Volume 10 Number 20 1982 Nucleic Acids Research

Directed mutagenesis of DNA cloned in filamentous phage: influence of hemimethylated GATC sites on marker recovery from restriction fragments

Wilfried Kramer, Klaus Schughart and Hans-Joachim Fritz

Institut für Genetik, Universität zu Köln, Weyertal 121, D-5000 Köln 41, FRG

Received 7 July 1982

ABSTRACT

Gapped duplex DNA molecules of recombinant genomes of filamentous phage are constructed *in vitro*. Denatured restriction fragments covering (part of) the precisely constructed gap are hybridized to the gapped duplex DNA molecules to form ternary duplices. The two strands of the ternary duplex molecules carry different genetic markers within the region spanned by the restriction fragment leading to a one base pair mismatch or to an insertion loop of 93 nucleotides, respectively. The two strands also vary with respect to A-methylation in GATC sites. In cases of asymmetrical methylation, transfection of *E.coli* with these heteroduplex molecules leads to marker recoveries with a pronounced bias in favour of the marker encoded by the methylated strand. This effect at least partly explains the comparably low marker yields achieved in previous directed mutagenesis experiments using filamentous phage as the vector. The results suggest how these procedures can be optimized. Precise construction of a 93 bp insertion in 9.5% marker yield is described.

INTRODUCTION

Chemical synthesis of polynucleotides of defined sequence and their application to genetic research were pioneered by H.G.Khorana. This tool has added to biology the central element of preparative chemistry: The exact chemical nature of the object of study is predetermined at will by the experimenter. This development culminated in the unraveling of the genetic code[1] and the first chemical synthesis of a gene fully functional *in vivo*[2].

An illustrative example of synthetic thinking applied to biological problems is the presently emerging technique of directed mutagenesis by marker salvage[3] from short (synthetic) polynucleotides. In this context the term "directed mutagenesis" implies specificity with respect to both the location of the mutation (at the nucleotide level) and the (predetermined) nature of the structural change induced in the target DNA.

Figure 1 illustrates the way the method is presently being used[5-17]; the basic principles were first outlined about ten years ago[3,4].

A second important feature of the method, besides its strictly controlled

0305-1048/82/1020-6475$ 2.00/0

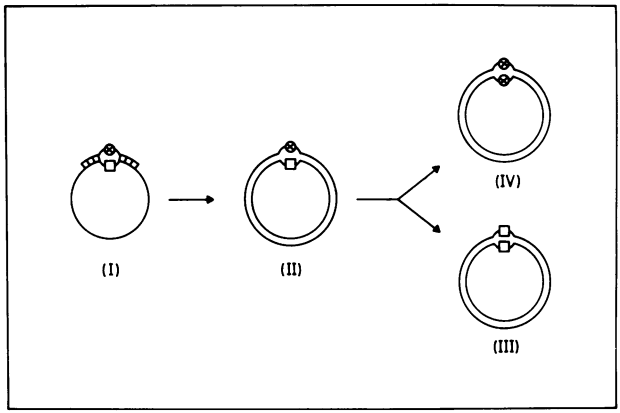

Figure 1:
Schematic outline of a typical directed mutagenesis experiment using a synthetic oligonucleotide: A synthetic primer containing a constructed structural deviation from wild type is annealed to a single-stranded wild type genome to give partial duplex (I) containing e.g. a single mismatch. Enzymatic *in vitro* reactions with DNA polymerase and DNA ligase yield a complete duplex (II), which upon transfection into *E. coli* host cells is replicated to fully homologous duplex molecules: (III), wild type and (IV), mutant type. Pure clones containing (III) or (IV), respectively, are obtained upon segregation.

chemistry, is the high frequency (in the percent range), with which the desired mutant type can be found in the progeny. This is especially important for the application of directed mutagenesis to e.g. eukaryotic gene fragments or control DNA segments, which cannot be expected to elicit easily observable phenotypes in *E. coli* host cells and may require DNA sequence analysis as the only possibility to identify a clone carrying the desired mutation.

In this situation, the amount of time and effort spent on screening individual clones will depend crucially on the abundance of the desired mutant type in the mixed progeny population as illustrated in figure 2. Marker yields as hitherto achieved with filamentous phage[13-17] and pBR 322[11] (typically less than 10%), either require the screening of dozens of clones: usually done by selective hybridization procedures if possible[11,16,17] or, alternatively, they call for biochemical marker enrichment (again resting on selective hybridization) prior to screening[9,10,15,16].

These methods suffer from several drawbacks: 1. The enrichment procedure consisting of repeated cycles of selectively primed *in vitro* synthesis of RF DNA plus subsequent transfection and phage DNA preparation is time consuming. 2. Optimal conditions for selective hybridization are different in each case

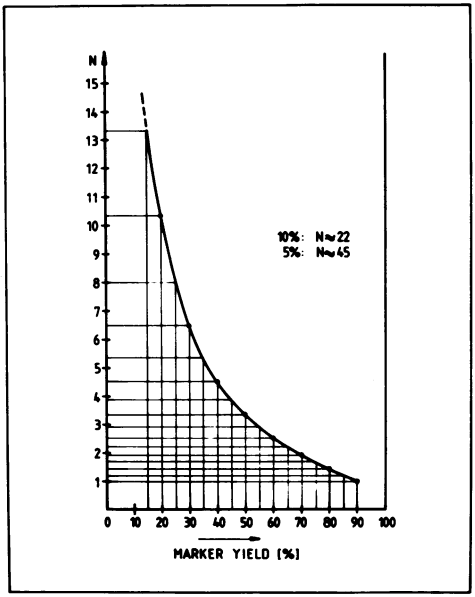

Figure 2:
Number, N, of candidate clones necessary to be screened for 90% probability to find the desired structure, plotted against marker yield (M.Y.) defined as percentage of those clones present in the progeny, which carry the constructed marker. N is calculated according to the equation:

$$N = \frac{-1}{\log\left(1 - \frac{M.Y.(\%)}{100}\right)}$$

and have to be worked out individually. 3. Even the most clear-cut hybridization response cannot be regarded as rigorous structural proof: DNA sequence analysis has to be executed additionally.

If marker yields could be increased to the range of 50 to 100%, then a generally applicable screening procedure could consist of sampling one or very few candidate clones at random and determining for each of them the nucleotide sequence of the DNA region in question.

This paper describes a model study that offers an explanation for the relatively low marker yields thus far achieved with the above mentioned vectors and demonstrates how marker recovery can be drastically increased. It also shows, that the marker salvage method can be used to precisely construct sizable DNA insertions in high yield.

MATERIALS AND METHODS

Bacterial strains: The following *E. coli* K12 strains were used: BMH 71-18 (Δ {lac-pro AB}, sup E, thi; F' lac I^q ZΔM15, pro A^+B^+) (source: B. Müller-Hill); GM 99 (F^-, mal 354, tsx, dam-4) (source: M. Marinus); Tn5-F1 (sm^r, Δ {nic-gal}, F' galT::Tn5 {kan^r}) (source: D. Pfeifer); GM 99-1 (F^-, mal 354, tsx, dam-4; pO1-14 {amp^r}) was derived from GM 99 by transformation with the plasmid pO1-14 carrying an ampicillin resistance marker; GM 99-F2 (mal 354,

tsx, dam-4; pO1-14{amp^r}; F' galT::Tn5{kan^r}) was constructed by F-crossing GM 99-1 with Tn5 - F1.

Phage and plasmid strains: M13mp2Z1, M13mp2/3Δ339 and M13mp3ΔM15 as well as an *E. coli* K12 strain harbouring pHBAB 339, a plasmid derived from pBR 322 and containing a 339 bp EcoRI/EcoRI fragment of the lacα-region were generous gifts of B. Gronenborn. For details of genome structures refer to figures 3 and 5 and description in results and discussion section.

Media: 2YT liquid medium and YT agar were prepared as described[18]. Other solid media had the following compositions: Standard top-agar: 10g Bacto-tryptone, 5g NaCl, 6.5g Bacto-agar, 1ml 1M $MgSO_4$, 1l H_2O; EHA-plates: 13g Bacto-tryptone, 8g NaCl, 10g Bacto-agar, 2g Na-citrate·$2H_2O$, 1l H_2O; Indicator plates: a) bottom-agar: YT agar containing IPTG at a concentration of 0.2mM; b) top-agar: 2.5ml standard top-agar (see above) per plate with 50µl of a 2% (wt/vol) solution of x-gal in N,N-Dimethylformamide added immediately before plating.

Enzymes and chemicals: Restriction enzymes were purchased from BRL and Boehringer, PEG 6,000, DEAE-cellulose, IPTG and x-gal from Serva. Hydroxyapatite was "DNA Grade Biogel HTP" from Bio-Rad. Low melting agarose was "Seaplaque Agarose" from Marine Colloids. Organic components of media were purchased from Difco, salts and buffers from either Merck or Serva.

Preparation of M13 RFI DNA (cccDNA): Recipient *E. coli* cells were grown in 2YT to an optical density (546 nm) of 0.8, then infected with M13 at a multiplicity of infection of ca 10, then grown for another 5 hours. Ccc DNA was prepared from lysozyme treated cells by CsCl gradient centrifugation according to Clewell and Helinski[19] with some minor modifications.

Preparation of M13 virion DNA (ss DNA): Phage infected cells were grown overnight (1l 2YT), and bacteria pelleted at 3,500 x g for 20 min. Phage particles were precipitated from the supernatant with PEG according to G. Heidecker et al.[20]. Half of the supernatant was decanted and the precipitate was pelleted at 16,000 x g for 30 min. The pellets were re-suspended in 40ml λ-buffer (1mM $MgSO_4$, 0.002% gelatine, 20mM sodium phosphate buffer pH 7.2) and the phage suspension centrifuged for 16h at 80,000 x g. The resulting pellets were re-suspended in 1-2ml λ-buffer. DNA was then isolated by phenol extraction and extensive dialysis against TES-buffer (50mM NaCl, 1mM EDTA, 10mM Tris-HCl, pH 7.2).

Construction of partial DNA duplex molecules with defined gaps was carried out as described[21]. In a typical experiment, 80 pmole of viral (+)strand were hybridized to 4 pmole of linearized RF DNA (obtained by restriction

cleavage as shown in figure 3) in a volume of 20 ml.

Preparation of DNA restriction fragments: Restriction fragments were separated by electrophoresis on 1.5% Seaplaque Agarose slab gels; the bands were made visible with ethidium bromide and cut out. The agarose was melted with 20 volumes of TES-buffer (see above) at 65°C and the DNA recoverd by chromatography on a column containing ca. 20 mg of DEAE-cellulose on a 10 mg bed of Sephadex G50 (P. Schreier, pers. comm.). Buffers: TES (see above) for sample application and washing; 1.5M NaCl, 10mM Tris-HCl, 1mM EDTA, pH 7.2 for elution.

Hybridization of restriction fragments to gapped DNA duplex molecules to form ternary duplices was generally carried out as follows: 200 fmole of restriction fragment in 1-2µl TES-buffer (see above) were sealed in a capillary tube and submersed in boiling water for 5 min., then immediately cooled to 0°C. The fragment was then mixed with a cold solution of gapped duplex (ca. 20 fmole in 20µl TES-buffer). The mixture was sealed again in a capillary tube, incubated for ca. 16 hours at 65°C in a water bath, then cooled to room temperature and used for transfection.

Transfection and segregation: Transfection of $CaCl_2$ treated bacteria (strain BMH 71-18) was carried out analogous to the procedure of Cohen et al.[22] with minor modifications. A 1/6 aliquot was removed immediately after the heat shock treatment and plated on EHA-plates (see above) to determine the total number of transfectants. Only experiments in which this number was in excess of 10^3 were followed up. The remainder was diluted with 25 ml 2YT and shaken overnight at 32 or 37°C. After sedimentation of the bacteria, phage particles in the supernatant were titrated by reinfecting strain BMH 71-18 cells at low multiplicity (ca. 10^{-5}, i.e. mixing 100µl of bacterial overnight culture with 100µl of phage suspension in λ-buffer diluted to 10^4 pfu per ml) and plating on indicator plates (see above). Phage carrying wild type or mutant lac(α) gene fragments respectively were scored as either blue or colourless plaques.

DNA sequence analysis was carried out as described[23] with minor modifications.

RESULTS AND DISCUSSION

Starting point of the present study was the observation that marker yields in experiments using ΦX 174 are markedly higher (up to 30 - 40%)[5,8] than with filamentous phage (typically less than 10%)[13,17]. It seemed to us that a possible explanation for this behaviour could be offered by the study of Radman et al.[24] on mismatch repair acting on duplex DNA molecules of the λ

genome containing one or two mismatch positions. In cases, were one strand of the heteroduplex was N^6-adenine methylated in its GATC sites and the other was not, these authors found highly preferential preservation of the genetic markers of the methylated strand. In directed mutagenesis experiments carried out according to figure 1, the same situation of hemimethylated GATC sites is created by *in vitro* DNA polymerase/DNA ligase reaction to ccc DNA, the methylation, however, being in the strand which contains the undesired original marker. The low marker yields with filamentous phage could thus be due to unfavourably oriented mismatch repair acting on the transfecting phage genomes before replication or methylation sets in. Since the ΦX 174 genome does not contain any GATC sites, marker yields achieved with this phage can be expected to be correspondingly higher (i.e. close to the theoretical value of 50%). The aim of our work was to investigate the influence of hemimethylated GATC sites on marker yields achieved with filamentous phage for both point mutation and deletion/insertion, and, if a significant influence could be found, to exploit it in a scheme to maximize the efficiency of directed mutagenesis.

In order to have a sound statistical basis for the assessment of mutational frequencies, it seemed desirable to screen large numbers of phages. Here the α-complementation brought about by a short N-terminal fragment of β-galactosidase encoded by a 789 bp fragment of lac DNA cloned in phage M13 as described by J. Messing et al.[25] proved to be very convenient. Wild type (α^+) and mutant alleles of the Z(α) gene fragment can easily be distinguished on the basis of plaque colour on x-gal indicator plates using an appropriate host strain[25].

All proposed structures of the mutant phages as well as those of wild type revertants were proven by DNA sequence analysis (data not shown).

In a first experiment we investigated the reversion of a poin mutation: Phage M13mp2Z1 carries the mutation Z1[26] in which codon number 23 of the lacZ gene (gln: CAA) is converted to a nonsense codon (ochre: TAA) (R. de la Haye and B. Gronenborn, unpublished). First, virion DNA of M13mp2Z1 was prepared from dam^+ and dam^- host strains and each hybridized to linear dam^+ RF DNA of phage M13mp2/3Δ339 as outlined in figure 3.

Two nearly identical gapped duplex DNA molecules differing only in the adenine methylation of the (+)strand resulted (for a review on the methylation of virion DNA from filamentous phage see K. Horiuchi et al.[27]). The genome of phage M13mp2Z1 contains 5 GATC sites: 3 in the M13 part[28] and two within the 339 bp stretch of the lac DNA insert, which remains single-stranded in the gapped duplex molecules (compare figure 4).

Two different fragments of wild type DNA were used for marker salvage: The

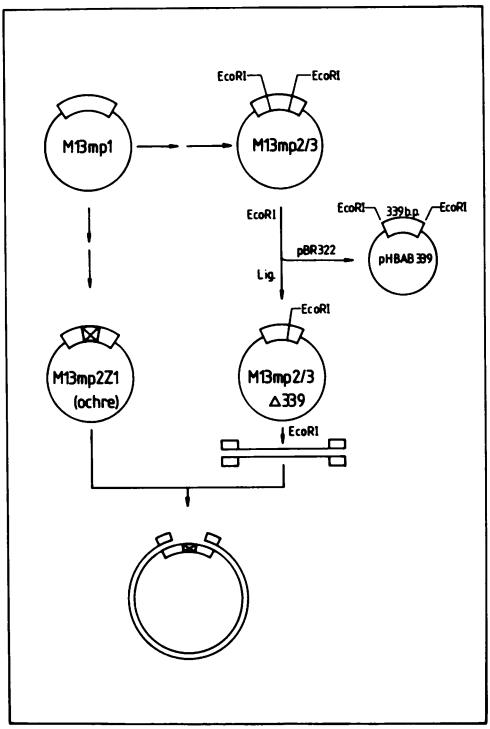

Figure 3:
Schematic diagram of construction of some replicons used in this study and of manipulations necessary to obtain a partial duplex DNA molecule with a defined gap, in which the position to be mutagenized is exposed within the single-stranded region.
M13mp2/3 differs from M13mp1 (ref. 25) by two G/C to A/T transitions creating two EcoRI restriction sites. M13mp2Z1 was constructed by recombination of M13mp2 (ref. 32) with an F'lac episome carrying the lacZ1 mutation (B. Gronenborn, unpublished) (Open bars represent cloned lac DNA, the crossed box symbolizes the ochre point-mutation).

339 bp EcoRI/EcoRI fragment of pHBAB 339 (see figure 3) that spans the entire single-stranded region of the gapped duplex molecules and the 92 bp EcoRI/AluI subfragment thereof, that covers only little more than the promoter proximal

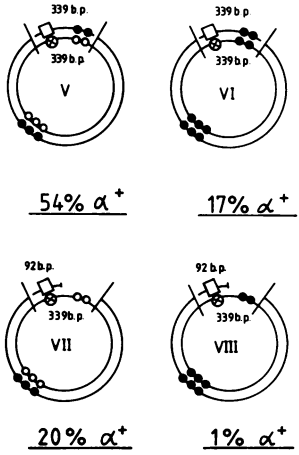

Figure 4:
Ternary duplices used for transfection in the point-mutation reversion experiment and percentage of wild type revertants found in the phage progeny. Open circles indicate unmethylated GATC sites, solid circles GATC sites carrying an N(6)-Me-adenine residue; crossed circles represent the Z1 ochre mutation, open boxes the respective wild type structure. Slanted lines crossing DNA strands indicate borders of single-stranded regions in the starting gapped duplex molecules. Marker yields are average values of several experiments, each one involving scoring of at least several hundred plaques.

fourth of the gap. These fragments were denatured, then annealed to the gapped duplex DNA and the resulting ternary duplex molecules (V - VIII in figure 4) used directly for transfection (fill-in and ligation reactions were omitted in order to keep the number of variables as small as possible).

Ternary duplices V and VI as a group give higher marker yields than VII and VIII. This could be explained on the basis of the longer fragment having a higher chance than the short one to be incorporated into a complete phage genome rather than being digested by exonuclease before the gap is filled and sealed *in vivo*.

Both sets of experiments display a marked influence of dam methylation on the marker yield (figure 4): Ternary duplex V, which has all the dam methylation in the (-)strand, gives more than three times the amount of reversion to α^+ as the symmetrically methylated duplex VI. This is in qualitative agreement with the analogous study by Radman et al.[24].

Gapped ternary duplex VII yields 20 times as many α^+ revertants as duplex VIII.Here the situation is more complicated than in the above case: For fragment induced reversion to take place, the remaining gap, which contains two GATC sites, has to be filled *in vivo*. This leads to the following configuration of GATC sites in the complete ds DNA: In duplex VII the two GATC sites close to the mismatch are completely unmethylated, the three distant ones are methylated only in the (-)strand . In duplex VIII the three distant GATC sites are symmetrically methylated, whereas the two nearby ones have the methyl groups in the (+)strand. In analogy to the findings with duplices V and VI, this can be expected to favour preservation of the (+)strand marker (i.e. α^- in duplex VIII). In duplex VII, in contrast, the three distant GATC sites will favour reversion to α^+. We interpret this to be the reason for the more pronounced difference between marker recoveries with VII and VIII as compared to V and VI. Note that GATC methylation at the two "nearby" sites in duplex VIII after completion *in vivo* resembles the situation created by *in vitro* reactions in previous directed mutagenesis studies (see above).

In an additional experiment, we constructed a precise insertion of 93 bp into the lac(α) region restoring an (α^-) deletion mutant to wild type. The starting point was phage M13mp3ΔM15 (B. Gronenborn, unpublished), which carries the ΔM15 deletion[29,30]. Construction of gapped duplices was carried out as in the reversion study on Z1 and also the same 339 bp wild type fragment was used for marker salvage.

In ternary duplices IX and X (figure 5), 93 nucleotides must be looped out from the mutagenic DNA fragment. As compared to the point mutation study, two

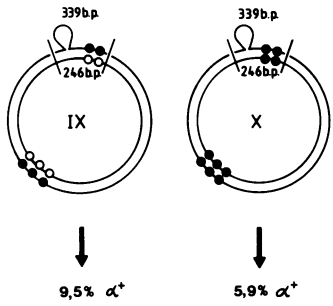

Figure 5:
Ternary duplices used in the directed insertion experiment. The looped out portion corresponds to 93 additional nucleotides carried by the wild type fragment. Other symbols are used as in figure 4.

effects can be seen: The marker yield is altogether lower and the influence of asymmetrical GATC methylation is less pronounced. The first effect might be due to cleavage of the single-stranded loop during transfection, the second could reflect a property of the repair system of the cell.

We would like to emphasize that our results are no rigorous proof of the involvement of DNA repair in the preferential marker preservation; other possibilities are left open[24]. The analogy to the findings of Radman et al.[24], however, is at least suggestive. No further studies were undertaken to clarify this question.

Extrapolating from our findings, it seems likely that construction of extended insertions will generally proceed with lower marker yields. This, however, should not be too detrimental, since in that case, unambiguous screening using "internal" DNA fragments can be expected to be a simple task.

In addition to opening the possibility of directing preferential marker preservation, use of gapped duplices in primer induced mutagenesis will clearly also be helpful in alleviating two other problems encountered in this method: a) Synthesis of complete duplex DNA by DNA polymerase/DNA ligase reactions *in vitro* will be facilitated since only short gaps have to be filled. b) Incorrect priming at inadvertant positions of the vector genome[12] is made impossible.

DNA duplices with precisely defined gaps as used in our study, will also be useful substrates for directed mutagenesis by single-strand specific mutagens and by "mutagenic fill-in reaction" (for review see D. Shortle et al.[31]).

We are presently extending the use of duplex DNA molecules with precisely defined gaps to other vector systems, such as pBR 322[33].

SUMMARY

We have demonstrated the influence of adenine methylation in GATC sites on

marker yield in directed mutagenesis of DNA cloned in filamentous phage. This effect explains why in previous studies the marker recovery was comparably low and also suggests how these experiments can be performed more efficiently: In a rather simple model experiment, we were able to push the marker yield beyond the 50% border by incorporating the desired marker into the methylated strand of an asymmetrically methylated DNA duplex.

We reckon that the 54% yield so far achieved can be further improved by enzymatic DNA manipulations *in vitro* to make sure that covalently closed DNA duplex molecules are used for transfection. This brings screening by DNA sequence analysis within the range of practicality.

For experiments that are less demanding with respect to marker yield, our findings suggest a slight modification of the scheme outlined in figure 1 as a simple minimal measure: Virion DNA should be isolated from phage grown in a dam^- host.

Our results also represent the first example of insertion construction by the marker salvage method.

ACKNOWLEDGEMENTS

We are deeply indebted to Dr. B. Gronenborn for generous gifts of various strains and continuous advice in genetics and biochemistry of phage M13. We also thank Dr. B. Kemper for communication of experimental protocols prior to publication. This work was supported by Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 74.

ABBREVIATIONS

bp: Base pairs; ccc DNA: Covalently closed circular DNA; DEAE: Diethylaminoethyl; ds DNA: Double-stranded DNA; EDTA: Ethylene diamine tetraacetic acid; IPTG: Isopropylthio-β-D-galactopyranoside; PEG: Polyethylene glycol; pfu: Plaque forming units; RF DNA: Replicative form DNA; ss DNA: Single-stranded DNA; Tris: Tris(hydroxymethyl)aminomethane; x-gal: 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside.

REFERENCES

1. Khorana, H.G. (1969) Angew. Chemie 81, 1027-1039
2. Khorana, H.G. (1979) Science 203, 614-625
3. Hutchison III, C.A. and Edgell, M.H. (1971) J.Virol. 8, 181-189
4. Schott, H. and Kössel, H. (1973) J.Amer.Chem.Soc. 95, 3778-3785
5. Razin, A., Hirose, T., Itakura, K. and Riggs, A.D. (1978) Proc.Natl.Acad. Sci.USA 75, 4268-4270
6. Hutchison III, C.A., Phillips, S., Edgell, M.H., Gillam, S., Jahnke, P. and Smith, M. (1978) J.Biol.Chem. 253, 6551-6560

7. Gillam, S., Jahnke, P., Astell, C., Phillips, S., Hutchison III, C.A. and Smith, M. (1979) Nucleic Acids Res. 6, 2973-2985
8. Gillam, S. and Smith, M. (1979) Gene 8, 81-97
9. Gillam, S. and Smith, M. (1979) Gene 8, 99-106
10. Gillam, S., Astell, C.R. and Smith, M. (1980) Gene 12, 129-137
11. Wallace, R.B., Johnson, P.F., Tanaka, S., Schöld, M., Itakura, K. and Abelson, J. (1980) Science 209, 1396-1400
12. Baas, P.D., Teertstra, W.R., Van Mansfeld, A.D.M., Jansz, H.S., Van der Marel, G.A., Veeneman, G.H. and Van Boom, J.H. (1981) J.Mol.Biol. 152, 615-639
13. Wasylyk, B., Derbyshire, R., Guy, A., Molko, D., Roget, A., Têoule, R. and Chambon, P. (1980) Proc.Natl.Acad.Sci.USA 77, 7024-7028
14. Kudo, I., Leineweber, M. and RajBhandary, U.L. (1981) Proc.Natl.Acad.Sci. USA 78, 4753-4757
15. Montell, C., Fisher, E.F., Caruthers, M.H. and Berk, A.J. (1982) Nature 295, 380-384
16. Temple, G.F., Andrée, M.D., Kenneth, L.R. and Kan, Y.W. (1982) Nature 296, 537-540
17. Miyada, C.G., Soberón, X., Itakura, K. and Wilcox, G. (1982) Gene 17, 167-177
18. Miller, J.H. (1972) Experiments in Molecular Genetics p. 433, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. Clewell, D.B. and Helinski, D.R. (1969) Proc.Natl.Acad.Sci.USA 62, 1159-1166
20. Heidecker, G., Messing, J. and Gronenborn, B. (1980) Gene 10, 69-73
21. Courage-Tebbe, U. and Kemper, B. (1982) Biochimica et Biophysica Acta 697, 1-5.
22. Cohen, S.N., Chang, A.C.Y. and Hsu, L. (1972) Proc.Natl.Acad.Sci.USA 69, 2110-2114
23. Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymology 65, 499-560
24. Radman, M., Wagner jr., R.E., Glickman, W. and Meselson, M. (1980) in "Developments in Toxicology and Environmental Sciences" Vol. 7, Progress in Environmental Mutagenesis, Alacevic, M. Ed. pp. 121-130, Elsevier North Holland, Amsterdam
25. Messing, J., Gronenborn, B., Müller-Hill, B. and Hofschneider, P.H. (1977) Proc.Natl.Acad.Sci.USA 74, 3642-3646
26. Zipser, D., Zabell, S., Rothman, J., Grodzicker, T. and Wenk, M. (1970) J.Mol.Biol. 49, 251-254
27. Horiuchi, K., Vovis, G.F. and Model, P. (1978) in: "The Single-Stranded DNA-Phages", Denhardt, D.T., Dressler, D. and Ray, D.S. Eds. pp. 113-137, Cold Spring Harbor Laboratory, New York
28. Van Wezenbeek, P., Hulsebos, T. and Schoenmakers, J. (1980) Gene 11, 129-148
29. Beckwith, J. (1964) J.Mol.Biol. 8, 427-430
30. Langley, K.E., Villarejo, M.R., Fowler, A.V., Zamenhof, P.J. and Zabin, I. (1975) Proc.Natl.Acad.Sci.USA 72, 1254-1257
31. Shortle, D., DiMaio, D. and Nathans, D. (1981) Ann.Rev.Genet. 15, 265-294
32. Gronenborn, B. and Messing, J. (1978) Nature 272, 375-377
33. Gleumes, B. and Fritz, H.-J., manuscript in preparation