Proc. Natl. Acad. Sci. USA
Vol. 75, No. 9, pp. 4268–4270, September 1978
Biochemistry

Efficient correction of a mutation by use of chemically synthesized DNA

(DNA sequence changes/triester DNA synthesis/heteroduplex synthesis *in vitro*/bacteriophage ϕX174 *am*3 mutant)

AHARON RAZIN*, TADAAKI HIROSE, KEIICHI ITAKURA, AND ARTHUR D. RIGGS

Division of Biology, City of Hope National Medical Center, Duarte, California 91010

Communicated by Ernest Beutler, July 3, 1978

ABSTRACT **The mutated base in the *am*3 lysis-defective mutant of the bacteriophage ϕX174 has been corrected by a combined *in vitro* enzymatic DNA synthesis and *in vivo* replication of the heteroduplex product. Chemically synthesized oligodeoxyribonucleotides carrying the wild-type sequence have been used to prime DNA synthesis with *am*3 ϕX174 DNA serving as a template. The resultant semisynthetic heteroduplex composed of an *am*3 (+) strand and a wild-type (−) strand, with one mismatched base pair at position 587 on the ϕX174 DNA sequence, was used to infect spheroplasts. The progeny phage were analyzed by a parallel plaque assay on wild-type host, *Escherichia coli* C, to screen for wild-type phenotype, and on *E. coli* HF4714, an amber suppressor strain, to determine the total progeny phage. When a 23-base-long synthetic primer was used, about one-third of total progeny were found to be wild type. Shorter primers yielded lower percentages of wild type; they also had poorer priming activity.**

The ability to specifically and efficiently substitute one or a few bases in a DNA sequence should greatly aid numerous genetic manipulations. Such a procedure would allow, for example: repair or creation of mutations; conversion of a gene of one species to the same gene of another species; and creation or elimination of special sequences such as restriction sites. In the present study, the *am*3 lysis-defective mutant of the bacteriophage ϕX174 has been used to test the feasibility of specifically changing one base in a DNA sequence. The *am*3 mutation was chosen because large amounts of DNA carrying it can be prepared easily, the complete DNA sequence of ϕX174 is known (1), the precise location of the mutation has been established (2), and the mutant progeny can be distinguished easily from wild type.

The strategy used in the present experiments involved the chemical synthesis of an oligodeoxyribonucleotide complementary to the wild-type DNA at the mutation region (2) (see Fig. 1) and its annealing to *am*3 ϕX174 single-stranded circular DNA. This partial duplex with one mismatched base pair was used for primed repair synthesis *in vitro* by *Escherichia coli* DNA polymerase I followed by ligation of the product to form a covalently closed circular duplex (RFI). These duplex molecules are infective to spheroplasts and genetically heterozygous, composed of an *am*3 mutant (+) strand and a wild-type (−) strand.

We report here that a substantial fraction of the progeny produced after infection of spheroplasts was found to bear the wild-type marker derived from the chemically synthesized oligonucleotide. The "salvage of mutations" of ϕX174 by hybridizing wild-type DNA with mutant DNA has been reported before (3–6). Results similar to those presented here were achieved with enzymatically synthesized primers (7).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

The bacteria and phage strains used in this study were: *E. coli* C (wild type), HF4714 (carrying a suppressor), *E. coli* W6; bacteriophage ϕX174 wild type and *am*3cs70 (a gene *E* mutant). All were obtained from R. L. Sinsheimer.

Preparation of DNA Substrates. Bacteriophage ϕX174 DNA was prepared as described (8). The oligodeoxyribonucleotides were synthesized by the modified triester method (9–11), which leaves free 5′-OH ends. The oligonucleotides were phosphorylated at the 5′ end by phage T4 polynucleotide kinase (Miles Laboratories) and [γ-^{32}P]ATP as described (12).

RFI Synthesis *In Vitro*. RFI was synthesized *in vitro* by using ϕX174 *am*3 DNA as template and a chemically synthesized oligonucleotide as primer. The reaction mixture contained the following ingredients in a final volume of 100 μl: 67 mM KPO_4 at pH 7.7; 6.7 mM $MgCl_2$; 1 mM 2-mercaptoethanol; 0.2–0.4 μg of template DNA; 10-fold molar excess of primer over template; 2.5 nmol each of dCTP, dGTP, dATP, and dTTP; 0.05 nmol of [^{3}H]dTTP (15 mCi/μmol); 2 units of phage T4 DNA ligase (a gift from H. Heyneker); and 0.3 unit of *E. coli* polymerase I (Worthington Biochemical Corp.) The reaction mixture was incubated at 12°C for 6 hr. To follow the kinetics of the synthesis, 5-μl aliquots were withdrawn and precipitated along with 20 μg of calf thymus DNA as carrier by trichloroacetic acid. The precipitated DNA was washed on GF/C filters with 5% cold trichloroacetic acid and its radioactivity was measured. The [^{3}H]dTTP incorporation usually started without a significant lag and leveled off after 4–6 hr of incubation.

Spheroplast Infection by Synthetic DNA. Spheroplasts were prepared from *E. coli* W6 cells as described by Guthrie and Sinsheimer (13). Twenty-five microliters of the original polymerase I reaction mixture was diluted to 0.4 ml with 50 mM Tris·HCl, pH 8.1, and mixed with 0.4 ml of spheroplast "stock." The infected mixture was incubated for 15 min at 33°C and 3.2 ml of PAM medium (13) was added, and incubation was continued for 90 min. Spheroplasts were lysed by adding a drop of chloroform and vortex mixing. The lysate was diluted in 50 mM sodium tetraborate to the appropriate dilutions for plaque assay on *E. coli* C and *E. coli* HF4714.

RESULTS AND DISCUSSION

Three oligonucleotides, of chain lengths 11, 14, and 17, and complementary to *am*3 ϕX174 DNA except for one mismatched base pair (see Fig. 1), were synthesized by the chemical triester method (9). These were used as primers for repair synthesis by polymerase I as shown schematically in Fig. 2. T4 DNA ligase was present during synthesis to convert completed

Abbreviations: RFI, covalently closed circular duplex DNA; RFII, RFI with a single-strand break.

* Visiting scientist from the Department of Cellular Biochemistry, The Hebrew University, Hadassah Medical School, Jerusalem, Israel 91000.

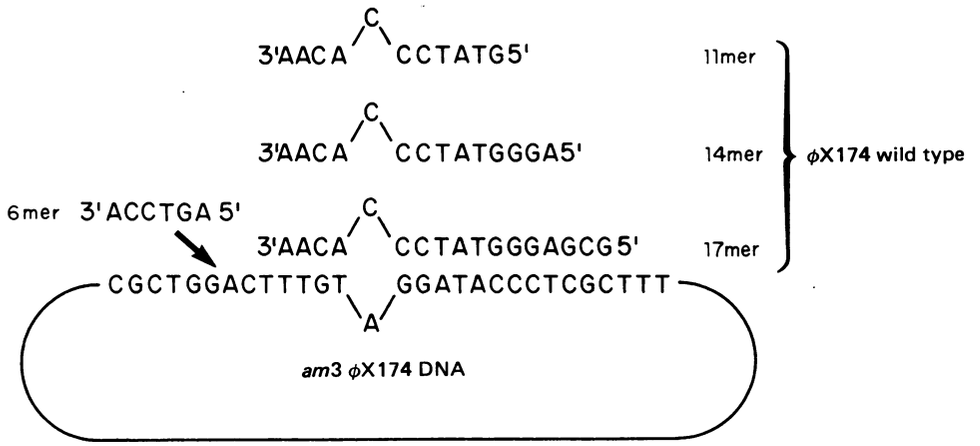

FIG. 1. The sequences of the oligodeoxyribonucleotides synthesized by the triester method (9–11) and the region in the φX174 (+) strand DNA where the amber mutation is located (position 587 on the φX174 DNA sequence).

but nicked duplexes (RFII) to their covalently closed circular form (RFI). In preliminary experiments the ligation was found to be essential; without T4 DNA ligase, no increase in wild-type progeny could be detected. The reason for this is believed to be nick translation (5′→3′ exonuclease) activity residing in the polymerase. Once the circle is completed, this activity will remove the primer with the desired wild-type DNA sequence.

Duplex phage DNA is infective to spheroplasts (14). Therefore, after infection of spheroplasts with heteroduplexes synthesized *in vitro*, replication *in vivo* will produce two phenotypes of progeny virus: *am*3 mutants derived from the template DNA and wild-type phage derived from the synthetic DNA primer (see Fig. 2). These progeny phage were scored by plaque assays on appropriate indicator strains.

With the experiment performed as described above, we observed a considerable production of wild-type phage (Table 1). However, even with a 17mer, the longest primer used, the yield of wild type was low (about 0.1%), despite the fact that the 17mer was a very potent primer. One likely explanation for the relatively low yield (0.1% compared to an expectation of about 50%) is that the 3′→5′ exonuclease activity of DNA polymerase I in concert with its synthetic activity, "corrects" the mismatched base in the primer, converting it to the undesirable *am*3 genotype. To reduce this editing action, we extended the 3′ end of the primer. An additional hexamer was synthesized chemically and joined to the primer by ligation on the *am*3 (+) strand template prior to DNA synthesis (see Fig. 1). As shown in Table 1, when these extended oligonucleotides are used as primers, they give rise to a severalfold increase in the yield of wild-type progeny. However the proportion of wild type remains low (<1%).

Analysis of the products of the combined polymerase I and ligase reaction on 1.2% agarose gels (Fig. 3) shows that about 50% of the original single-stranded DNA molecules are converted to RFI. The remaining single-stranded circular molecules carrying the amber mutation are 20-fold more infective than the RFI molecules (14) and will interfere with the assay by artificially lowering the value of the wild-type progeny when expressed as percent of total progeny. Therefore, a limited digestion with nuclease S1 was introduced to eliminate infection by single-stranded DNA molecules and incomplete duplexes. The reaction mixtures after polymerization and ligation were treated with 0.1% diethyloxydiformate (Eastman) for 10 min at 37°C to inactivate polymerase I and then incubated for 60 min at 37°C with an amount of nuclease S1 (Sigma) sufficient to cleave 99% of the single-stranded DNA molecules, as judged by infectivity to spheroplasts in control experiments. More extensive treatment was not attempted, to avoid the possibility of nicking at the mismatch. After this treatment, most of the infective molecules are RFI heteroduplexes composed of *am*3 (+) strand and wild-type (−) strand. When these molecules are used to infect spheroplasts, up to 32% of the progeny viruses are wild type (Table 1), the best yield again being obtained with the longest primer (17 + 6).

The single-burst experiments of Baas and Jansz (5) and Weisbeek and Van Arkel (6) indicated that about 70% of *am*3

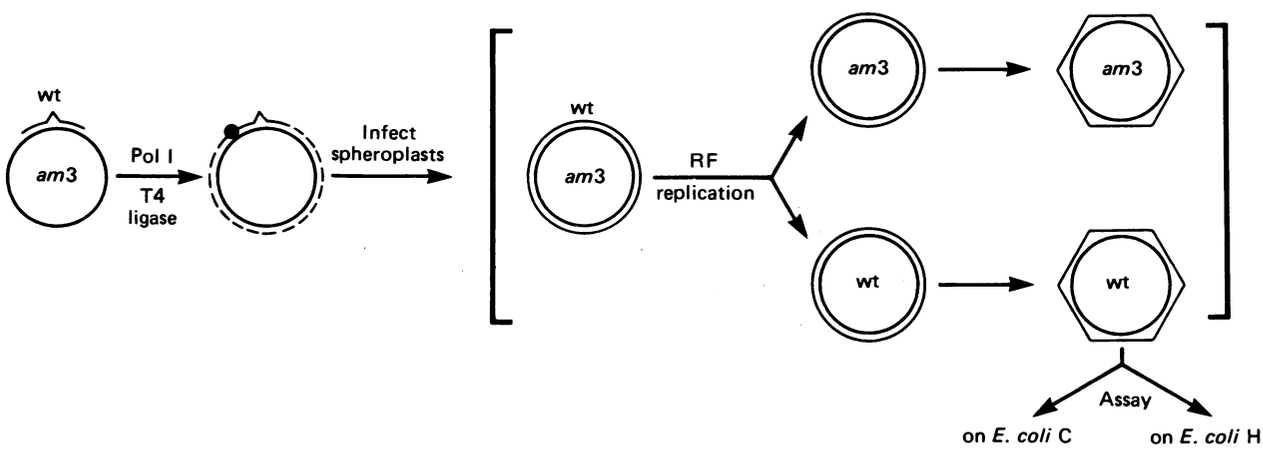

FIG. 2. Preparation of a semisynthetic RFI molecule composed of *am*3 (+) strand and synthetic wild-type (wt) (−) strand. The replication of these RFI molecules in spheroplasts results in production of two phenotypically distinct progeny phages. Pol I, DNA polymerase I.

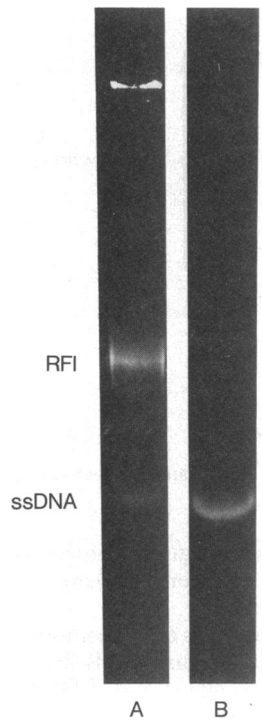

FIG. 3. Analysis by agarose gel electrophoresis of the products of the ligation and polymerase I reactions. A synthetic 5′-^{32}P-labeled 17mer was ligated for 12 hr at 12°C to a synthetic 5′-^{32}P-labeled 6mer with *am*3 (+) strands serving as templates. Then polymerase I and dNTP mix were added and incubation was continued for 6 hr at 12°C. An aliquot of the reaction mixture was mixed with an equal volume of 14% Ficoll 40/1% sodium dodecyl sulfate/0.4% bromophenol blue/0.05 M EDTA; the mixture was heated for 10 min at 60°C and applied to 1.2% agarose slab gel (Seakem, Rockland, ME). Electrophoresis was conducted at 7.5 V/cm for 180 min in 40 mM Tris acetate/5 mM sodium acetate/1 mM EDTA, pH 7.8. The gel was stained in a solution of 0.75 μg of ethidum bromide per ml for 30 min, then photographed. Lane A, product of *in vitro* primed DNA synthesis, lane B, single-stranded φX174 DNA (ss DNA) marker.

RFII heteroduplexes are repaired to homoduplexes *in vivo* prior to replication. We have observed that the burst size of *am*3 is 3 times larger than that of wild type. Therefore, the observed proportion of wild type is a minimum value for the successful *in vitro* synthesis of heteroduplex. A yield of 32% wild-type progeny thus suggests that any desired specific change in the base sequence of a natural DNA can be achieved efficiently by using primed synthesis with a suitable synthetic primer.

To obtain high correction efficiencies, the strand displacement, nick translation, and editing activities of DNA polymerase must be minimized. Nick translation (or strand displacement) was clearly a problem in our earlier experiments. However, nick translation is greatly reduced by using primers extended at their 5′ end (11mer → 17mer, see Fig. 1) together with T4 DNA ligase to seal the nick generated when the circle is completed. Editing activity (3′→5′ exonuclease) was also significant, cleaving back four base pairs and then eliminating the mismatch (see Fig. 1). Our solution to this problem was to place the mismatch 10 bases from the 3′ end. Thus, long primers are advantageous. As shown here, it is possible to make them by ligation of shorter oligonucleotides on the natural DNA as template. Ligation will proceed even with a nearby mismatch. With the new improvements in triester DNA synthesis methods (10, 11), the necessary primer oligonucleotides (about 21 bases total) can be made in a very short time.

Table 1. Yield of wild-type progeny phage after infection of spheroplasts with heteroduplex RFI composed of mutant (+) strand and wild-type (−) strand

Primer oligonucleotide*	Progeny wild type, % of total progeny†	
	No S1 nuclease treatment	S1 nuclease treatment
11mer	<0.02	<0.02
14mer	0.02	1.9
17mer	0.1	13.9
11mer + 6mer	<0.02	1.2‡
14mer + 6mer	0.06	14.0
17mer + 6mer	0.4	32.5

Ligation of different oligomers with 6mer and primed synthesis were as described in legend to Fig. 3. *am*3 φX174 single-stranded DNA control had less than 0.003% wild type. Minus polymerase controls after S1 treatment had no detectable wild type.

* See Fig. 1.

† In the case of no S_1 nuclease treatment, the total progeny plaque count was 2–7 × 10^7 plaques per μg of DNA. After S1 nuclease treatment, total progeny was 1–5 × 10^6 plaques per μg of DNA. The transfection efficiency of single-stranded DNA was 2 × 10^{-4} plaques per DNA molecule and 10^{-5} plaques per DNA molecule for RF DNA.

‡ Note that the 17mer generated by joining the 11mer with a 6mer is different from the 17mer made directly without ligation. The 11 + 6 primer has the mismatch located 10 nucleotides from the 3′ end but only 6 base pairs from the 5′ end. The 17mer primer has the mismatch 4 nucleotides from the 3′ end but 12 nucleotides from the 5′ end.

The efficiencies reported here are high enough that even sequence changes for which there is no method of selection can be made. Molecular cloning followed by direct DNA sequencing or other methods of specific sequence detection could be used to identify the desired phage or plasmid DNA molecules.

This work was supported by contract 729 from Genentech, Inc.

1. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., III, Slocombe, P. M. & Smith, M. (1977) *Nature (London)* **265,** 687–695.
2. Barrell, B. G., Air, G. M. & Hutchison, C. A., III (1976) *Nature (London)* **264,** 34–41.
3. Weisbeek, P. J. & Van de Pol, J. H. (1970) *Biochim. Biophys. Acta* **224,** 328–338.
4. Hutchison, C. A., III & Edgell, M. H. (1971) *J. Virol.* **8,** 181–189.
5. Baas, P. D. & Jansz, H. S. (1972) *J. Mol. Biol.* **63,** 569–576.
6. Weisbeek, P. J. & Van Arkel, G. A. (1976) *Virology* **72,** 72–79.
7. Hutchison, C. A., Phillips, S., Edgell, M. H., Gillam, S., Jahnke, P. & Smith, M. (1978) *J. Biol. Chem.* **253,** 6551–6560.
8. Razin, A., Sedat, J. W. & Sinsheimer, R. L. (1970) *J. Mol. Biol.* **53,** 251–259.
9. Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F. & Boyer, H. W. (1977) *Science* **198,** 1056–1063.
10. Hirose, T., Crea, R. & Itakura, K. (1978) *Tetrahedron Lett.* **28,** 2449–2452.
11. Crea, R., Hirose, T. & Itakura, K. (1978) *Tetrahedron Lett.*, in press.
12. Heyneker, H. L., Shine, J., Goodman, H. M., Boyer, H. W., Rosenberg, J., Dickerson, R. E., Narang, S. A., Itakura, K., Lin, S. & Riggs, A. D. (1976) *Nature (London)* **263,** 748–752.
13. Guthrie, G. D. & Sinsheimer, R. L. (1963) *Biochim. Biophys. Acta* **72,** 290–297.
14. Goulian, M., Kornberg, A. & Sinsheimer, R. L. (1967) *Proc. Natl. Acad. Sci. USA* **58,** 2321–2328.