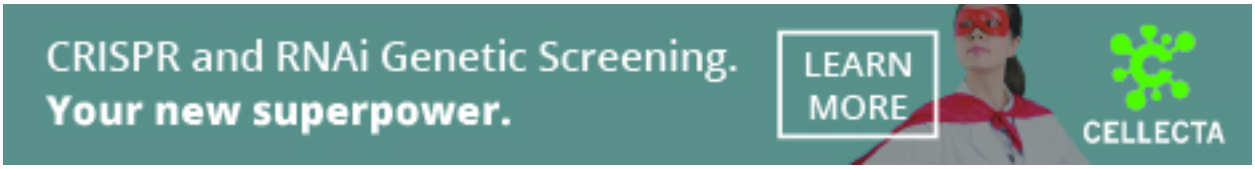
A PCR-derived library of random point mutations within the V3 region of simian immunodeficiency virus.

F Kirchhoff and R C Desrosiers

Genome Res. 1993 2: 301-304
Access the most recent version at doi:10.1101/gr.2.4.301



A PCR-Derived Library of Random Point Mutations within the V3 Region of Simian Immunodeficiency Virus

Frank Kirchhoff and Ronald C. Desrosiers

New England Regional Primate Research Center, Harvard Medical School, Southborough, Massachusetts 01772-9102

Oligonucleotide primers corresponding to variable region 3 (V3) of simian immunodeficiency virus (SIV) were randomly mutagenized during synthesis by doping each of the four nucleoside phosphoramidites with a small amount of the other three. PCR was then used to incorporate the altered sequences into larger, clonable DNA fragments by spliced overlap extension (SOE). With the composition of the phosphoramidites used, 53 of the 100 clones analyzed were unique, having one or more point mutation within the 84-bp target sequence. These 53 unique clones contained an average of 2.1 nucleotide substitutions and 1.5 amino acid substitutions per clone within the target V3 sequence. Of the internal 25 amino acid positions within the V3 domain, 23 were changed at least once. This method should be generally useful for the construction of libraries of random point mutations within a defined target DNA sequence.

Among the methods that have been developed for the specific mutagenesis of DNA sequences, PCR is one of the most powerful.[1–6] It allows rapid preparation of relatively large quantities of mutated DNA with the use of mutated oligonucleotide primers. Because of the error frequency of *Taq* and other polymerases used in PCR, it is generally most convenient to use PCR-based mutagenesis strategies which minimize the length of the PCR-derived DNA,[1–3] rather than amplifying entire plasmids.[4–7] Recently, Morrison and Desrosiers used PCR with spliced overlap extension (SOE) in the presence of mixed populations of oligonucleotide primers to mutagenize selected sites in a target DNA sequence.[7]

In some cases, it is desirable to analyze an extensive collection of mutations over the full length of a defined target DNA sequence. Hutchinson et al. described the generation of a library containing all possible point mutations within a 30-bp target sequence.[8] They doped each of the four nucleoside phosphoramidites with a small amount of the other three and cloned the two complementary mutagenic oligonucleotides bearing cohesive ends directly into an M13 vector. This specific approach does not have general utility because of the need for conveniently located restriction sites at the ends of the target sequence. However, it did set the stage for development of protocols for extensive mutagenesis of target DNA sequences in phage vectors.[9,10]

Here, we present a method that combines the advantages of the SOE PCR method with the doping procedure originally used by Hutchinson et al. for the generation of potentially complete libraries of point substitutions. This procedure is likely to be generally useful for saturation mutagenesis of small target regions. Because it is easy to generate a large number of variants in a defined area with this method, it may also be useful for biological selection experiments.

MATERIALS AND METHODS

Oligonucleotide Synthesis

Oligonucleotides were synthesized with a Cyclone DNA synthesizer (Milligen-Biosearch, Burlington, MA). Prior to synthesis, each vial of phosphoramidite was doped with 1.5% (vol/vol) of each of the other three. The primers were purified on oligonucleotide purification cartridges (Applied Biosystems Inc., Foster City, CA). The four mutagenic primers that were used span nucleotides 7535–7618 of the simian immunodeficiency virus (SIV)mac239 sequence.[11] The primer sequences are shown in Table 1. The expected number of nucleotide changes per oligonucleotide was calculated on the basis of the concentration of the other three nucleotides [4.5% (vol/vol)] versus the concentration of the wild-type nucleotide [95.5% (vol/vol)].

PCR Conditions

Twenty-five nanograms of the p316EM* plasmid served as the template for PCR amplification; p316EM* contains the *env* gene of a macrophage tropic variant of the SIVmac239 clone.[12] Reactions were performed using a DNA thermal cycler (Perkin-Elmer Cetus, Norwalk, CT) and the manufacturer's PCR reagent kit.

2:301–304©1993 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/93 $3.00

TABLE 1 Oligonucleotides Used for Mutagenesis of the SIVmac V3 Region

Primer	Sequence (5′ → 3′)	Position in SIVmac239 sequence	Expected number of nucleotide changes per oligonucleotide
F5	CCACTGTAACACTTCTG	7230 → 7246	0
F6	CTGTATTCCTATCTTCTACC	7852 → 7833	0
F7	GGTGACTGGTAAAACTGTCTTATTTCCTGGTCTTCTAC	7572 → 7535	1.7
F8	AGACAGTTTTACCAGTCACCATTATGTCTGGATTGGTTTTCC	7553 → 7594	1.9
F9	GGAAAACCAATCCAGACATAATGGTGACTGGTAAAACTGTC	7594 → 7554	1.8
F10	TATGTCTGGATTGGTTTTCCACTCACAACCAATCAATGATAGGC	7575 → 7618	2.0

For the synthesis of oligonucleotides F7–F10, 1.5% (vol/vol) of each of the other phosphoramidite nucleoside mixes was added to each vial. Thus, the probability for a change at each position is 4.5%.

The PCR mutagenesis strategy is shown schematically in Figure 1. The first round of amplification using primer pairs F5/7, F5/9, F6/8, or F6/10 was performed in a 100-μl volume using 60 pmoles of each primer (25 cycles, 1 min at 94°C, 1 min at 45°C, 1 min at 68°C, 2.5 mM $MgCl_2$). Thereafter, 25 μl of the PCR products was analyzed by electrophoresis in a 1.5% agarose gel. PCR products of the expected size, ranging from 280–362 bp, were isolated with the GeneCleanII kit (BIO101 Inc., La Jolla, CA). To generate a linear double-stranded template, 0.1 pmole of the left and right half fragments were combined and six cycles of amplification without any additional primer were performed. Then, 50 pmoles of outer primers pF5 and pF6 were added and another 28 cycles of amplification were performed to generate the three classes of 622-bp products, which differ slightly in the positions of mutations (Fig. 1).

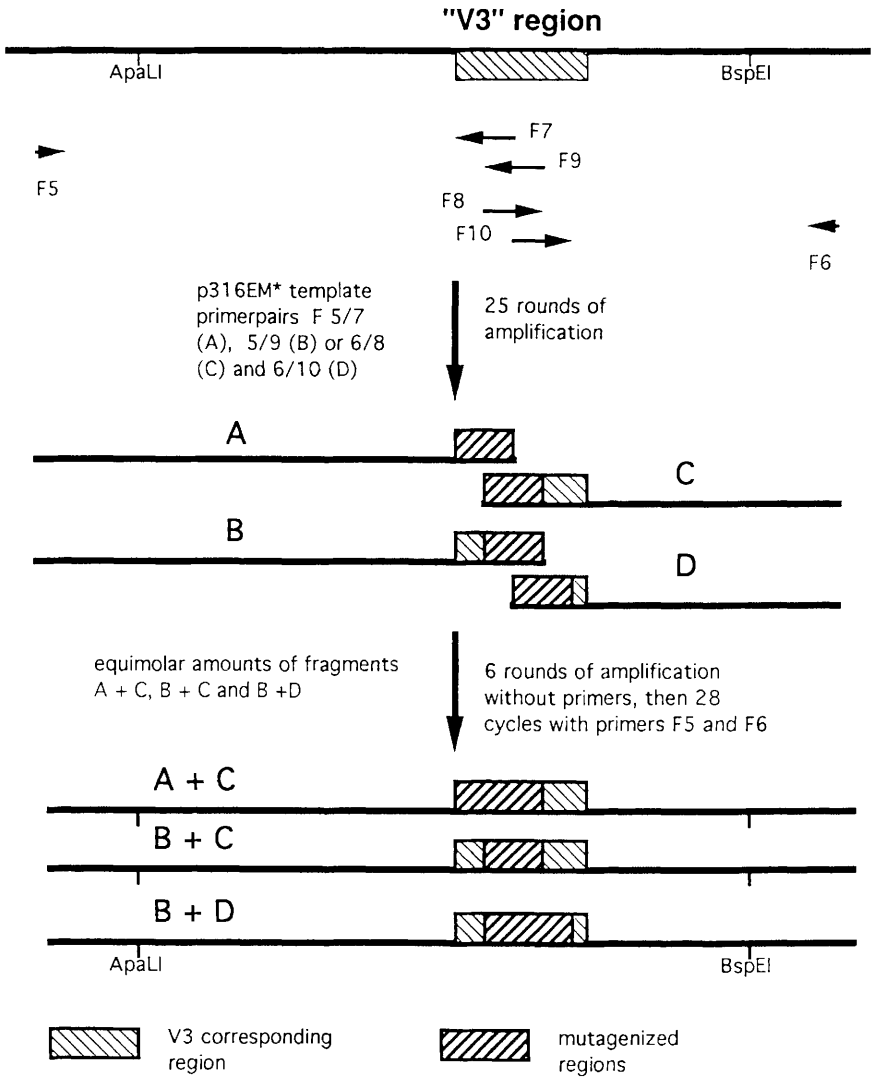


FIGURE 1 Strategy for random mutagenesis of the SIVmac V3 region. The p316EM* plasmid was used as a template for amplification with doped oligonucleotides (see Table 1). First round, PCR-generated DNA fragments A–D were subsequently used for the synthesis of DNA fragments containing random changes in the V3 region at three slightly different locations, as indicated.

Cloning and DNA Sequencing

The PCR products were isolated from agarose gels and digested with *Apa*LI and *Bsp*EI. Because three *Apa*LI sites are present in the p316EM* plasmid, a gel purified subfragment of p316EM* was used for insertion of the mutant PCR products. The p316EM* vector DNA was obtained after transformation of the dam$^-$ strain of *Escherichia coli* JM110 (Stratagene, La Jolla, CA). The unmethylated p316EM* DNA was digested with *Bsp*EI, dephosphorylated, and digested with *Sph*I. The 7.7-kbp restriction fragment was isolated and ligated to the dephosphorylated 0.85-kpb *Sph*I/*Apa*LI restriction fragment derived from the same vector. The 8.6-kbp ligation product was isolated from an agarose gel, ligated to a fivefold molar excess of an equimolar mixture of the three different 457-bp *Apa*LI/*Bsp*EI PCR fragments, and transformed into supercompetent *E. coli* XL-1 Blue cells (Stratagene). About 10% of the transformation mix was spread on LB-Amp plates in serial dilutions, and the rest was used directly for plasmid

DNA preparation. The resulting plasmid products varied considerably in size and 95% were smaller than expected. Plasmids migrating at the correct size were purified from an agarose gel and retransformed into *E. coli* XL-1 Blue. Plasmid DNA was prepared from single colonies, and the V3 target sequence was determined. For a subset of clones containing changes within V3, the entire sequence between the *Apa*LI and *Bsp*EI sites was determined.

Serial dilutions of the transformation mixture showed that the mixed plasmid population represented ~4000 colonies. Agarose gel analysis revealed that only ~5% of the plasmid population obtained had the correct size. Therefore, the plasmid population used to retransform *E. coli* after the purification step represented only ~200 different PCR fragments. The low proportion of correctly sized plasmids obtained after the first round of transformation was probably related to the unmethylated status of the vector DNA. In recent experiments using normally methylated p316EM* DNA cut with *Mro*I (Boehringer-Mannheim) instead of *Bsp*EI, ~90% of the plasmid population after initial transformation had the correct electrophoretic mobility on an agarose gel. In addition, cloning efficiencies were ~20-fold higher than in the initial experiment.

RESULTS AND DISCUSSION

The analysis of 105 single colonies showed that 100 of them contained the insertion of a 0.46-kbp restriction fragment. Sequencing of the variable region 3 (V3) target region within these 100 clones showed that 69 contained at least one nucleotide change from the parental. Of these 69 clones, 53 had a unique sequence (Table 2). Eight sequences were present twice and four sequences were present three times in the plasmid population.

TABLE 2 Numbers of Amino Acid and Nucleotide Substitutions in the 53 Unique V3 Mutant Clones

Number of nucleotide changes/number of amino acid changes	Number of clones
1/0	10
1/1	10
2/1	9
2/2	9
3/1	1
3/2	4
3/3	1
4/2	2
4/3	4
4/4	1
5/4	2
Total	53

Among the 53 unique clones, there were 112 nucleotide substitutions and 77 predicted amino acid substitutions (Table 2); thus, the average number of nucleotide substitutions per V3 target sequence was 2.1 per clone, and the average number of translated amino acid substitutions per V3 target sequence was 1.5 per clone. If one assumes no selection against incorporation of nucleotide substitutions into cloned DNA, we would have expected an average of 1.9 nucleotide substitutions per clone. This expectation is based on an average of 1.9 nucleotide misincorporations per single-stranded oligonucleotide primer (Table 1), 3.8 single-stranded nucleotide substitutions per double-stranded DNA fragment, and it assumes equidirectional resolution to 1.9 matched base-pair substitutions per clone by the PCR process and mismatch repair in *E. coli*. Thus, the 53 unique clones had an average nucleotide substitution frequency (2.1) that was close to what would be expected (1.9) based on these assumptions. If, however, one includes the total 100 clones, 31 of which had no substitutions at all in the target sequence, the observed frequency was somewhat less than expected. Other investigators have observed a somewhat lower than expected mutation frequency among the final clones analyzed following oligonucleotide doping.[7,8] Inefficient PCR amplification with mismatched primers and inefficient survival of mismatched sequences in *E. coli* may contribute to a lower than expected frequency of nucleotide substitutions in the final clones. PCR amplification may be especially inefficient when the sequences near the 3′ end of the primer are not perfectly matched.

Despite the fact that the 84-bp target sequence was AT rich (59.5%), 73 of the 112 nucleotide substitutions (65%) were changes to A or T (Fig. 2). Hutchinson et al.[8] also observed a higher than expected frequency of changes to A or T. The reasons for this are not clear but could possibly involve the kinetics of phosphoramidite condensation during oligonucleotide synthesis.

Of the 53 mutant clones, 20 had a single nucleotide change, 18 had two changes, 6 had three changes, 7 had four changes, and 2 had five changes in the V3 target sequence (Table 2). Two clones

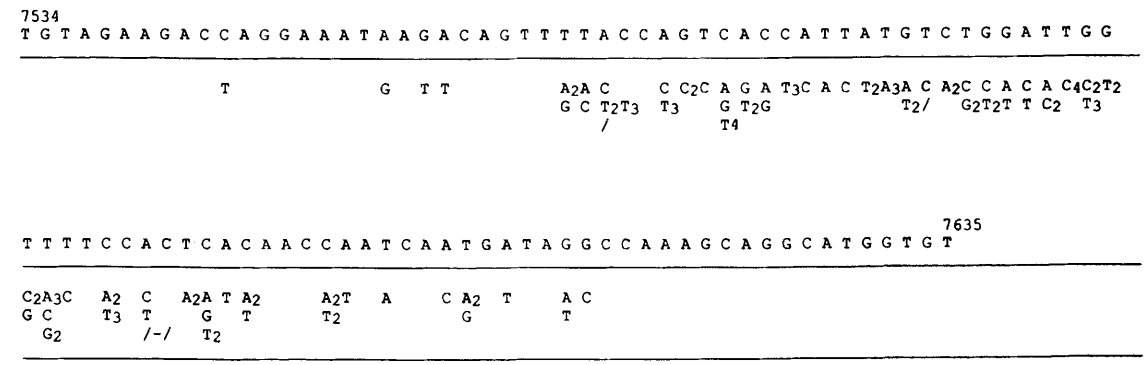


FIGURE 2 Nucleotide changes in the SIVmac p316EM* V3 region resulting from PCR-directed mutagenesis. Numbers indicate the frequency of mutations if greater than one. (/) A single nucleotide deletion; (/-/) -a double nucleotide deletion. The complete length of the V3 cysteine loop is shown. The target sequences for mutagenesis include bases 7535–7618.

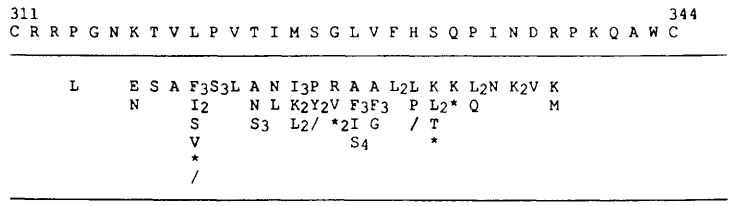


FIGURE 3 Amino acid changes in the SIVmac p316EM* V3 region resulting from PCR-directed mutagenesis. Numbers indicate frequency of the mutation when greater than one. (*) Change to a stop condon. (/) Location of nucleotide deletions resulting in a frameshift of downstream sequences.

had a deletion of a single nucleotide and one clone had a deletion of two nucleotides (Fig. 2). The use of PCR fragments A and C, B and D, and A and D allowed nucleotide substitution over the full 84 bp of target sequence (Fig. 1). This procedure was also intended to obtain higher frequencies of point mutations within the central part of the V3 region due to the overlapping nature of the mutagenic F7, F8, F9, and F10 primers (Fig. 1). A higher frequency of mutations was concentrated in this central region (Figs. 2 and 3). This basic mutagenesis strategy can be easily modified to distribute point mutations in any manner desired.

The vast majority of changes occurred in the region covered by the mutagenic oligonucleotides. However, eight additional changes, which were most likely due to errors by the *Taq* polymerase, were identified elsewhere in the 0.46-kb PCR fragments analyzed. This random substitution rate was ~1 per 3300 bp and, thus, similar to the rate of 1 per 4000 bp reported by Ho et al.[1] and much lower than the 1 per 840 bp found by Morrison and Desrosiers.[7]

For the random mutagenesis of short target regions within longer DNA sequences, the method described here has some advantages over those described previously.[1,2,7,8] The doped oligonucleotide synthesis, PCR, and cloning steps are easy to perform and require no unusual advanced preparation. The average number of mismatches in each oligonucleotide and any preference for desired changes (e.g., by adding only one additional nucleoside solution) can easily be varied. The choice of the external primers makes it possible to minimize the size of the mutated PCR fragment and also makes it easy to reclone the mutated fragments into longer sequences.

The retrieval of 53 unique mutants from a population of 100 clones shows that this approach is useful for the extensive random mutagenesis of defined target sequences. PCR fragments obtained with this mutagenesis strategy potentially contain all possible point substitution mutations. With high cloning efficiencies, this approach can also be useful for biological selection experiments. For example, we have used mixed plasmid populations containing a high number of changes within V3 to select variants with altered cell tropism and to identify changes within V3 that can compensate for defects in other regions of the SIV envelope.

ACKNOWLEDGMENTS

We thank Hilary G. Morrison and Dean Regier for their helpful advice. This work was supported by U.S. Public Health Service grants AI 25328 and RR00168, and by a fellowship of the Deutsches Krebsforschungszentrum to F.K.

REFERENCES

1. Ho, S.N., H.D. Hunt, R.M. Horton, J.K. Pullen and L.R. Pease. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77:** 51–59.
2. Horton, R.M., H.D. Hunt, S.N. Ho, J.K. Pullen, and L.R. Pease. 1989. Engineering hybrid genes without the use of restriction enzymes: Gene splicing by overlap extension. *Gene* **77:** 61–68.
3. Higuchi, R., B. Krummel, and R.K. Saiki. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: Study of protein and DNA interactions. *Nucleic Acids Res.* **16:** 7351–7367.
4. Jones, D.H., K. Sakamoto, R.L. Vorce, and B.H. Howard. 1990. DNA mutagenesis and recombination. *Nature* **344:** 793–794.
5. Hemsley, A., N. Arnheim, M.D. Toney, G. Cortopassi, and D.J. Galas. 1989. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res.* **17:** 6545-6551.
6. Jones, D.H. and B.H. Howard. 1991. A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction. *BioTechniques* **10:** 62–66.
7. Morrison, H.G. and R.C. Desrosiers. 1993. A PCR-based strategy for extensive mutagenesis of a target DNA sequence. *BioTechniques* **14:** 454–457.
8. Hutchinson, C.A., S.K. Nordeen, K. Vogt, and M.H. Edgell. 1986. A complete library of point substitution mutations in the glucocorticoid response element of mouse mammary tumor virus. *Proc. Natl. Acad. Sci.* **83:** 710–714.
9. Zoller, M.J. and M. Smith. 1987. Oligonucleotide-directed mutagenesis: A simple method using two oligonucleotide primers and a single-stranded DNA template. *Methods Enzymol.* **154:** 329–350.
10. Loeb, D.D., R. Swanstrom, L. Everitt, M. Manchester, S.E. Stamper, and C.A. Hutchinson III. 1989. Complete mutagenesis of the HIV-1 protease. *Nature* **340:** 397–400.
11. Regier, D.A. and R.C. Desrosiers. 1990. The complete nucleotide sequence of a pathogenic molecular clone of simian immunodeficiency virus. *AIDS Res. Hum. Retroviruses* **6:** 1221–1231.
12. Mori, K., D.J. Ringler, T. Kodama, and R.C. Desrosiers. 1992. Complex determinants of macrophage tropism in env of SIV. *J. Virol.* **66:** 2067–2075.

Received December 14, 1992; accepted in revised form on February 22, 1993.