesis of both strands. Because the orientation of the cDNA is determined by its synthesis, one of the three clones should use the correct frame of the insert and therefore express a fusion protein. Another advantage of the new vectors is the high level of expression of the cloned DNA. As shown in Results (Figure 4), the fusion protein is very abundant.

The amount and the stability of the fusion protein is the same or better when compared with the fusion proteins produced from DNA cloned into the *Eco*RI site at position 1006 of the *lacZ* gene (Itakura *et al.*, 1977; Shine *et al.*, 1980; Charnay *et al.*, 1980; Davis *et al.*, 1981; Young and Davis, 1983). The β-galactosidase-α-casein fusion and the β-galactosidase-lysozyme fusion, which is only found in the membrane fraction, are partially degraded to wild-type size β-galactosidase (Figure 4). Perhaps there is a specific proteolytic cleavage of the latter chimera after its secretion into the membrane since part of the signal sequence of the lysozyme is still present in the cDNA sequence (Sippel *et al.*, 1978).

Because in the vectors all cloning sites exist in all frames, any cDNA should be expressed in these vectors. The resulting fusion proteins should be easy to purify since the active β-galactosidase part can be purified by affinity chromatography on a column carrying covalently bound substrate analogues (Steers *et al.*, 1971; Ullman, personal communication).

The inserted cDNA itself can be analysed by restriction or sequence analysis. It can be isolated through its flanking restriction sites to be used as a hybridization probe. The combination of the sites allows direct sequencing of the inserts as described by Rüther *et al.* (1981) and Rüther (1982). Finally, the insert can be isolated in large amounts by the use of *Eco*RI alone, one of the cheapest restriction enzymes.

Materials and methods

E.. coli K12 strains F' 11 *recA* and BMH 71-18 are described by Rüther *et al.* (1981) and Messing *et al.* (1977), plasmid pUR 250, pUC 9 and pBR322 in Rüther (1982), Vieira and Messing (1982) and Bolivar *et al.* (1977).

Restriction enzymes were purchased from BRL, Biolabs or Boehringer. *Bal*31 was obtained from Biolabs.

*Bal*31 reaction. 7.5 pmol ends of plasmid DNA were incubated in 30 μl buffer (600 mM NaCl, 12 mM $MgCl_2$, 20 mM Tris-HCl pH 8.0) with 0.1 units of *Bal*31 for 10 min at 37°C. The reaction was stopped by addition of 30 μl H_2O and 150 μl ethanol. The DNA was precipitated after chilling.

Acknowledgements

We thank Lothar Hennighausen for providing α- and β-casein cDNA and antibodies against milk proteins. This work was supported by a grant of the Deutsche Forschungsgemeinschaft (Schwerpunkt Gentechnologie).

References

Bolivar,F., Rodriguez,R.L., Greene,P.J., Betlach,M.C., Heynecker,M.L., Boyer,H.W., Crosa,J.M. and Falkow,S. (1977) *Gene,* **2**, 95-113.

Broome,S. and Gilbert,W. (1978) *Proc. Natl. Acad. Sci. USA,* **75**, 2746-2749.

Büchel,D.E., Gronenborn,B. and Müller-Hill,B. (1980) *Nature,* **283**, 541-545.

Charnay,P., Gervais,M., Louise,A., Galibert,F. and Tiollais,P. (1980) *Nature,* **286**, 893-895.

Davis,A.R., Nayak,D.P., Keda,M., Hiti,A.L., Dowbenko,D. and Kleid, D.G. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 5376-5380.

Hennighausen,L. and Sippel,A.E. (1982) *Eur. J. Biochem.,* **125**, 131-141.

Itakura,K., Hirose,T., Crea,R., Riggs,A.D., Heynecker,H., Bolivar,F. and Boyer,H.W. (1977) *Science (Wash.),* **198**, 1056-1063.

Kemp,D.J. and Cowman,A.F. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 4520-4524.

Koenen,M., Rüther,U. and Müller-Hill,B. (1982) *EMBO J.,* **1**, 509-512.

Laemmli,U.K. (1970) *Nature,* **227**, 680-685.

Messing,J., Gronenborn,B., Müller-Hill,B. and Hofschneider,H.-P. (1977) *Proc. Natl. Acad. Sci. USA,* **74**, 3642-3646.

Miller,J.M. (1972) *Experiments in Molecular Genetics,* published by Cold Spring Harbor Laboratory Press, NY.

Okayama,M. and Berg,P. (1982) *Mol. Cell Biol.,* **2**, 161.

Roychoudhury,R. and Wu,R. (1980) *Methods Enzymol.,* **65**, 43-62.

Rüther,U. (1982) *Nucleic Acids Res.,* **10**, 5765-5772.

Rüther,U., Koenen,M., Otto,K. and Müller-Hill,B. (1981) *Nucleic Acids Res.,* **9**, 4087-4098.

Rüther,U., Koenen,M., Sippel,A.E. and Müller-Hill,B. (1982) *Proc. Natl. Acad. Sci. USA,* **79**, 6852-6855.

Shine,J., Fettes,J., Lam,N.C.Y., Roberts,J.L. and Baxter,J.D. (1980) *Nature,* **285**, 456-461.

Sippel,A.E., Land,H., Lindenmaier,W., Nguyen-Huu,M.C., Wurtz,T., Timmis,K.N., Giesecke,K. and Schütz,G. (1978) *Nucleic Acids Res.,* **5**, 3275-3294.

Steers,E., Jr., Cuatrecasas,P. and Pollard,H.B. (1971) *J. Biol. Chem.,* **246**, 196-200.

Suggs,S.V., Wallace,R.B., Hirose,T., Kawashima,E.H. and Itakura,K. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 6613-6617.

Vieira,J. and Messing,J. (1982) *Gene,* **19**, 259-268.

Williams,J.G. (1981) in Williamson,R. (ed.), *Genetic Engineering,* Vol. **1**, Academic Press, London, pp. 1-55.

Young,R.A. and Davis,R. (1983) *Proc. Natl. Acad. Sci. USA,* **80**, 1194-1198.