Novel methods in molecular cloning

Summary of Ph.D. Thesis

Candidate: József Gál

Environmental Biochemistry and Biotechnology Doctoral Program

Supervisor: Miklós Kálmán, Ph.D.

Institute for Biotechnology, Bay Zoltán Foundation for Applied Research and Institute of Genetics, Biological Research Center of the Hungarian Academy of Sciences

University of Szeged

2000.

1. Introduction and Aims

During molecular cloning, the DNA fragment to be multiplied is attached to a cloning vector, a DNA molecule which is capable of self-replication within the host cell. The submitted Ph.D. thesis describes the development of two new methods in molecular cloning: the development of a *gltS*-based cloning system enabling positive selection of recombinant clones and a method for the directional cloning of native products of the polymerase chain reaction (PCR).

Positive selection (also called "direct selection") cloning vectors enable the easy identification of insert-harboring clones. Among selective conditions, cells harboring the unaltered vector are counterselected, and only the insert-harboring derivative of the vector support colony formation of the transformants.

Most of positive selection cloning vectors utilize a conditionally lethal gene as the marker for the selection of insert-harboring clones. Insertional inactivation of the gene enables transformants to grow, but cells transformed by the recircularized vector are counterselected among selective conditions.

Identification of the correct recombinant clones is often laborious despite the availability of the previously reported blue/white screening and positive selection cloning systems. Most of these systems are not free from particular problems which limit their general applicability. The most frequent problems are the need for a special host strain, the low number of available restriction sites for cloning, and the expensive selection medium. Sometimes the selection system does not completely exclude the growth of insert-free cells, just causes a difference in the sizes of the appearing colonies. We had decided to develop a new set of positive selection cloning vectors free from the above mentioned limitations. The new vectors utilize a novel selection principle,

namely the conditional inhibitory effect of the *gltS* gene from *E. coli* K-12 on the growth of *E. coli* cells.

The polymerase chain reaction (PCR) is a method of central importance in molecular biology. Difficulties are frequently encountered during the cloning of PCR products, although numerous methods have been developed for the purpose to date.

The simple blunt ligation of the PCR product to a vector is not directional, and rather inefficient. *Pfu* polymerase generates PCR products with blunt ends, but *Taq* polymerase possesses 3' terminal extendase activity, so end-polishing of *Taq* products with Klenow, T4 or *Pfu* polymerase is advisable.

The extendase activity of *Taq* polymerase is exploited by the method called T-cloning. Since almost exclusively a dAMP residue is added to the 3' end, native PCR products can be ligated to vectors displaying single 3' TMP overhangs. The T-cloning is fairly efficient, but not directional.

Original restriction sites of the template rarely support cloning. The directional cloning of PCR products is frequently accomplished by the inclusion of extra restriction sites in the 5' end of amplification primers and restriction enzyme digestion of the product before ligation to a suitably cleaved vector. Since many restriction enzymes cut rather inefficiently close to DNA termini, often it is necessary to include several further overhanging nucleotides in the primer, elevating the cost of oligonucleotide synthesis.

Limited degradation of the PCR product by uracil DNA glycosylase to form 3' overhangs was reported, but the method requires special primers, enzymatic modification and alkaline treatment of the product. The method can be applied using *Taq* polymerase, but archaeal DNA polymerases, e. g. the *Pfu* or Vent polymerase are seriously inhibited by uracil-containing primers.

The 3'-5' exonuclease activity of T4 DNA polymerase is rather strong. Limited degradation of PCR products amplified by

suitably designed primers results in the formation of 5' single-stranded overhangs on the ends of the PCR product. Unfortunately, enzymatic modification of the product is necessary, and controlling of the limited degradation by T4 DNA polymerase is rather difficult.

The fusion of insert and vector sequences is also possible by PCR, but the methods are dependent on more than one primer pair and more than a single amplification reaction.

A staggered reannealing method to form single-stranded overhangs on PCR product ends was published, but it is also based on two primer pairs.

Recombination of the PCR product and the cloning vector can be accomplished *in vivo*, but a special host strain and more than one primer pair (or rather long primers) are needed.

We decided to develop a novel method for the directional cloning of PCR products using a single primer pair without the

need of any modification of the product after PCR. The method was named Autosticky PCR (AS-PCR).

2. Results and Discussion

The *gltS* gene product GltS is a glutamate-specific permease which is also responsible for the transport of the toxic glutamate analog α-methyl glutamate (αMG). This analog interferes with the conversion of glutamic acid to glutamine and inhibits the growth of *E. coli* on glucose minimal medium. Inactivation of the *gltS* gene by the insertion of foreign DNA abolishes the αMG sensitivity allowing recombinant clones to grow.

The wild type *gltS* gene offers a limited set of unique restriction sites for cloning. To overcome this problem, the KS and SK polycloning sites of the pBC vectors were inserted into the *gltS*, leaving the original reading frame unaltered. The position of insertion of the polycloning sites corresponds to a

hydrophylic loop located cytoplasmatically. The function of the permease was left intact by this in-frame insertion. The resulting *gltS*-derivative marker genes offer a wide set of restriction sites for cloning, enable *in vitro* transcription of the insert sequence and the expression of protein fusions to the N-terminal part of the GltS, even that of proteins incapable of secretion from *E. coli* cells.

The vectors described here combine the versatility of the Bluescript vectors with the possibility of positive selection. Special host strains are not required, the αMG is commercially available and inexpensive. Since the conditionally toxic *gltS* gene is expressed from a weak, constitutive promoter, there is no need to use an inducer in combination with a host strain overproducing the corresponding repressor (e. g. IPTG and $lacI^q$ host). No special strain is required with respect to the chromosomal *gltS* allele either, since the latter does not interfere with the selection. There are no limitations as to the propagation of the empty

cloning vectors. Although overproduction of GltS is toxic to the cells even without added αMG, the low-level synthesis in our system allows stable propagation of the vectors in commonly used laboratory strains. These features, together with the highly efficient selection mechanism, should make our system an attractive alternative to other positive selection cloning systems.

Abasic sites in a DNA template are non-instructional for a DNA polymerase, so it is stalled at the site during synthesis of the complementary strand. The 5' ends of PCR product strands contain built-in amplification primers. Abasic sites within the primers result in the formation of 5' single-stranded overhangs at the ends of the PCR product, enabling its direct ligation to a suitably cleaved cloning vector without the need of any further modification. During our experiments, a stable, tetrahydrofuran derivative abasic site was used.

According to our results, *Taq* polymerase was stalled at tetrahydrofuran abasic sites in high temperature sequencing reactions, and a dAMP residue was added opposite the sites.

Cloning experiments with the Autosticky PCR products showed that the behavior of the products is similar to that of a DNA fragment generated by restriction enzyme digestion, but not to that of a conventional PCR product. In the resulting clones, the abasic site was either deleted or repaired as a TMP, less frequently as a dCMP residue. The possible deletion of the abasic position has to be taken into consideration during the planning of cloning experiments.

In one of the clones, the *Bam*HI site was mutated in a part of the plasmids. Rare mutation events at flanking positions of abasic sites were also encountered previously.

The repair events observed are unlikely to be SOS-dependent, since the used *E. coli* strain DH5αF' is RecA deficient. *E. coli* DNA polymerases I, II and III core each can

bypass a tetrahydrofuran abasic site unassisted by other proteins, although bypass at abasic sites is less efficient under non-SOS circumstances than after SOS-induction. In a previous study, no differences were found in the mutagenic spectra of synthetic abasic sites in RecA deficient and proficient strains.

The original AS-PCR method was described with *Taq* polymerase using primers with a single tetrahydrofuran abasic site and a 5-nucleotides long overhang comprising an *Eco*RI or *Bam*HI sticky end and a site-regenerating nucleotide. Products amplified by *Pfu* or Vent polymerase using the above primers were cloned with very low efficiency, although tetrahydrofuran abasic sites were reported to block a variety of DNA polymerases, including *Pfu*. A set of primers was synthesized to test the effect of various overhang structures on the efficiency of cloning. Our results show that depending on the 3' non-templated extra

nucleotide addition activity of the used polymerase, products with different overhang structures are cloned efficiently.

The tested polymerases behave similarly at abasic sites and at the ends of a linear template. *Taq* polymerase possesses 3' non-templated extra nucleotide addition activity, whereas *Pfu* and Vent polymerases are stalled before the abasic position. In the ligation junction, the abasic site should be opposed either by a polymerase-added extra nucleotide, or the 5' terminal residue of the vector-borne overhang. From these results we conclude that during AS-PCR primer design, the non-templated 3' extra nucleotide addition activity of the desired polymerase should be taken into consideration.

Our AS-PCR method has several advantages. It allows directional cloning of PCR products. There is no need to modify the amplification product before ligation, so the end sensitivity of restriction enzymes does not disturb cloning. Theoretically any desired 5' overhang can be generated, including overhangs that

correspond to restriction sites occurring within the amplified sequences.

List of Publications

Gál, J., Szekeres, S., Schnell, R., Pongor, S., Simoncsits, A., and Kálmán, M. (1999) A positive selection cloning system based on the *gltS* gene of *Escherichia coli*. *Anal. Biochem.* **266**, 235-238.

Gál, J., Schnell, R., Szekeres, S., and Kálmán, M. (1999) Directional cloning of native PCR products with preformed sticky ends (Autosticky PCR). *Mol. Gen. Genet.* **260**, 569-573.

Gál, J., Schnell, R., and Kálmán, M. (2000) Polymerase dependence of Autosticky PCR. *Anal. Biochem.*, accepted for publication.

Patent pending

Gál, J., Schnell, R., Szekeres, S., and Kálmán, M. (1998) „Synthetic oligonucleotides containing abasic sites and their application", Nº P9801320, Hungarian Patent Office