Site-directed mutagenesis by overlap extension using the polymerase chain reaction

(Genetic engineering; sequencing; recombinant DNA; Taq polymerase; oligodeoxyribonucleotide primers; major histocompatibility complex mutants)

Steffan N. Ho¹, Henry D. Hunt¹, Robert M. Horton², Jeffrey K. Pullen¹ and Larry R. Pease¹

¹ Department of Immunology and ² Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN 55905 (U.S.A.)

Received by M.R. Culbertson: 15 December 1988
Accepted: 19 December 1988

SUMMARY

Overlap extension represents a new approach to genetic engineering. Complementary oligodeoxyribonucleotide (oligo) primers and the polymerase chain reaction are used to generate two DNA fragments having overlapping ends. These fragments are combined in a subsequent ‘fusion’ reaction in which the overlapping ends anneal, allowing the 3’ overlap of each strand to serve as a primer for the 3’ extension of the complementary strand. The resulting fusion product is amplified further by PCR. Specific alterations in the nucleotide (nt) sequence can be introduced by incorporating nucleotide changes into the overlapping oligo primers. Using this technique of site-directed mutagenesis, three variants of a mouse major histocompatibility complex class-I gene have been generated, cloned and analyzed. Screening of mutant clones revealed at least a 98% efficiency of mutagenesis. All clones sequenced contained the desired mutations, and a low frequency of random substitution estimated to occur at approx. 1 in 4000 nt was detected. This method represents a significant improvement over standard methods of site-directed mutagenesis because it is much faster, simpler and approaches 100% efficiency in the generation of mutant product.

INTRODUCTION

The alteration of genes and/or the proteins they encode through the substitution of specific nucleotides within a gene sequence by site-directed mutagenesis represents a fundamental tool of modern recombinant DNA technology. It not only allows for the analysis of the structural basis of gene and protein function, but also facilitates the generation of novel gene products. Several techniques designed to introduce specific mutations into cloned genes through oligo-mediated mutagenesis have been described (Smith, 1985; Botstein and Shortle, 1985; Wu and Grossman, 1987). These techniques require...
multiple steps in the introduction of the target sequence into the appropriate bacteriophage or plasmid vector systems and the generation of heteroduplex substrates for mutagenesis. The mutant products often occur at low frequency and must be isolated and reintroduced into the original vector for expression.

An alternative method to create site-directed mutations within the nucleotide sequence employs the PCR (Saiki et al., 1985; Mullis and Faloona, 1987) which uses two synthetic oligos as primers to amplify a nucleotide sequence of interest. These primers anneal at either end of the targeted nucleotide sequence and are oriented in opposite directions. Exponential amplification of the target sequence occurs over the course of multiple rounds of denaturation, annealing and 3' extension by DNA polymerase. PCR can be used to introduce additional sequences, such as restriction sites, by incorporating these into the oligo primers (Mullis and Faloona, 1987). However, this use of PCR as a means of site-directed mutagenesis is limited because all sequence alterations must be introduced within the primer located at the ends of the targeted sequence. Given that restriction sites must be located at the ends of fragment to permit cloning, this approach requires that the sites of mutagenesis be located near these restriction sites. The introduction of mutations at other sites within the amplified gene sequence is therefore not possible.

Mutagenesis by overlap extension, as described here, employs the PCR as a means of creating altered genes from cloned DNA. This method makes possible the introduction of specific mutations into the nucleotide sequence directly from a cloned gene in its original vector with essentially 100% efficiency in a few simple steps. The requirement for restriction endonucleases or DNA ligase is eliminated by generating two PCR fragments having overlapping ends that can be effectively fused by recombining them in a subsequent PCR reaction. This method can be used for the purpose of site-directed mutagenesis or alternatively, as described in the accompanying paper by Horton et al. (1989), for generating hybrid gene constructs between two independent genes, a method we have designated SOE. In this report we demonstrate the use of the overlap extension technique to create site-directed mutations in a murine MHC class-I gene.

MATERIALS AND METHODS

(a) Reagents

DNA polymerase chain reactions were performed using the GeneAmp kit (Perkin Elmer Cetus). Restriction enzymes HindIII and XhoI were obtained from Boehringer-Mannheim. Sequencing of ds plasmid or ss M13 subclones was performed by the dideoxy chain-termination method using Sequenase (U.S. Biochemicals Corp.). Sequencing of ds DNA was performed as previously described (Kraft et al., 1988).

(b) Generation of PCR products

PCRs were carried out using *Taq* polymerase as specified by the manufacturer (Perkin Elmer Cetus). Briefly, amplification of DNA fragments from the plasmid template was achieved by adding 0.1 to 100 ng of template DNA, 50 mM KCl/10 mM Tris·HCl pH 8.3/1.5 mM MgCl2/0.01% (w/v) gelatin/200 μM each of dNTP/1 μM of each primer and 2.5 units of *Taq* polymerase in a final volume of 100 μl. These samples were overlaid with 100 μl of light mineral oil (Sigma Co.) and subjected to 30 cycles of denaturation (1 min, 94°C), annealing (2 min, 50°C), and extension (3 min, 72°C) using a DNA Thermal Cycler (Perkin Elmer Cetus). The products of the reaction were analyzed on an agarose gel containing 3% Nusieve agarose (FMC), 1% Seakem agarose (FMC) and 0.5 μg Et dBs/ml in Tris·acetate buffer/40 mM Tris·acetate/1 mM EDTA pH 8.0).

Although products were obtained by PCR using template concentrations ranging from 1 pg to 1 μg, subsequent sequence analysis from cloned PCR products of MHC class-II mutant genes suggests that there may be a higher observed error frequency among clones generated from reactions using low initial template concentrations as compared with those from reactions using higher template concentrations (S.N. H., and D.J. McKean, unpublished observations). The reasons for this are discussed further below. Thus, to minimize undesired nucleotide changes initial template concentrations of 1 μg are recommended. Because high concentrations of initial template are used, gel purification of the reaction products is routinely performed to ensure
no carryover of the wt template and to avoid amplification by-products, as described below.

(c) **Mouse class-I MHC gene constructs**

The DNA template used in the mutagenesis experiments consisted of a 4.3-kb DNA fragment containing the H-2 K\textsuperscript{b} target sequence (Schulze et al., 1983; Weiss et al., 1983) cloned into the plasmid pUC18. The unique HindIII and XhoI restriction sites were introduced earlier by a conventional gapped heteroduplex mutagenesis procedure (J.K. P., data not shown).

(d) **Synthetic oligos used as PCR primers**

Synthetic oligos were synthesized by the phosphoramidite method using a model 380A Applied Biosystems synthesizer. Oligos were purified by column chromatography using Sephadex G50 (Pharmacia) equilibrated with distilled water. The sequences of the mutagenic oligos are shown in Fig. 3B. The oligos were designed so that the length of the portion of the oligo that was to overlap with the other fragment in the fusion reaction had a calculated denaturation temperature that matched the annealing temperature used in the PCR reaction (50°C). The ‘melting’ temperature (T_\text{m}) was calculated as follows:

\[
T_\text{m} (\degree C) = 4(G + C) + 2(A + T)
\]

as described by Suggs et al. (1981).

(e) **Joining DNA fragments by overlap extension**

PCR-generated DNA fragments from an initial set of reactions were either used directly in a subsequent overlap extension reaction or first purified. In the former case, 10 \mu l (or dilutions thereof) from the two PCR reactions containing the overlapping fragments were mixed and subjected to PCR amplification using the external oligo primers (analogous to 'a' and 'd' in Fig. 1). Alternatively, the DNA fragments were first purified to eliminate contaminating PCR-amplified products generated in the first PCR reactions. Purification of PCR products was performed by size-fractionating the DNA by electrophoresis through an agarose gel as described for the analysis of PCR reaction products. The band of the appropriate size was cut from the gel and purified using the GeneClean kit (Bio101) according to the manufacturers specifications. The GeneClean kit is based on a previously described method (Vogelstein and Gillespie, 1979). The reaction conditions used for the fusion of the two PCR-generated fragments were identical to those used to generate the fragments.

(f) **Hybridization analysis**

DNA fragments were transferred onto nylon membranes (MSI) and hybridized at 70°C with the mutagenic oligo probe 5' end-labeled with \(^{32}\)P by polynucleotide kinase (NEN) as described (Duran and Pease, 1988).

**RESULTS AND DISCUSSION**

(a) **Description of the overlap extension technique**

Mutagenesis by overlap extension involves the generation of DNA fragments that, by virtue of having incorporated complementary oligo primers in independent PCR reactions, can be effectively 'fused' anywhere along the gene sequence by combining them in a second primer extension reaction. The method is illustrated in Fig. 1. In separate PCRs two fragments of the target gene sequence are amplified. Each reaction uses one flanking primer that hybridizes at one end of the target sequence (primer 'a' or 'd' in Fig. 1) and one internal primer that hybridizes at the site of the mutation and contains the mismatched bases (primer 'b' or 'c' in Fig. 1). Since the product generated in a PCR incorporates the primers, the wt sequence will not be amplified. By using two internal primers that overlap, the two fragments AB and CD, generated in the first PCR, can be fused by denaturing and annealing them in a subsequent primer extension reaction. The overlap allows one strand from each fragment to act as a primer on the other, and extension of this overlap results in the mutant product (‘mutant fusion product’ in Fig. 1). Even though the annealing of the short overlap between the two fragments may occur at low frequency, the inclusion of additional flanking primers (‘a’ and ‘d’ in Fig. 1) allows the ‘fusion’ product that is formed to be amplified by PCR.
Fig. 1. Schematic diagram of site-directed mutagenesis by overlap extension. The ds DNA and synthetic oligos are represented by lines with arrows indicating the 5'→3' orientation. The site of mutagenesis is indicated by the small black rectangle. Oligos are denoted by lower-case letters and PCR products are denoted by pairs of upper-case letters corresponding to the oligo primers used to generate that product. The boxed portion of the figure represents the proposed intermediate steps taking place during the course of reaction (3), where the denatured fragments anneal at the overlap and are extended 3' by DNA polymerase (dotted line) to form the mutant fusion product. By adding additional primers 'a' and 'd' the mutant fusion product is further amplified by PCR.

As an additional advantage, site-directed mutagenesis by overlap extension is extremely flexible in the variety of sequence alterations that can be achieved. In addition to point mutations, insertions and deletions can be incorporated into the overlapping oligo pair (Fig. 2). Insertions are possible because the 'b' and 'c' overlapping oligos need only be complementary to the template in the 3' portion of the oligo. Additional sequence present in the 5' portion of these oligos will be incorporated into the AB and CD fragments and, upon generation of the AD fusion product, will represent an insertional mutation (Fig. 2A). Deletions are similarly obtained by using paired 'b' and 'c' oligos that correspond to the sequence after the deletion (Fig. 2B).

(b) Site-directed mutagenesis of a mouse class-I MHC gene by overlap extension

To demonstrate this method we have introduced mutations into the mouse major histocompatibility complex class-I Kb gene that correspond to the 5 aa changes found in the naturally occurring Kb mutant bm10 (Nathenson et al., 1986; recent sequence analysis of the bm10 mutation has demonstrated the presence of an additional aa change at position 167, in which the wt tryptophan is replaced by a serine; J. Schneck and D. Margulies, personal communi-
cation). A schematic representation of the class-I gene illustrating the regions targeted for PCR amplification and mutagenesis is shown in Fig. 3A, with the oligo primers labeled as in Fig. 1. The external primers ‘a’ and ‘d’ were located approx. 100 bp outside of the unique HindIII and XhoI restriction sites, thus making it possible to easily ligate the fusion product from the overlap extension reaction back into the expression vector containing the remainder of the wt class-I gene construct. The overlapping oligo pair used to introduce the mutation for the mutant designated bm10-1 (oligos ‘b1’ and ‘c1’) results in the substitution of Glu for Lys. These 16-mer oligos are completely overlapping, in contrast to the other oligo pairs which consist of a 35-mer (‘b2’) and a 26-mer (‘c2’ and ‘c3’) that overlap by 16 nt. The oligo pair ‘b2’ and ‘c2’ introduced 4 aa substitutions with 5 nt changes to give rise to the mutant designated bm10-2, while the oligo pair ‘b2’ and ‘c3’ introduced 5 aa substitutions with 6 nt changes to give rise to the bm10 mutant (Fig. 3B). These constructs are designed to investigate the functional importance of specific amino acid residues hypothesized to be involved in antigen and/or T cell receptor

Fig. 3. Mutagenesis strategy for the class-I MHC gene. (Panel A) The oligos used as primers for the PCR reactions are shown in relation to the map of the H-2 K\(^{b}\) class-I gene. The locations of the unique HindIII and XhoI sites are shown. The flanking primers corresponding to ‘a’ and ‘d’ in Fig. 1 anneal to sequences within the second exon and third intron, respectively, outside the unique HindIII and XhoI restriction sites. The sequences of these oligos are as follows: ‘a’, 5’-GGCTACTACAACCAGAG-3’; ‘d’, 5’-TGGGAGAGCCCTGAGCT-3’. The products of the first two PCR reactions (AB and CD) and the fusion reaction (AD) are shown in relation to their location within the class-I gene. Digestion of the AD fragment with HindIII + XhoI results in the formation of three fragments of 101, 561 and 108 bp in length. (Panel B) The complementary oligo pairs used to introduce the mutations were either entirely or partially overlapping, although in each case the overlap consisted of 16 bp.
Fig. 4. Generation of a mutant fusion product from overlapping PCR-generated fragments by overlap extension. (Panel A) Electrophoretic analysis of PCR amplification products forming mutant 6×1. 123-bp ladder markers (BRL), wt AD product generated by PCR using primers 'a' and 'd' and the K template (AD wt), fragment AB generated by PCR using primers 'a' and 'bl' (refer to Fig. 2B), fragment CD generated by PCR using primers 'cl' and 'd', AD fusion product generated by combining fragments AB and CD directly with additional primers 'a' and 'd' (AB + CD titration; these samples were generated by adding ten-fold dilutions of the samples containing AB and CD, starting with 10 µl of each undiluted sample). (Panel B) Electrophoretic analysis of PCR amplification products forming mutant 6m10-2. AD wt fragment as above, fragment AB generated by PCR using primers 'a' and 'b', fragment CD generated by PCR using primers 'cl' and 'd', AD fusion product generated by combining fragments AB and CD that had been gel purified, fusion fragment AD digested with HindIII, fusion fragment digested with HindIII + XhoI, wt HindIII-XhoI insert, 123-bp ladder marker, HindIII-digested phage λ DNA marker (New England Biolabs). (Panel C) Southern-blot analysis of the PCR-amplified products. The gel shown in Fig. 4B was probed with the oligo 'c2'.

binding to the class-I molecule. The results of functional studies using cell lines transfected with these mutant class-I constructs will be reported elsewhere.

The products from the PCR using primer 'a' or 'd' in combination with the overlapping oligo primers 'bl' and 'cl' (used to introduce the mutations designated bm10-1) are shown in Fig. 4A. The sizes of the AB product and the CD product are consistent with the map shown in Fig. 3A. The products from the first set of PCR reactions appeared relatively homogeneous based on EtBr staining, although additional PCR-generated products of differing sizes are visible at much lower concentrations. The fusion reaction in which dilutions of the first PCR containing the AB and CD fragments were mixed directly, with no further purification, resulted in the generation of multiple products including the AD fusion product of the predicted 770-bp length (Fig. 4A; AB + CD titration). For comparison, the AD wt fragment generated by PCR using the original template and primers 'a' and 'd' gave rise to a 770-bp fragment (Fig. 4A, AD wt). The presence of a predominant 350-bp product generated in this fusion reaction suggests that under these conditions the fusion or amplification of minor PCR products may be favored in the second reaction. To increase the efficiency and yield of the fusion reaction, the input fragments AB and CD were gel-purified. The product of the PCR fusion reaction using gel-purified AB and CD fragments contained predominantly the AD fusion product and no visible 350-bp band (Fig. 4B). Because this modification in the protocol eliminated the 350-bp fragment, the band was not characterized further.

To determine whether the product from the fusion reaction contained the intended nucleotide substitutions, the reaction products shown in Fig. 4B were transferred to a nylon membrane and hybridized with end-labeled oligo 'c2'. The CD fragment and the
AD fusion product hybridized strongly to the mutagenic oligo (Fig. 4C, AD, AD HindIII; AD HindIII-XhoI), while the PCR product using primers 'a' and 'd' on wt template and the wt HindIII-XhoI insert showed faint hybridization. Hybridization to CD was faint due to the lower concentration of DNA loaded. Fragment AB showed no hybridization because it only matched 16 nt of the 26-nt oligo 'c2'. These results indicate that the fusion product contains the mutagenic sequence. By gel-purifying the AB and CD fragments, the wt template is not carried over into the PCR fusion reaction, thereby eliminating the possibility of also amplifying wt sequences.

To quantitatively assess the efficiency of this mutagenesis procedure, colony lifts from bacteria transformed with the ligation of the vector with the PCR fusion product were hybridized with the mutagenic oligo. Among 168 colonies probed there were only three colonies that did not hybridize with the mutagenic oligo (data not shown). This represents at least a 98% efficiency of mutagenesis. The three negative colonies were not analyzed to determine whether they represented failures in ligation or mutagenesis. Furthermore, complete sequence analysis of PCR-generated DNA fragments revealed no clones containing the wt sequence. Thus, the efficiency of this mutagenesis procedure is essentially 100% in terms of generating altered product.

(c) Sequence analysis of cloned mutant genes generated by overlap extension

The mutagenic AD fragment was then cut sequentially with HindIII and XhoI (Fig. 4B) and ligated back into the original plasmid vector. The resulting mutant clones were either sequenced directly by ds sequencing of the mutant plasmid or by subcloning into M13. All clones sequenced contained the desired mutations introduced by the paired mutagenic oligos. Sequence data demonstrating the mutations forming bm10, bm10-1 and bm10-2 in comparison with the wt sequence is shown in Fig. 5. In sequencing the bm10 PCR-generated mutant clones as well as three addi-

Fig. 5. Sequence analysis of bm10 mutants generated by the overlap extension technique. The three bm10 site-directed mutants generated by the overlap extension method and the wt were sequenced from ss M13 templates. The region shown encompasses nt encoding aa 160 through 175, where the mutations were directed. The altered nt are indicated by the horizontal dashes along the right margin of each mutant sequence.
tional mutants, there was one undesired nucleotide change, detected out of more than 3900 nt sequenced. This represents sequence data from a total of nine independently derived clones generated from five separate overlap extension reactions needed to produce the three bnl0 mutant genes and three additional mutant genes.

(d) Analysis of error frequency

The method of generating site-directed mutants described here has recently been suggested independently by Higuchi et al. (1988) who used direct sequencing analysis of the PCR-generated DNA fragments to demonstrate the presence of molecules containing an introduced nucleotide change. However, these mutated products were not cloned and no conclusions could be made concerning the frequency of undesired nucleotide changes resulting from the misincorporation rate per nucleotide per cycle of the Taq polymerase. Here we have analyzed cloned, PCR-amplified DNA fragments generated using the overlap extension technique and have observed that the 0.026% error frequency (frequency of undesired nucleotide changes observed in the cloned PCR product) derived from the results presented here (one error in over 3900 nt sequenced) is much lower than the error frequency of 0.25% previously reported (Saiki et al., 1988). Assuming this error frequency of 0.25% applies here, the probability of obtaining one or fewer undesired mutations in 3900 nt is 0.0006, calculated by the binomial distribution (3900 x 0.9975 exp 3899 x 0.0025). Therefore, the error frequency observed here is significantly less than previously reported.

This low error frequency may be due to the use of cloned template rather than genomic DNA template. In the latter case the template sample actually contains only a few copies of the target sequence. A misincorporation in an early amplification cycle would be propagated through subsequent cycles and represent a significant portion of the final product. Starting with more of the specific target template allows for the generation of the final product after fewer cycles of PCR amplification, thereby decreasing the chance of an incorrect base being introduced by the polymerase. Assuming the misincorporation rate per nucleotide per cycle of the Taq polymerase is constant, the only other variable affecting the error frequency is the number of cycles of amplification. Thus, our observed error frequency of 0.026% may be much lower than the previously reported error frequency of 0.25% because, by using cloned template, the number of cycles of PCR amplification required is much less. Whatever the explanation for the low error frequency of 0.026% observed here may be, it is clear from these results that the error frequency of 0.25% published previously (Saiki et al., 1988), which would have been prohibitively high for purposes of cloning, does not apply in this case. Therefore, the overlap extension technique is a practical alternative to conventional methods of mutagenesis or gene splicing, as described below.

(e) Advantages of the overlap extension technique

The use of overlap extension for site-directed mutagenesis represents a significant technical improvement over current methods. The standard methods (Smith et al., 1985) involve annealing the mutagenic oligo to an ss wt sequence. This requires a cloning step to transfer the wt sequence into an ss vector system such as M13. The wt sequence then serves as a template and the mutagenic oligo serves as the primer for the synthesis of a new mutagenic strand by DNA polymerase. Despite the use of selection systems, the efficiency of this process (i.e., the % of output molecules that are not of the wt sequence) is significantly less than 100% because of the presence of the wt template throughout the mutagenesis procedure. This makes it necessary to screen the products using the mutagenic oligo as a hybridization probe. The major advantages of site-directed mutagenesis by overlap extension are its simplicity and virtual 100% efficiency. Site-directed mutagenesis by overlap extension eliminates the need for ss template and viral vector intermediates, thus eliminating a cloning step. The virtually 100% efficiency forgoes the need to screen the products of the mutagenesis for the presence of the wt changes. The elimination of these steps reduces the time and effort required to generate site-directed mutations.

Although the method described here represents a novel approach to site-directed mutagenesis, an essential feature of this method is the ability to fuse or recombine two independent fragments of DNA without the use of a DNA ligase. The method of overlap extension can be used to generate recomb-
nant DNA products from two independent genes without the use of restriction endonucleases. Gene splicing by overlap extension eliminates the need for the introduction of restriction sites and is independent of the targeted sequences. We have used this technique to construct chimeric genes comprised of segments of different class-I genes (Horton et al., 1989) and are currently involved in constructing additional fusion proteins. These constructs have involved the fusion of as many as four DNA fragments having a combined size of more than one thousand bases. Given the substantial savings in time, the increased versatility in design strategies for introducing site-directed mutations and recombinig gene sequences, and the low rate of undesired nucleotide changes, gene splicing by overlap extension represents a significant advance in recombinant DNA technology.

ACKNOWLEDGEMENTS

We would like to thank Dr. S.S. Sommer and his lab for the use of their DNA Thermal Cycler and Drs. R.T. Abraham, B.N. Beck, D.J. McKean and E.D. Wieben for reviewing the manuscript. S.N.H. was supported by a medical scientist scholarship from the Life and Health Insurance Medical Research Fund. H.D.H. and J.K.P. were supported by NIH training grant CA-09127. Additional funding for this work was obtained through NIH grants AI-22420, AI-00706 and CA-42199.

REFERENCES


