Regular

International Journal of Integrative Biology
A journal for biology beyond borders

IJIB
ISSN 0973-8363

A method to introduce RE recognition site within coding region of a gene without altering amino acid frame continuity

A Nagee [1], PN Mukhopadhyaya [2,*], Arpan Acharya [2], BA Aich [2], IL Kothari [3]

[1] ARIBAS, *Anand, India*
[2] *Medical Genetics Division, geneOmbio Technologies, Pune, India*
[3] *Department of Biosciences, Sardar Patel University, Vallabh Vidyanagar, Anand, India*

Submitted: 23 Jan. 2008; Accepted: 18 Mar. 2008

Abstract

cDNA synthesized from a filamentous fungus was used as template to perform 3' RACE to incorporate *Hin*d III restriction endonuclease (RE) recognition sequence within the multiple cloning site (MCS) of pGEM T-Easy plasmid DNA molecule. The method involved use of a specifically-designed set of oligonucleotide primer pairs for first and second strand cDNA synthesis which had a *Hin*d III site at the 5' end of the forward primer along with a critical nucleotide strategically placed upstream of it for maintaining amino acid frame continuity. The method relies on TA cloning technology and takes into account the additional bases that are incorporated by template-independent incorporation of a nucleotide base (Adenine) by *Ta*q DNA Polymerase during PCR. The methodology can be used to insert any short restriction endonuclease recognition sequence within the MCS to upgrade multiple cloning site of existing plasmid DNA cloning vector molecules.

Keywords: vector, restriction endonuclease, multiple cloning sites, molecular cloning.

INTRODUCTION

Cloning vectors assumed importance since the advent of recombinant DNA technology in biological research. The first cloning vector developed was called pBR322 (Bolivar, 1977; Sutcliffe, 1979; Watson, 1988; Sambrook *et al*., 1989), which latter paved way for a vast array of other vectors with different attributes and functions.

One of the differences between pBR322 and other new generation cloning plasmids is the mode of selection of the insert. While pBR322 relies on insertion inactivation of antibiotic resistant genes for selection, other vectors are available that provide the convenience of blue-white color selection for this purpose (Yanisch-Perron *et al*., 1985; Lin-Chao *et al*., 1992). The latter have clear advantage over the former because of the ease in screening bacterial population for recombinant plasmids. Another important attribute of cloning vectors is the number of unique restriction recognition sites available for inserting foreign DNA. For pBR322 vector and its derivatives, these sites should be internal to the antibiotic genes such that selection through insertion inactivation can be performed (Balbas *et al.*,1986). For vectors providing the facility for blue-white selection, such sites should be internal within the open reading frame of a functional gene (Chaffin *et al.*, 1998).

Discovery of new restriction endonucleases (REs) is a continuous phenomenon. However, it leads to the need for incorporation of new recognition sites within existing vectors to upgrade them and make them compatible with novel enzymes. This work presents a method for incorporation of new RE recognition sequence within the multiple cloning site of a multi copy-cloning vector that retains the feature of blue-white selection for

[*] ***Corresponding author:***
Pratap Narayan Mukhopadhyaya, Ph.D.
geneOmbio Technologies, 4th Floor,
Krishna Chambers,, Pashan-Sus Road, Pashan
Pune-411 021, Maharashtra, INDIA
Email: pnm6619@yahoo.co.in

screening recombinants exploiting a eukaryotic cDNA as template.

MATERIALS AND METHODS

Growth of fungal mycelium

Arthobotrys oviformis was grown in duplicate in liquid CMA (Corn Meal Agar) medium using spore suspension as inoculum. The culture was grown at 28^0C temperature under shaking condition (200 revolutions per minute) for four days followed by isolation of total RNA.

Extraction of total RNA

Total RNA was extracted from freshly grown fungal mycelium (100 mg) with RNAzol (Gibco BRL Life Technologies, USA) as recommended by the manufacturer. RNA was digested with RNase free DNase I (Promega Corporation, USA) for 30 minutes at 37°C in 50 mmol/L Tris buffer (pH 6.5) with 10 mmol/L $MgCl_2$ and 10 mmol/L dithiothreitol (DTT). The reaction was terminated with 10× termination mix (0.1 M EDTA [pH 8.0] and 1 mg/mL glycogen), and RNA extraction was performed using the phenol method (Ding *et al.,* 1999).

First & second strand cDNA synthesis

First strand cDNA synthesis was performed using the "First Strand cDNA Synthesis Kit" (Invitrogen Corporation, USA) employing reverse primer (RP) (5'-AAGCTTTTTTTTTTG-3'), according to the manufacturer's protocol. The reaction was stopped by heating at 70°C for 10 minutes and was immediately followed by 2^{nd} strand synthesis according to the method described by Gubler and Hoffmann, 1983.

5 µl of the 2^{nd} strand cDNA was amplified using the forward primer (FP) (5'- CAAGCTT GCGATCAGGACC– 3') and the reverse primer (RP) in a 50 µl of reaction volume using Qiagen *Ta*q PCR master mix (Qiagen, Germany) and 200 µMolar each of dTTP, dGTP, dCTP supplemented with 0.5 µl α-P^{32}-dATP (3000 Ci/mmol). The thermal cycling conditions for second strand cDNA synthesis were as follows: 94^0C, 1 minute; 55^0C, 1 minutes; 72^0C, 2 minutes (30 cycles) followed by a hold of 72^0C, 5 minutes. After PCR, the reaction mixture was stored at -20^0C until further use.

AFLP (Amplified Fragment Length Polymorphism) Protocol

The AFLP method was performed as described by Vos *et al*. (1995), with minor adaptations. PCR amplicons generated from the fungal cDNA were loaded and separated in 12% denaturing polyacrylamide gel using a BRL Life Technologies make vertical electrophoresis system (Model: S2). Following a 30 minutes pre-run and a 1 hour run at 1700 volts, the gel was transferred on to Whatman 3MM filter paper and vacuum dried using a Hoefer gel dryer (model SE 1160) at 70°C for 50 min. The dried gel along with the filter paper base was then stapled to a Kodak Biomax MR Sheet, and alignment holes were made in the corners of the paper filter as well as the film. After 1-2 days of exposure in dark conditions, the Kodak film was separated from the dried gel, developed and fixed in appropriate photography-grade solution. Dried autoradiogram was then analyzed for the presence of fragments and one of size ~ 300 bp was identified, re-aligned with the dried polyacrylamide gel and excised from dried gel behind a pexiglass shield for protection. It was then purified, re-amplified using RP and FP primer pair and cloned as described by Santos and Simon, (2002) in pGEM T Easy vector (Promega Corporation, USA) to generate the recombinant plasmid called pGEMH3-1.

Fluorescent DNA sequencing, restriction digestion and electrophoresis

Cloned fragments were sequenced with T7 (5'-GTA ATA CGA CTC ACT ATA GGG C-3') and SP6 (5'-TAT TTA GGT GAC ACT ATA G-3') promoter specific primers using BigDye Terminator v3.1 Cycle Sequencing Kit and analyzed with an Automated Genetic Analyzer, model 3100 (Applied Biosystems, USA). Restriction digestion was performed with Hind III (Bangalore Genei Private Limited, India) according to the manufacturer's instructions. Fragments were resolved on a 12% non denaturing polyacrylamide gel using a vertical electrophoresis apparatus and photographed using a gel documentation system (Bio Rad Laboratories, USA).

Sub cloning of pGEMH3-1 to construct pGEMH3-2

pGEMH3-1 was restricted with Hind III, purified using a commercial PCR product purification kit (Promega Corporation, USA), re-ligated in a

volume of 30 µl using a rapid ligation kit (Fermentas Inc., MD) and used to transform competent E. coli cells (Prat 201) (Nagee *et al*., 2002). Successful transformants harboring the recombinant plasmid were screened through blue white selection using standard protocol as described by Sambrook *et al.* (1989). DNA Insert within these deletion plasmids were sequenced using T7 and SP6 primers. The deletion plasmid thus constructed was designated as pGEMH3-2.

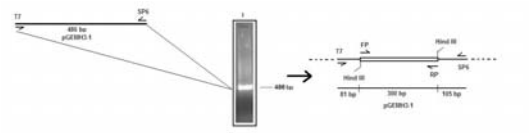

Figure 1: A 486 base pair PCR amplicon generated when the pGEMH3-1 plasmid DNA was used as template for thermal amplification using T7-SP6 PCR primer pair, the annealing sites of which are resident on the vector DNA molecule. The anatomy of the amplicon with regard to annealing position of different primers and Hind III recognition site are indicated.

Molecular cloning of Lambda virus derived 2.2 Kb Hind III fragment in pGEMH3-2

Genomic DNA from lambda bacteriophage (Bangalore Genei Private Limited, Bangalore, India) was restricted with *Hin*d III (Bangalore Genei Private Limited) according to the manufacturer's instructions and resolved on a 1% agarose gel (Promega Corporation, USA). A 2.2 Kb fragment was eluted using conventional method and ligated to pGEMH3-2 linearized with *Hin*d III. Competent *E.coli* cells (Prat 201) were transformed with 2 µl of ligation mixture and positive recombinants were selected using the blue-white selection procedure. Plasmid DNA from transformed *E.coli* was extracted by the method of Birnboim and Dolly, 1979, restricted with *Hin*d III enzyme and resolved on a 1% agarose gel using standard procedures (Sambrook *et al.*, 1989).

RESULTS AND DISCUSSION

Multiple cloning sites in modern DNA cloning vectors such as those of pGEM series (Promega Corporation, USA) provide advantage of access to large number of RE sites arranged within a short stretch of DNA. The most important attribute perhaps is their occurrence within a functional gene *viz*., *Lac* Z. However, it is also observed that common and useful RE sites are often absent from the list. An example of this is the absence of the recognition site for the RE *Hin*d III within the MCS of pGEM T Easy vector (Promega Corporation, USA).

The template for generating the insert DNA for construction of pGEMH3-1 was chosen to be a 300 bp well resolved cDNA amplicon. Further, the reverse primer was specifically designed to amplify a subset of cDNA ends. To enable this, the 3'end of the reverse primer was replaced with a non-T base (G), a strategy typical of AFLP (Vos *et al.,* 1995). This reduced the number of amplified fragments to a reasonable one such that satisfactory resolution was achieved upon electrophoresis. Molecular cloning of the 300 bp eluted cDNA fragment into pGEM T vector generated pGEMH3-1 that was confirmed by a PCR amplification using T7 and SP6 primers. This PCR generated an anticipated fragment of size 486 bp. This fragment comprised of the central portion of cloned cDNA fragment flanked by two ends of the multiple cloning site of pGEM T vector. The anatomy of pGEMH3-1 is summarized in fig. 1.

pGEMH3-2 was generated following deletion of the cDNA insert in pGEMH3-1 with *Hin*d III the cutting site of which was synthetically incorporated in the 5' end of the FP and RP primers respectively. Restriction digestion mapping of the PCR amplicon generated using T7 and SP6 primers indicated that in pGEMH3-1, there were two *Hin*d III recognition sites (Fig. 2) while in pGEMH3-2 it was converted into a unique one.

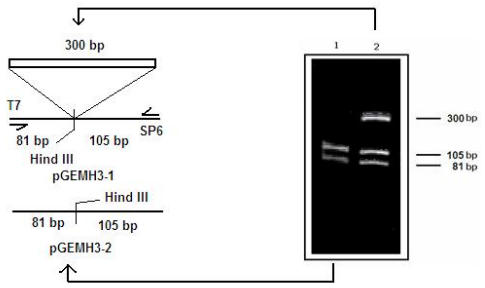

Figure 2: Non denaturing polyacrylamide gel electrophoresis of PCR amplicons generated by T7-SP6 primers and restricted with Hind III using pGEMH3-2 (Lane 1) and pGEMH3-2 (Lane 2) as template DNA. Restriction digestion of the pGEMH3-1 amplicon results in excision of a 300 bp fragment originating for fungal cDNA along with generation of the two flanking fragments of size 81 bp and 105 bp respectively. pGEMH3-2, a deletion plasmid originating from pGEMH3-1 has a unique Hind III recognition site

within the amplicon and thus generated 81 bp and 105 bp fragments only.

analysis. Results indicated that all white colonies

Table 1: Incorporation of sequence of bases within pGEMH3-2 deletion construct, their origin, amino acids coded by the three triplet codons and their category.

Source →	Incorporated by native *Taq* DNA polymerase	Additional base in the 5' end of the forward primer upstream of the *Hin*d III recognition sequence	*Hin*d III recognition site generated as a result of its integration in the 5' ends of forward and reverse primers.							Incorporated by native *Taq* DNA polymerase
Base →	***T***	**C**	**A**	**A**	**G**	**C**	**T**	**T**		***A***
Coding amino acids →		Serine			Cysteine			Leucine		
Nature of amino acid		Nucleophilic			Nucleophilic			Hydrophobic		

The modification at the 5' end of cDNA amplification primers and its relation to TA cloning needs further discussion. The recognition sites for *Hin*d III were incorporated at 5' end of both forward (FP) and reverse (RP) primers. This resulted in specific cleavage of a large section internal to the PCR amplicon generated with these two primers. Since the cloning of the fragment was facilitated by the TA-cloning method, two additional bases were anticipated for incorporation within the deletion plasmid pGEMH3-2. Further, a specific base ("C") strategically incorporated upstream of the *Hin*d III site of the forward primer (FP) only ensured that the total number of bases incorporated were 3 in number. *Hin*d III being a hexacutter had six base-long recognition sequences and addition of 3 extra bases as described above resulted in total incorporation of 9 bases in total in the final construct pGEMH3-2 which is a multiple of 3. The three amino acids coded by these 9 bases can be predicted *in silico* prior to experimentation as indicated in Table 1.

E.coli (Prat 201) transformed with plasmid pGEMH3-2 developed blue color when subjected to blue white selection (Fig. 3). This indicated that *Lac* Z gene was functioning optimally in pGEMH3-2. In order to confirm this, a 2.2 Kb *Hin*d III fragment generated by restricting Lambda bacteriophage DNA was cloned in the *Hin*d III site of pGEMH3-2 and *E.coli* transformants harboring ligated pGEMH3-2 were successfully screened through blue white selection.

Three each of white and blue colonies were subjected to plasmid DNA extraction and further harbored recombinant plasmids while blue colonies represented transformants harboring re-ligated pGEMH3-2 plasmids (Fig. 4). Further, restriction enzyme mapping of the pGEMH3.2 recombinant plasmids containing the 2.2 Kb *Hin*d III – cut fragments indicated successful release of the insert (Fig. 4).

It could thus be concluded that the recognition sequence for *Hin*d III could be successfully incorporated as a unique recognition site in pGEMH3.1 vector and further, the new construct could be successfully used to exploit the blue white selection feature for identifying positive recombinants.

This work forms part of larger research programme of identifying expressed sequence tags (EST) in nematophagous fungi. For this reason, cDNA was used as a template for demonstrating insertion of *Hin*d III recognition sequence within pGEM T Easy vector. Although pGEM series of vectors with special reference to pGEM T Easy is a popular cloning vector, it does not have the recognition sequence for *Hin*d III. However, one of our research objectives warranted generation of library containing *Hin*d III restricted genome fragments in a pGEM T vector. Therefore, the plasmid construct named pGEMH3-2 was engineered to address this need without compromising with the ease and convenience of blue white selection for identifying recombinants in *E. coli*.

However, we have confirmed that this method can be adapted by targeting any region of any genome provided the 5' end modifications as described in this text are maintained. Further, in case TA cloning method is not followed, the 5'

end modifications will be in need to be altered depending upon the type of PCR product treatment followed prior to cloning. If the ends are filled with Klenow fragment then a single base addition at the 5' end of the forward primer (after the RE recognition sequence) will be required. For the reverse primer, only the RE recognition sequence should be added at its 5' end.

Figure 3: Blue color of the E.coli (Prat 201) harboring the pGEMH3-2 recombinant plasmid growing on Luria agar plate amended with X-Gal and IPTG. The color is due to the phenomenon of µ-complementation, which indicates that the amino acid frame of the µ subunit of the LacZ gene within the vector is continuous and functional.

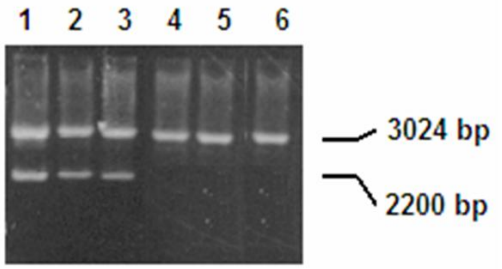

Figure 4: Restriction mapping of recombinant (Lane 1, 2 and 3) and re-ligated (Lane 4, 5 and 6) plasmids with Hind III after a cloning experiment involving ligation of a 2.2 Kb Hind III fragment from Lambda Bacteriophage DNA into non dephosphorylated, Hind III-digested pGEMH3-2 vector.

However, if the ends are treated with S1 nuclease, there will be no need of adding any additional base upstream of both forward and reverse primers apart from the RE recognition sequence-specific bases assuming that the enzyme in question is a hexacutter and total number of bases constitute a multiple of 3 and that the six bases together do not code for any stop codons. In case the RE in question is not a hexacutter, care should be taken to design the primers in such a way that after considering the method of cloning (TA or blunt end) and number of bases that constitute the RE recognition sequence should be three or a multiple of it, preferable 9 and none of the triplet codons be coding for a nonsense codon.

In principle, this method can be adapted to insert virtually any RE recognition sequence within critical region of a cloning vector to broaden its range for accepting RE restricted inserts.

References

Birnboim HC and Dolly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**: 1513-1523.

Bolivar F, Rodríguez RL, Greene PJ, Betiach MC, Heyneker HL, Boyer HW, Crosa JH and Falkow S (1977) Construction and characterization of new cloning vehicles II. A multipurpose cloning system. *Gene* **2**: 95-113

Ding Q, Dong Z, Evers BM (1999) Enterocyte-like differentiation of the Caco-2 intestinal cell line is associated with increase in AP-1 protein binding. *Life Sci* **64**: 175–182.

Gubler, U., Hoffmann, B.J (1983) A simple and very efficient method for generating cDNA libraries, Gene, **25**, 263-269.

Lin-Chao S, Chen WT and Wong TT. (1992) *Mol. Microbiol.* **6**: 3385-3393.

Nagee A., Mukhopadhyaya PN and Kothari IL. (2002) *Rom. Arch. Microbiol Immunol* **60**: 337-348.

Sambrook J, Fritsch EF and Maniatis T. (1989) *Molecular cloning: A Laboratory Manual 2nd Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Santos, C.A.F. and Simon , P.W (2002) Some AFLP amplicons are highly conserved DNA sequences mapping to the same linkage groups in two F_2 populations of carrot. *Genet. Mol. Biol.* **25:** 195-201.

Sutcliffe J.G (1979) *Cold spring Harbor Symposium on Quantitative Biology* **43**: 77-90.

Vos P, Hogers R, Bleeker M, Reijans M, Lee T van der, Hornes M, Frijters A, Pot J, Peleman J, Kupier M and Zabeau M (1995) AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res **23**:4407-4414.

Watson, N (1988) A new revision of the sequence of plasmid pBR322, *Gene* **70**: 399-403.

Yanisch-Perron, C., Veira, J.C. and Messing, J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103-119.

Balbas P, Soberon X, Merino E, Zurita M, Lomeli H, Valle F, Flores N, Bolivar F (1986). Plasmid vector pBR322 and its special-purpose derivatives-a review. *Gene* **50**:3-40.

Chaffin DO, Rubens CE (1998) Blue/white screening of recombinant plasmids in Gram-positive bacteria by interruption of alkaline phosphatase gene (phoZ) expression. *Gene* **219**:91-9.