Correction. In the article "A model for the origin of biological catalysis" by D. L. Stein and P. W. Anderson, which appeared in number 6, March 1984, of Proc. Natl. Acad. Sci. USA (81, 1751-1753), the authors' affiliations and the contributed line were omitted. They are as follows:

D. L. Stein* and P. W. Anderson*†

*Department of Physics, Princeton University, Princeton, NJ 08544; and †Physical Research Laboratories, AT&T Bell Laboratories, Murray Hill, NJ 07974

Contributed by P. W. Anderson, November 16, 1983

Correction. In the article "Enzymatic techniques for the isolation of random single-base substitutions in vitro at high frequency" by Patricio Abarzúa and K. J. Márians, which appeared in number 7, April 1984, of Proc. Natl. Acad. Sci. USA (81, 2030-2034), the designation of the symbols in the legend to Fig. 5 was reversed. The correct designation is (○) without added ssb and (●) with added ssb.
Enzymatic techniques for the isolation of random single-base substitutions in vitro at high frequency

(Escherichia coli exonuclease III/heteroduplex/targeted misincorporation/φX174-type single-stranded circular DNA synthesis)

PATRICIO ABARZÚA AND K. J. MARIANS

Department of Developmental Biology and Cancer, Division of Biological Sciences, Albert Einstein College of Medicine, Bronx, NY 10461

Communicated by Jerard Hurwitz, December 27, 1983

ABSTRACT A general and efficient method has been developed to generate large numbers of single-base substitution mutations simply and rapidly. A unique f1 phage recombinant DNA cloning vector is described, which contains the φX174 origin of viral strand DNA synthesis and allows one to direct mutagenesis to any specific segment of DNA. Gapped circular DNA is constructed by annealing single-stranded circular DNA [ss(c) DNA] with a mixture of linear duplex DNAs that have had their 3'-OH termini processively digested with Escherichia coli exonuclease III under conditions in which the resulting, newly generated 3'-OH termini present in the various hybrid molecules span the region of interest. Base changes are induced by misincorporation of an α-thiodideoxynucleoside triphosphate analog onto this primer-template, followed by DNA repair synthesis. The asymmetric segregation of mutants from wild-type sequences is accomplished by double-stranded replicative form DNA → ss(c) DNA synthesis in vitro, initiated from the φX174 viral strand origin sequence present on the vector DNA. Mutated ss(c) DNA is screened by the dideoxy chain termination method. In one mutagenesis experiment, 21 independent single-base substitutions were isolated in a 72-nucleotide-long target region. DNA sequence analysis showed that all possible base transversions and transitions were represented.

Recently, in vitro mutagenesis techniques have entered the repertoire of standard laboratory practices. There are, however, still some drawbacks to the most popular techniques. When the exact position of a nucleotide change required to generate a mutant with a particular phenotype is unknown, or when many base changes need to be introduced over a wide area, oligonucleotide directed mutagenesis (1–3) becomes technically difficult. Similarly, deletion loop mutagenesis induced by bisulfite (4) is likely to yield multiple-base substitutions when a large single-stranded loop is used. In addition, the type of mutation induced by this technique is limited.

Here the development of a mutagenesis procedure that combines the use of a unique recombinant DNA cloning vector with the use of several different enzymatic reactions is described. This technique can be used to induce, at very high efficiency, single-base changes in a predetermined region of DNA of up to several hundred nucleotides in length. This method has been used successfully to generate a library of 21 independent single-base substitutions within a 72-nucleotide-long site from a single mutagenesis experiment.

MATERIALS AND METHODS

Bacteria and Cloning Vectors. The Escherichia coli strain K38 (HfrC) was used for phage growth and for transformation. φ1R229, which contains a single EcoRI site within the intergenic space (5), is the parent vector of f1YA-5, which contains the pBR322 H-strand Y DNA site (6).

Deoxynucleoside Triphosphate Derivatives and Enzymes. 2'-Deoxyadenosine 5'-O-[(1-thiotriphosphatate) (dATP[αS]) and thymidine 5'-O-[(1-thiotriphosphatate) (dTTP[αS]) were the gifts of S. Benkovic (7). In the experiments reported here, the A isomer of dATP[αS] was used, whereas the preparation of dTTP[αS] contained both the A and B isomers. [α-32P]Deoxynucleoside triphosphates were purchased from Amersham; deoxynucleoside triphosphates were from P-L Biochemicals. T4 polynucleotide ligase was obtained from Bethesda Research Laboratories; the large fragment of E. coli DNA polymerase I was from Boehringer Mannheim; E. coli exonuclease III (exo III) was purchased from P-L Biochemicals; all restriction enzymes were obtained from New England Biolabs. E. coli DNA polymerase I, the φX174 gene A protein, the E. coli rep protein, E. coli DNA polymerase III*, and E. coli single-stranded DNA binding protein (ssb) were purified as described (8–13). The E. coli dnaN gene product (14) was purified by an unpublished procedure, as were the subunits of E. coli DNA gyrase (M. Gelbert, personal communication). Units are those of the manufacturer for commercial preparations or as those published for all other preparations.

Preparation of DNAs. f1 single-stranded circular DNA [ss(c) DNA] and double-stranded DNA in the replicative form (RF DNA) were prepared as described (15). The recombinant phage f1YA-68 was constructed by cloning the φX174 origin of viral strand DNA replication into f1YB-5 RF DNA (6). The φX174 Hae III b restriction fragment carrying EcoRI linkers was excised from recombinant C39-27, a pBR322 DNA carrying two φX174 viral strand origins of DNA replication (16), and ligated with f1YA-5-RF DNA that had been partially digested with the EcoRI restriction enzyme. The ligation mixture was used to transform E. coli K38 directly and plaques were obtained by plating with E. coli K38 indicator bacteria. Recombinant phage DNAs containing the φX174 viral strand origin of DNA replication were identified by plaque hybridization using nick-translated φX174 RF1 DNA as a probe; restriction enzyme analysis of the recombinant RFI DNA confirmed the orientation of the inserted fragment. In f1YA-68 DNA, the φX174 (+) strand is contiguous with the f1 viral strand and the pBR322 H strand and is upstream from the replication factor Y DNA site.

Digestion of DNA with exo III and Heteroduplex Formation. f1YA-68 RFI DNA was linearized by digestion with the Pvu II restriction enzyme for 3 hr at 37°C. The reaction was stopped by the addition of EDTA, followed by phenol-extracted DNA.
traction and ethanol precipitation. exo III digestion of the linear duplex DNA was performed in 250 μl of 66 mM Tris-HCl, pH 7.9/5 mM MgCl₂/10 mM dithiothreitol/90 mM NaCl at 37°C for 30 min. After phenol extraction and ethanol precipitation, the exo III-treated DNA was mixed with a 10- to 20-fold molar excess of f1YA-68 ss(c) DNA and alkali denatured for 2 hr at 37°C in 0.3 M NaOH. The solution was then neutralized with HCl and adjusted to 0.5 M NaCl/100 mM Tris-HCl pH 7.5/1 mM EDTA/0.05% Sarkosyl. The DNAases were annealed by heating the mixture to 100°C for 2 min, followed by slow cooling to room temperature. The DNA was precipitated with ethanol, resuspended in TEN buffer (10 mM Tris-HCl, pH 7.5/1 mM EDTA/10 mM NaCl), loaded onto an 11-ml 5–20% neutral sucrose gradient in TEN buffer containing 1 M NaCl and 0.05% Sarkosyl, and sedimented at 29,000 rpm for 18 hr at 4°C in a Beckman SW 41 rotor. Fractions (0.15 ml) were collected from the bottom of the tube. Aliquots of each fraction were electrophoresed through a 0.8% agarose gel to identify the contents containing the gapped circular DNA. These fractions were then pooled, dialyzed against TEN buffer and concentrated by ethanol precipitation. Contamination with ss(c) DNA was <1%.

Materials

A microgram of gapped circular DNA was incubated for 12–18 hr at 15°C in a 100-μl reaction containing 50 mM Tris-HCl (pH 7.5), 0.4 mM MnCl₂, 10 mM 2-mercaptoethanol, 100 μg of bovine serum albumin per ml, and either 100 μM ATP[α-S] or 200 μM dTTP[α-S]. The reaction was stopped by adding the mixture to 10 mM EDTA, followed by phenol extraction and ethanol precipitation. The DNA pellet was resuspended in 10 μl of 10 mM Tris-HCl (pH 7.5).

DNA Repair Synthesis. DNA recovered from the misincorporation reaction described above was gap-repaired by incubating with 1 unit of DNA polymerase I and 0.5 unit of T4 DNA ligase in a 50-μl reaction mixture containing 50 mM Tris-HCl (pH 7.5), 1 mM MgCl₂, 2 mM MnCl₂, 10 mM 2-mercaptoethanol, 100 μg of bovine serum albumin per ml, 0.5 mM ATP, and 100 μM (each) of the four deoxyribonucleoside triphosphates for 18–24 hr at 0°C. EDTA was then added to a concentration of 10 mM and an aliquot of the mixture was electrophoresed through a 0.8% agarose gel containing 0.5 μg of ethidium bromide per ml and run in 50 mM Tris-HCl/40 mM sodium acetate/1 mM EDTA, pH 7.9/0.5 μg of ethidium bromide per ml to determine the percentage of molecules converted to a completely closed form.

Superc elongation of Relaxed DNA. The mutated relaxed covalently closed circular RF DNA (RF DNA) was supercoiled in a reaction mixture (40 μl) containing 25 mM Tris-HCl (pH 7.5), 25 mM potassium phosphate buffer (pH 7.5), 10 mM MgCl₂, 5 mM dithiothreitol, 5 mM spermidine, 50 μg of bovine serum albumin per ml, 100 μg of trNA per ml, 0.25–0.5 μg of RFI DNA, and 3 units of DNA gyrase reconstituted from purified subunits. Incubation was at 30°C for 45 min. The reaction was terminated by the addition of excess EDTA, followed by phenol extractions.

The DNA was then isolated by ethanol precipitation and resuspended in 10 mM Tris-HCl, pH 7.5/0.1 mM EDTA.

RF → ss(c) DNA Synthesis with Purified Proteins and Purification of ss(c) DNA. Reaction mixtures (50 μl), containing 50 mM Tris-HCl (pH 7.5), 10 mM MgCl₂, 10 mM dithiothreitol, 4 mM ATP, 40 μM [α-32P]dCTP (300–1000 cpm/ml), 1 Ci = 37 GBq), 40 μM (each) dGTP, dATP, and dTTP, 10 μg of rifampicin per ml, 200 μg of bovine serum albumin per ml, 6% sucrose, 0.1–0.25 μg of RFI DNA, 0.32 unit of 43X47 gene A, 0.26 unit of E. coli tRNA, 0.4 unit of DNA polymerase III elongation system, and 4 μg of ssb, were incubated at 30°C for 60 min. Aliquots (1 μl) were processed periodically to determine the extent of formation of acid-insoluble radioactivity. When the reaction was complete, EDTA, protease K, and NaCl were added to 10 mM, 0.1 μg/ml and 0.1%, respectively, and the mixture was incubated at 37°C for 90 min, followed by phenol extraction and ethanol precipitation. The DNA pellet, after washing twice with ethanol, was resuspended in TEN buffer and applied to a neutral sucrose gradient (5–20%) formed in TEN buffer containing 1 M NaCl and 0.05% sarcosyl. Centrifugation was at 25,000 rpm for 15 hr at 4°C in a Beckman SW 50.1 rotor. Fractions were collected from the bottom and the peak of ss(c)125PJDNA was determined by Cerenkov radiation. The ss(c) DNA was pooled, dialyzed against TEN buffer and used to transform E. coli K38.

RESULTS AND DISCUSSION

This laboratory has been studying the interaction of E. coli replication factor Y with short specific DNA sequences from the viral (+) strand of φX174 and the H and L strands of plasmid pBR322 (19). Although these DNA regions are functionally identical, they exhibit no extensive sequence homology (15). To understand, at the nucleotide level, the interaction of factor Y with its effector sites, the construction of libraries of single-base substitutions along these DNA effector sequences has been initiated. To do so, a site-specific mutagenesis procedure has been developed that could be applied to any given DNA segment such that every nucleotide within the target sequence would have the same probability of being mutated. In addition, because the isolation of both inactive and silent mutations was of interest, the method had to be efficient enough to allow rapid screening of the mutated DNAs by DNA sequence analysis techniques. The use of this method to direct the in vitro mutagenesis of the 72-nucleotide-long H-strand factor Y effector site from pBR322 DNA is described in this report. Characterization of the mutant phenotypes will be published elsewhere.

Summary of the Technique. The basic steps of the mutagenesis method are illustrated in Fig. 1. RFI DNA, linearized with a restriction endonuclease that cuts once near the region of interest, is synchronously and processively digested with exo III. The resulting digest is annealed to excess ss(c) DNA and the isolated heteroduplexes are mutagenized by the forced misincorporation (20) of an excision-resistant α-thiodeoxynucleoside triphosphate (21), followed by the repair of the single-strand gap. The RFI DNA is then supercoiled and used as a template for φX174-type RF → ss(c) DNA synthesis. The purified ss(c) DNA, representing a nonsegregatable mutant population, is used to transform E. coli K38. Plaques are picked at random, a small amount of viral DNA is prepared, and a single-lane dideoxy sequence analysis reaction is performed for each DNA to screen for base substitution.
exo III Digestion of DNA. Linear duplex DNA was found to be a much better substrate than uniquely nicked circular RF DNA for the processive and synchronous digestion by exo III. To determine the rate of digestion by this enzyme, DNA linearized by the Pvu II restriction enzyme was digested with exo III and samples were taken at several points during the reaction. These samples were end-labeled by DNA polymerase I (Klenow fragment) in the presence of two $^{32}$P-labeled deoxynucleoside triphosphates. After inactivation of the polymerase at 65°C for 10 min, a second restriction enzyme, which cuts the DNA asymmetrically, was added and the incubation was continued at 37°C. The products were isolated and electrophoresed through a 10% sequencing gel. Fig. 2 shows a comparison of the rate of exo III digestion at 23°C and at 37°C. Eight to ten nucleotides per min were removed from each 3'-OH termini at 23°C, in agreement with previous reports (22, 23), whereas at 37°C the rate was 10-fold higher. Though exo III is a 3'→5' exonuclease with some sequence specificity (24), it has been shown that, depending on the extent of the digestion, molecules having 3'-OH ends of every length within a given range can be obtained (23). However, we have found that at 23°C exo III appears to slow down or stop before entering regions of the DNA that have the potential to form secondary structure. The intense bands evident in Fig. 2a after 5 min of digestion, which are still the major products even after 40 min of digestion, correspond to the base of the hairpin that the H-strand Y DNA site can potentially form. This effect is less pronounced at 37°C, though the rate of digestion is much faster (Fig. 2c). This phenomenon is not observed at the other end of the molecule (Fig. 2b) and cannot be overcome by adding more enzyme (data not shown). Therefore, the optimal conditions of digestion to generate 3'-OH ends at every position within the target site must be determined for each particular sequence. A complete discussion of the characteristics of exo III digestion of linear duplex DNA can be found in ref. 23.

Mutagenesis of the H-Strand Y Site from pBR322 DNA. Heteroduplexes constructed by annealing the linear duplex DNA with shortened 3'-OH termini to an excess of ss(c) DNA were used as primer templates for misincorporation of either dTTP[αS] or dATP[αS] onto the 3'-OH end, followed by repair synthesis and ligation. The overall extent of the reaction, assayed by conversion of the input DNA to a

FIG. 1. Summary of the mutagenesis procedure. The construction of the recombinant phage f1YA-68 is described in the text. Stippled box, φX174 Hae III 6b restriction fragment containing the viral strand origin of DNA replication; black box, the Rsa I–Hae III restriction fragment from pBR322 DNA containing the H-strand factor Y effector sequence; v, EcoRI site; v, Pvu II site.

FIG. 2. exo III digestion of linear duplex DNA. f1YA-68 RF2 DNA was linearized by digestion with the Pvu II restriction enzyme. As shown schematically in Fig. 1, the target site (the H-strand factor Y effector site) is on the right-hand side of the linear DNA. Aliquots were taken at the indicated times during incubation with 40 units of exo III per pmol of DNA (as molecules) at either 23°C (a and b) or at 37°C (c) and were end-labeled by incubation with the Klenow fragment of DNA polymerase I in the presence of [α-32P]dCTP and [α-32P]dTTP. The DNA used in the experiment shown in a and c was then digested with Acc I, which gave 32P-labeled fragments of 179 base pairs (bp) and 6388 bp from the right and left sides of the intact linear RF DNA (RFIII DNA), respectively. The DNA used in the experiment shown in b was digested with HpaI, which gave 32P-labeled fragments of 320 bp and 5753 bp from the left and right sides of the intact RFIII molecule, respectively. The samples were then denatured and electrophoresed through a 10% polyacrylamide sequence analysis gel. The positions of size markers are indicated on the autoradiogram.
closed circular form, was >90%. To keep the frequency of mutagenesis as high as possible, the mismatch repair mechanisms of the host were circumvented by segregating the mutant from wild-type sequences in vitro. Thus, the mutagenized DNA was used as a template for ss(c) DNA synthesis in vitro, catalyzed by the resident φX174 viral strand origin sequence. The φX174 viral strand origin of DNA replication, located within the Hae III 6b restriction fragment of φX174 RFI DNA, is recognized and cleaved specifically, when it is in a supercoiled form, by the φX174 gene A protein, and, if supplemented with the E. coli rep protein, E. coli DNA polymerase III elongation system, and ssb, can be used to catalyze the synthesis of single-stranded copies of any DNA into which it has been cloned (16, 25). Thus, the mutagenized RFI DNA was supercoiled by DNA gyrase and then inoculated with these four proteins under conditions in which 8–10 rounds of replication occurred. Only one of the strands [the (−) strand] serves as template in this particular mechanism of DNA replication. Thus, the segregation of sequences is asymmetric and, therefore, every misincorporation event will give rise to a base substitution within the progeny ss(c) DNA.

After transformation of E. coli K38 with the synthetic ss(c) DNA (1–5 × 10^6 plaque-forming units/μg of DNA), plaques were picked at random and viral DNAs were prepared for DNA sequence analysis. Screening was carried out by either the ddATP or ddTTP chain-terminating DNA sequence analysis reaction (18). Therefore, any base substitution arising as a result of misincorporation of either dATP[os] or ddTTP[os] will give rise to an extra band in the sequence analysis lane on the gel. A typical screening gel is shown in Fig. 3. Five of the 12 DNAs had base substitutions of the type expected and 1 unexpectedly had a missing band. Complete sequence analysis of these 6 DNAs confirmed that all of them had a single-base change from the wild-type sequence, 5 were within the limits of the target region, while the sixth (lane 6) was two nucleotides away from the 5' end of the site. Although a statistical analysis of the frequency of base substitution cannot be done with this method, the frequency is high enough to encourage mutant screening by biochemical assays. Other lesions, such as unexpected nucleotide changes similar to the one shown in Fig. 3, and double substitutions have also been observed. Their frequency, however, is rather low (<2%) and they probably originate from misincorporation of either a normal nucleoside triphosphate or an α-thiooxygenucleoside triphosphate analog during repair synthesis. However, double-base substitutions in adjacent positions, a phenomenon that has been reported to occur when misincorporation is induced by the error-prone avian myeloblastosis virus reverse transcriptase in conditions of infinite bias pool (26), or deletions or insertions, lesions commonly induced by the mismatch repair mechanism of the host cell (20) have never been observed.

A summary of single-base substitutions isolated from a single preparation of exo III-treated primer-template is shown in Fig. 4. Because every mutagenized DNA molecule undergoes several rounds of replication during in vitro DNA synthesis, almost every mutant shown here has been independently isolated more than once. Eighteen nucleotide positions (25% of the target sequence) have been mutated by misincorporation of an α-thiooxygenucleoside triphosphate analog onto a family of 3'-OH ends “predetermined” by exo III digestion. All possible transversions and transitions are represented and, even though all of the α-thiooxygenucleoside triphosphate analogs have not been used, it seems reasonable to expect that all types of single-base substitutions can be forced by α-thiooxygenucleoside triphosphate misincorporation.

RF → ss(c) Replication of Mutated RFI DNAs in a Crude Extract. To make this method less dependent upon the avail-

---

**Fig. 3.** Autoradiogram of a screening gel. Twelve phage plaques resulting from a mutagenesis experiment using dATP[os] were picked at random and ss(c) DNA was prepared as described (15). t1R229 RF DNA linearized by the EcoRI restriction enzyme was annealed to the viral DNA and the resulting heteroduplex was used as primer-template for the ddATP chain termination reaction (18). The dideoxy-terminated DNA products were digested with the EcoRI restriction enzyme. The samples were denatured and then electrophoresed through a 10% polyacrylamide gel containing 50% urea run in 50 mM Tris borate pH 8.3/1 mM EDTA. The wet gel was exposed directly to x-ray film for 10–14 hr at room temperature. The arrows indicate the limits of the target sequence. Lanes 2, 4, 5, 7, and 10 show an extra band; lane 6 has a missing band; however, no shift in the sequence pattern was observed; therefore, it is not a deleted nucleotide; instead an unexpected base substitution has occurred.

---

**Fig. 4.** DNA sequences of constructed point mutants. The nucleotide sequence of the H-strand Y site from pBR322 DNA is displayed on the top line (15). The remaining lines are the sequences of the mutants. Dashed lines indicate the wild-type nucleotide. A total of 124 plaques resulting from mutagenesis of a single preparation of exo III-treated primer-template were screened. Forty-seven isolates had single-base substitutions within the target sequence; 21 of these were independent events, representing 18 different positions of the target sequence.
cells (dnaB ts), when supplemented with φX174 RFI DNA and the φX174 gene A protein, support RF → ss(c) DNA synthesis via a rolling circle mechanism similar to the one observed with purified proteins (17), although the yield of ss(c) DNA is roughly 1/10th of that obtained using purified proteins. However, it was found that if one supplemented the crude extract with ssb, the only protein that acts stoichiometrically in this reaction, the production of ss(c) DNA could be increased dramatically (Fig. 5). DNA mutagenized and supercoiled as described above was replicated in a heat-inactivated BT1029 ammonium sulfate receptor fraction supplemented with the φX174 gene A protein and ssb. The ss(c) DNA produced was isolated and transfected into E. coli K38 as described above. The frequency of base substitution in the ss(c) DNA progeny was found to be similar to that of DNA replicated with purified proteins (data not shown). The possibility of using a crude extract prepared from BT1029 cells that have been transformed with a low-copy-number plasmid bearing the ssb gene, thus supplying the system in vivo with the only limiting factor, must be explored. Nonetheless, this protein is now commercially available and is easy to purify from overproducer strains. Therefore, the φX174 gene A protein is left as the only purified component that is not widely disseminated that is absolutely required for the system to work.

In conclusion, the mutagenesis procedure described here can be used to induce single-base substitutions at very high efficiency in any DNA segment cloned into this single-stranded DNA phage vector. Exo III digestion, used to generate an assortment of 3'-OH termini that can act as primers along the target sequence, can be initiated from any unique restriction site as far as 1000 nucleotides away (22). However, if such a unique restriction site is not available, a purified restriction fragment that spans the target region can be used as a substrate for exo III digestion. The resulting digest, when annealed to ss(c) DNA, would approximate the heteroduplex primer-template DNAs used in the study reported here. Although there are not, presently, any experimental data that can be used to determine the limit of the target site that can be mutagenized by this technique, it should be possible to induce single point mutations within a predetermined region of DNA of up to several hundred nucleotides in length. This site-specific mutagenesis procedure should prove very useful in the study of protein-DNA interactions, where it is difficult to predict in advance what base should be altered. In addition, this method could readily be adopted to isolate conditional missense and nonsense mutants by using appropriate conditions and strains.

These studies were supported by National Institutes of Health Grant GM31624. K.J.M. is the recipient of an American Cancer Society Faculty Research Award (FRA261).