UNIVERSITY OF ILLINOIS

April 16, 1990

THIS IS TO CERTIFY THAT THE THESIS PREPARED UNDER MY SUPERVISION BY

Michael S. Bartlett

ENTITLED Comparison of Two Methods of Generating

Random DNA Fragments for Cloning and Sequencing

IS APPROVED BY ME AS FULFILLING THIS PART OF THE REQUIREMENTS FOR THE

DEGREE OF Bachelor of Science in Microbiology

Instructor in Charge

APPROVED:

HEAD OF DEPARTMENT OF Microbiology

Comparison of Two Methods of Generating Random DNA Fragments for Cloning and Sequencing

By

Michael S. Bartlett

Thesis

for the
Degree of Bachelor of Science
in
Microbiology

College of Liberal Arts and Sciences
University of Illinois
Urbana, Illinois

1990

ACKNOWLEDGEMENTS

Many thanks to Dr. Gary Olsen, the overseer of this scientific foray, for his helpful input and numerous rough draft revisions. Also, thanks to Chris Olson, Amy Whetsell, and Gisbert Meuter, my cohabitants in the Center for Prokaryotic Genome Analysis, for their patience with me, especially on Fridays. Special thanks to Dr. Pat Gillevet for his much valued opinions and constant moral support, and for his relentless good humor.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

DNA—deoxyribonucleic acid

M13—bacteriophage M13

RF—Replicative Form, describing the M13 genome when it is double-stranded

lacZ—the gene coding for the enzyme β-galactosidase

X-gal—5-bromo-4-chloro-indolyl-β-D-galactopyranoside, a sugar which, when cleaved by the enzyme β-galactosidase, dimerizes and creates a blue color by virtue of its indole groups

IPTG—isopropyl-β-D-thiogalactopyranoside, a non-metabolically active sugar which acts as an inducer of the lacZ gene

EDTA—ethylenediaminetetraacetate, a chelator of divalent cations

U—units, as in units of an enzyme

DNase I—bovine pancreatic deoxyribonuclease I

ng—nanogram

µl—microliter

Taq polymerase—DNA polymerase from Thermus aquaticus

OD_{260}—optical density at 260 nm

psi—pounds per square inch

bp—base pairs

pfu—plaque forming units

INTRODUCTION

In sequencing a large fragment of DNA, it is of utmost importance that the sequence obtained completely cover the entire fragment, with no missing sequence information. Without such uniformity of coverage the usefulness of any sequence determined is dubious. Random breakage of the DNA fragment into smaller, more manageable and cloneable pieces for sequencing (so-called "Shotgun" DNA sequencing) is one approach in completely sequencing large DNAs.

It has already been shown that there are dependable strategies for randomly fragmenting large pieces of DNA and obtaining complete coverage of sequence. Anderson (1981) effectively used DNase I in the presence of Mn^{++} to randomly fragment a 4257 bp restriction fragment of bovine mitochondrial DNA, and was able to obtain the complete sequence of the restriction fragment with about 50 M13 clones. DNase I is an enzyme that normally introduces random single-stranded nicks into DNA, but that, in the presence of Mn^{++}, introduces essentially random double-stranded cuts (Melgar and Goldthwait, 1968).

A recent development has seen the use of "nebulizers" (or any device which will create an aerosol) to generate shear forces sufficient to introduce double-stranded breaks into DNA, fragmenting it to a desired size. It has been discovered that the size of the resultant DNA fragments is directly dependent on the gas pressure used to aspirate the DNA through the nebulizing apparatus (Stefan Surzycki, Department of Biology, Indiana University, unpublished results)

In this thesis, I compare these two methods of fragmentation and attempt to clone the fragments into M13 RF for sequencing. The DNase I method of fragmentation was attempted, but the first try yielded no recombinant ligation products. Another try yielded supposedly recombinant ligation products, but after sequencing it was determined that the "recombinant" white plaques were actually formed by M13 with small deletions rather than inserts of the digested DNA.

Fragmentation of DNA using a nebulizer was then attempted. The resulting fragments were also ligated into M13, and this time white plaques from the transfection yielded genuinely recombinant clones. The clones were sequenced using a fluorescent DNA sequencer and were found to yield novel sequence data.

MATERIALS AND METHODS

Bacterial strains, vectors, and media. Escherichia coli strain DH5αF', the host organism for the propagation of recombinant phage vector, was obtained from Bethesda Research Laboratories, Inc. (BRL). The double-stranded, or replicative form (RF), of the vector bacteriophage M13mp19 was obtained from Boehringer Mannheim Biochemicals (BMB). Recombinant phage were detected by the α-complementation reaction, whereby the host bacterium contains DNA coding for the omega (Ω) fragment of the lacZ gene product and the transfecting phage genome contains the DNA coding for the alpha (α) fragment of that gene product. In the presence of IPTG, a non-metabolically active inducer of the lacZ gene, non-recombinant phage transcribes and translates the α-fragment normally, which acts in concert with the Ω-fragment, enzymatically cleaving the indolic sugar X-gal, creating a blue plaque on the lawn of host cells. In recombinant phage, this sequence is interrupted by a foreign piece of DNA and may not form an active α-fragment of the lacZ gene product. Such phage will grow on a lawn of host bacteria, but the plaques formed will be white rather than blue, allowing for easy screening for recombinant phage particles (Yanisch-Perron et al., 1985). The DH5αF' cells were grown in Luria-Bertani (LB) top agar poured onto plates of solidified 1.5% LB agar, both prepared according to Sambrook et al. (1989).

Agarose gel electrophoresis. The DNA products of the various enzymatic manipulations were examined using agarose gel electrophoresis. Most often the gels were 1% agarose and were run at 4-6 volts per centimeter in a 1X tris-borate EDTA (TBE) buffer, prepared according to Sambrook et

al. (1989). After electrophoresis was complete, the gels were stained for 10 minutes on a rotary shaker with about a 2 µg/ml solution of the fluorescent dye ethidium bromide. The gels were then destained for 10 minutes with distilled water on the rotary shaker. The stained DNA was photographed under ultraviolet illumination.

DNase I digestion of double-stranded DNA. Partial digestions of DNA were performed according to Anderson (1981) with DNase I (from Sigma Chemical Company). Mn^{++}, which is required for DNase to act, was added to a final concentration 0.01 mM. The stock of DNase I was standardized with control DNA to determine its relative activity prior to digesting target DNA.

Fragmenting DNA with a nebulizer. DNA was also fragmented using the Model No. 1731 T UP-DRAFT II™ NEB-U-MIST™ NEBULIZER, from the Hudson Oxygen Therapy Sales Company. A 25 µg/ml solution of the DNA in H_2O was cycled for 2 minutes through the apparatus using nitrogen at 30 psi. The mist created by the nebulizer was recovered in a coil of sterile plastic tubing.

Purification of DNA. All DNA was purified by phenol extraction and recovered by ethanol precipitation as described in Sambrook et al. (1989). Ammonium acetate (NH_4OAc) was the salt used in all of the precipitations.

DNA Tailing. Vector and insert DNAs were tailed with dCTP and dGTP, respectively, using terminal deoxynucleotidyltransferase (Roychoudhury et al. 1976) and appropriate buffers from International Biotechnologies, Inc. (IBI) according to the protocol specified by the company. Controls were run in which a short (24 nucleotide) piece of DNA was tailed with [α-^{32}P] dATP and the reaction products resolved on a

polyacrylamide gel. Autoradiographic detection of the labelled products was used to determine the length of the tails added.

DNA ligation. Vector and insert DNAs were ligated using T4 DNA ligase, which catalyzes the formation of phosphodiester bonds between adjacent 3'-hydroxyl and 5'-phosphate termini in DNA (Weiss et al. 1968). The enzyme was from IBI and was used according to the protocol recommended by the company.

Transfection with double-stranded M13 DNA. Transfection with exogenous bacteriophage M13 DNA was performed using competent E. coli DH5aF' cells which were either purchased from BRL or prepared according to a protocol, modified from Cohen et al. (1972), from the BRL catalog. The competent cells that were prepared were stored in 40% glycerol at -70° C.

Klenow repair of DNA. DNA with damaged ends from the fragmentation steps of preparation was repaired using the Klenow fragment of E. coli DNA polymerase I (from BMB), using the protocol described by Ausubel et al. (1989). The Klenow fragment's DNA polymerase activity fills in 5' overhangs and its 3'-to-5' exonuclease activity degrades 3' overhangs.

Screening phage for DNA inserts. White plaques were picked from transformation plates and grown overnight in 15 ml LB broth. The cells were removed by centrifugation for 10 minutes at 3000 rpm (Sorvall RT6000B Refrigerated Centrifuge). The supernatant (which contains the phage particles) was raised to 1% sodium dodecyl sulfate and 20 µl were electrophoresed on a 1% agarose gel in order to screen the clones for inserts and determine the sizes of the inserts.

Preparation of single-stranded M13 template for sequencing. DNA was purified from clones that appeared to have inserts in order to make suitable templates for sequencing, according to a protocol from Sambrook et al. (1989). The following modifications in the protocol were made: 10 ml of supernatant was used; the manipulations were done in 50 ml round bottomed centrifuge tubes and centrifugation was done using a Beckman J2-21 Centrifuge with a JA-20 rotor; 2 ml 20% polyethylene glycol/3.75 M ammonium acetate (NH_4OAc) were added to precipitate the phage particles; the phage particles were resuspended in 500 µl H_2O; phenol extraction of the DNA was performed once using 3 ml phenol, followed by 3 ml 24:1 chloroform/isoamyl alcohol; DNA was precipitated using 1/2 volume 7.5 M NH_4OAc and 2 volumes 95% ethanol; and the precipitated DNA was resuspended in 50 µl sterile H_2O.

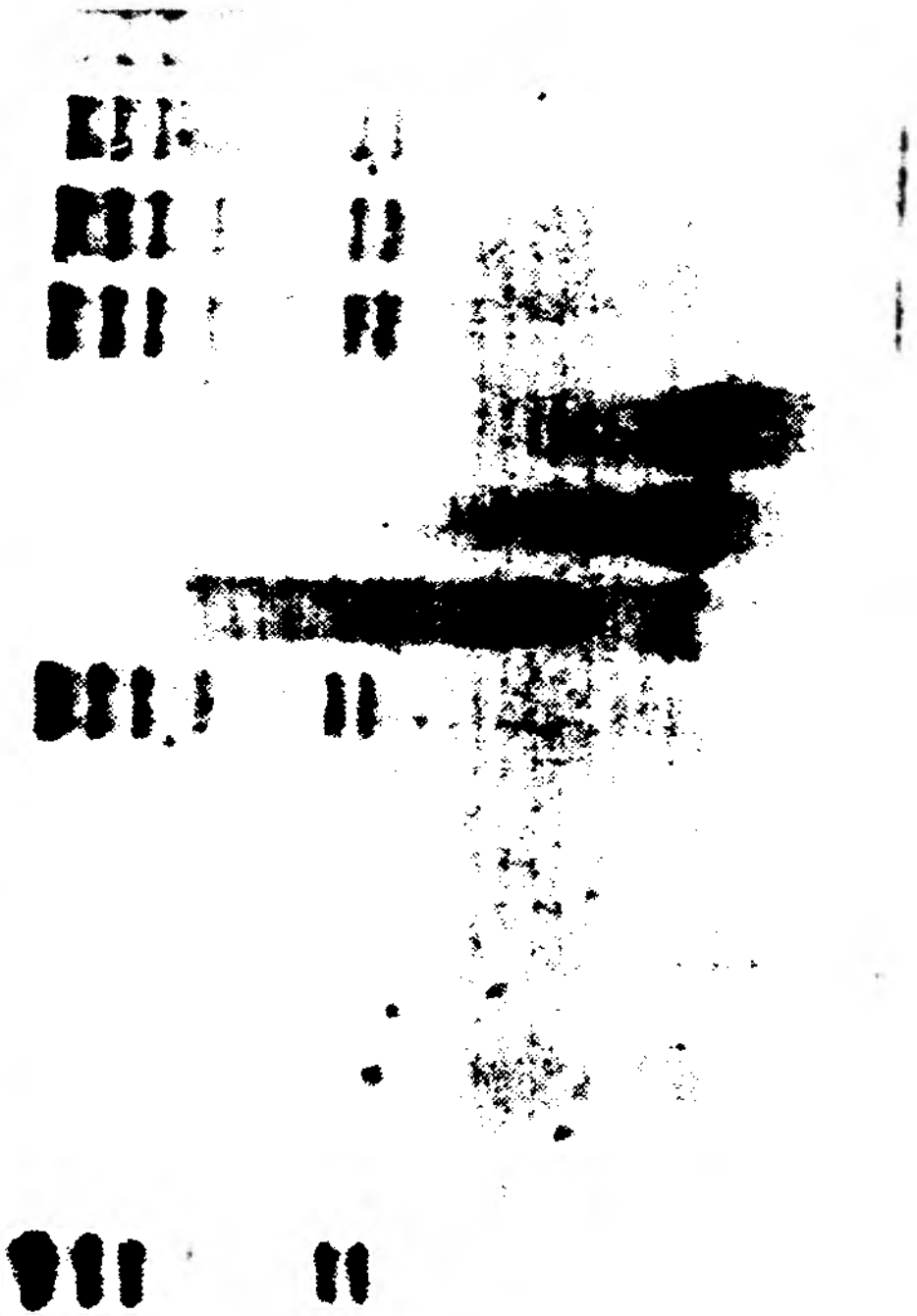
Sequencing using M13 and the fluorescent sequencer. Single-stranded M13 templates were sequenced on an ABI 370A DNA sequencer. Sequencing reactions were carried out using "Applied Biosystems Model 370A DNA Sequencing System/Taq Polymerase," a kit from the Promega Corporation. The protocol supplied with the kit (dated 1/89) was followed, except for the following changes: reactions were performed in a Falcon® 3911 MicroTest III™ Flexible Assay Plate (from Becton Dickinson), rather than using a microfuge tube for each individual reaction; Taq polymerase was added before the annealing step, rather than after it as the protocol specified; and the "chase" step mentioned in "Notes and Comments" was omitted. The products of the sequencing reactions were electrophoresed on a 6% denaturing polyacrylamide gel in the ABI 370A DNA sequencer.

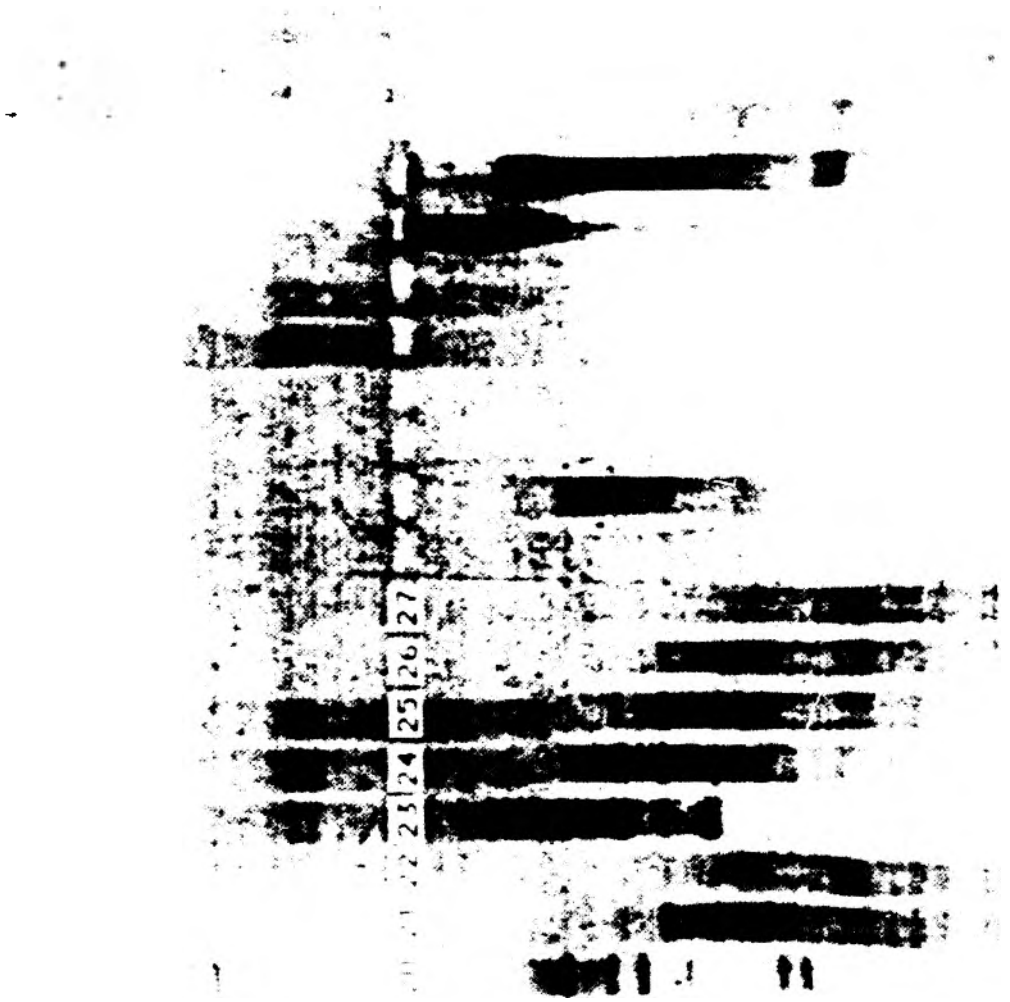
Sequence data analysis. The raw data obtained from the ABI 370A sequencer was initially processed using version 1.3 of the ABI Fluorescent Analysis Package. The processed data was then run through the Fragment Assembly program in version 6.0 of the Wisconsin GCG Sequence Analysis Software Package. The MacVector Package (from IBI) was used to search the sequences obtained for open reading frames. The FASTA program of the GCG package was used in comparing obtained sequences with the entire sequence database.

RESULTS

Random fragmentation of DNA using DNase I. DNase I was used in the presence of Mn^{++} to partially digest a target DNA so that a random library of fragments covering the length of the DNA could be cloned into bacteriophage M13 RF DNA and ultimately sequenced. The relative activity of a 2000 U/ml DNase I stock solution was assayed by performing controlled digestions of 2 µg of HindIII digested bacteriophage λ DNA. Reactions were stopped at various timepoints by adding EDTA (to a final concentration of 15 mM) which complexes with the Mn^{++} ions necessary for enzymatic activity. As shown in Figure 1, the optimal control over digestion was achieved with a reaction time of 15' and a DNase I concentration of 0.02 U/90 µl reaction volume.

The experimental DNA consisted of five plasmids containing inserts of DNA from the proton translocating ATPase gene from Methanococcus voltae (plasmids supplied by R. Dharmavarani, in the laboratory of J. Konisky, University of Illinois). The plasmids (pRD-1, pRD-7, pRD-8, pRD-9, and pRD-9-1) were digested using 0.02 U of DNase I and the reaction was carried out on aliquots of each plasmid for 0, 5, 10, 15, and 20 minutes. The products of this digestion were then examined by agarose gel electrophoresis (figure 2). Plasmid pRD-7 was omitted from further study due to lack of sufficient amounts of DNA. The following time points for each plasmid yielded appropriately sized DNAs for further cloning purposes, and were thus pooled: pRD-1, 15 and 20 minutes; pRD-8, 5 and 10 minutes; pRD-9, 20 minutes; and pRD-9-1, 20 minutes. After phenol extraction and ethanol precipitation, the fragmented plasmid DNAs were tailed with dCTP, and again phenol extracted




27
26
25
24
23

and ethanol precipitated. Bacteriophage M13mp19 RF was linearized using the blunt cutting restriction enzyme SmaI, and was then phenol extracted and ethanol precipitated. The linearized phage DNA was tailed with dGTP, and again phenol extracted and ethanol precipitated. The DNAs were quantitated based on their OD_{260} using a Beckman DU 70 Spectrophotometer, so that appropriate ligation ratios could be calculated.

Ligations were carried out between the linearized and dG-tailed M13 and the fragmented and dC-tailed plasmid DNAs. The ligation mixtures were then directly used to transfect competent DH5αF' E. coli. Control M13 gave blue plaques on a lawn of cells, but none of the ligation mixes gave any plaques, blue or otherwise. It was determined by agarose gel electophoresis that the DNA (which there was very little of to begin with) was probably lost during the phenol extractions. After each phenol extraction/ethanol precipitation the amount of DNA recovered had been measured with the spectrophotometer, but it obviously gave erroneous results. Possibly residual phenol from the extraction step or NH_4OAc from the precipitation step had caused these errors in quantitation.

The DNase method of DNA fragmentation was tried again, but this time on two inserts that had been enzymatically cut out of their plasmids (pRD-9, pRD-9-1). The inserts were each digested for 0, 5, 10, and 15 minutes, with EDTA being added at the appropriate times to halt the reaction, as before. After examining the reactions on an agarose gel (figure 3) the last two timepoints for each insert were pooled and phenol extracted, this time using 1 µl of glycogen stock solution (20 mg/ml) from BMB as a carrier to aid in the ethanol precipitation of the purified DNA.

DNase I Digestion of Inserts

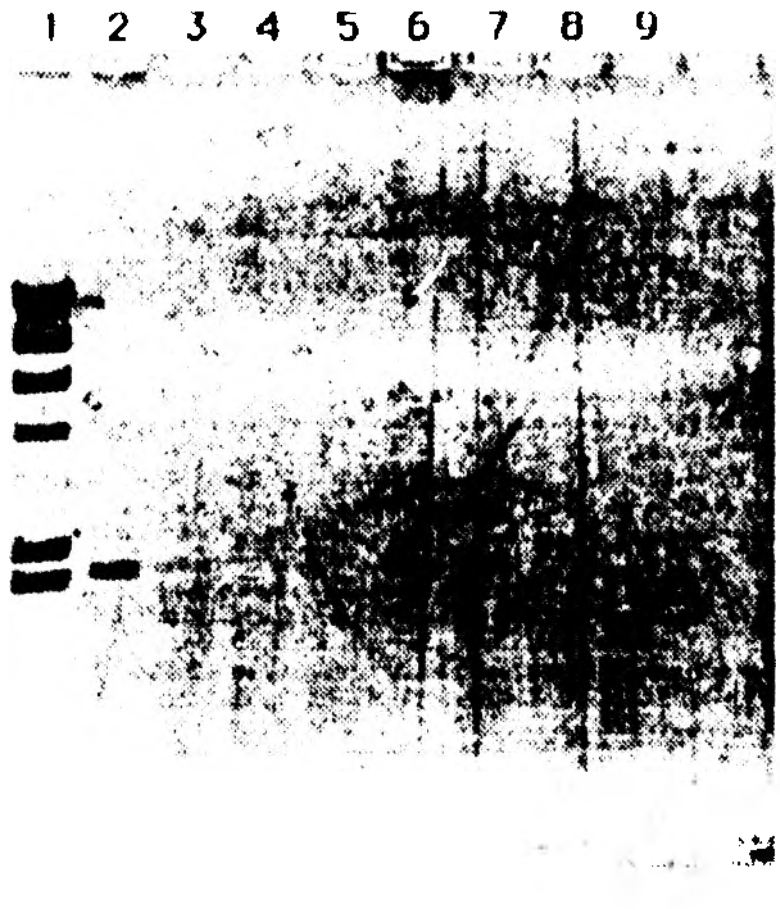


Figure 3

Insert DNA digestion with DNase I. Lane 1 contains 355 ng of λ HindIII fragments as size markers. Lanes 2-5 contain about 125 ng of insert RD-9 digested with 0.02 U DNase I for, from left to right, 0', 5', 10', and 15'. Lanes 6-9 contain about 150 ng of insert RD-9-1 digested with 0.02 U DNase I for 0', 5', 10', and 15'. For both inserts, 10' was sufficient to completely digest the DNA, as suggested by the complete disappearance of the insert band.

The digested inserts were not tailed, but instead were blunt end ligated into M13 RF. Since it is likely that digestion with DNase leaves ragged, rather than blunt ended DNA, the ends of the DNA were "repaired" using the Klenow enzyme. Blunt-ended ligations, with vector to insert molar ratios of 1:1, 1:3, and 1:5, were performed between the digested, repaired inserts and M13 that had been linearized with SmaI. Competent DH5αF' E. coli was transfected with about 50 ng of DNA from each of the ligation mixes. Plaques were observed, and the transfection efficiency of the ligation mixes was about $3x10^3$ pfu/μg vector. Optimal recombination occurred with a 1:1 vector to insert ratio, where the frequency of white plaques was about 25%.

Twenty-eight white plaques were picked from the transformation plates (using sterile toothpicks) and grown for 9 hours in 15 ml LB broth. The phage in the culture supernatants were screened for phage inserts (see MATERIALS AND METHODS). Seventeen of the M13 clones appeared to possess very small inserts, and eleven did not. Regardless, twenty were purified into template to be used for sequencing. Sequencing reactions were performed on sixteen templates and they were run on the Applied Biosystems, Inc. 370A DNA Sequencer. All of the clones were determined to be just M13, with small deletions causing a frame-shift and resulting in incorrect translation of the α-fragment of the lacZ gene product.

Testing the ligation and transfection protocols. To see whether or not there were problems with the basic techniques of enzyme restrictions, ligations, or phenol extractions/ethanol precipitations, a control experiment was set up in which HindIII digested bacteriophage λ was ligated with M13 with either blunt or cohesive ends. First, M13 RF was linearized with SmaI, phenol extracted and ethanol precipitated, and was ligated, under blunt-ended

ligation conditions, with λ HindIII fragments that had been end-filled with the Klenow enzyme. Special attention was paid to ligation ratios and total amounts of DNA per reaction volume, so that optimal ligation conditions might be determined. About 50 ng of vector was used in each reaction, and insert was added in molar ratios (vector to insert) of 1:1, 1:5, and 1:10. Transfections into competent E. coli DH5αF' were performed with DNA from each ligation ratio. Efficiency of transfection (using 0.25 ng of untreated M13mp19 RF from BMB) was about 2.5×10^6 pfu/μg vector. Optimum recombination (number of white plaques/total number of plaques) occurred with a vector to insert ratio of 1:5, using about 5 ng of DNA in the transfection, and was about 24%. However, screening 10 of these clones for inserts (as described in MATERIALS AND METHODS) indicated that only one or two actually possessed inserts, and that the others likely formed white plaques as a result of frame-shifting deletions.

Alternatively, M13 RF was linearized using HindIII, leaving cohesive ends to ligate with unaltered λ HindIII fragments. The DNA was phenol extracted and ethanol precipitated, and ligations were carried out, again with 50 ng of vector used in each ligation, with insert added to achieve vector to insert molar ratios of 1:1, 1:5, and 1:10. Cells were transfected as above. Optimum recombination (about 32%) occurred with a vector to insert ratio of 1:10, again using about 5 ng of DNA to transfect. Ten white plaques were screened for recombinant status, and it appeared that seven of the clones did possess sizeable inserts. Thus it appeared that the standard restriction and ligation techniques were working, so it was decided to continue with the experiment.

DNA fragmentation using a nebulizer. The experimental DNA was purified from a bacteriophage λ clone, obtained from L. Achenbach (in the laboratory of C. Woese, University of Illinois), with an insert of about 15,000 base pairs from Thermococcus celer genomic DNA that includes the rRNA genes. As described in Materials and Methods, 700 µl of the DNA (25 ng/µl) was run through a nebulizer for 2 minutes at 30 psi of pressurized nitrogen. About 325 µl were recovered from the starting volume (a 46% recovery). Some of the product of this manipulation was run on a 1% agarose gel to check the efficiency of shearing and the average size of the resultant fragments. As can be seen in the gel photograph (figure 4), the fragments are concentrated at around 300 base pairs (as compared to the last band in the HindIII/λ marker lane), a fairly good length for random sequencing.

Since it is probable that fragmentation of DNA in this way will leave damaged ends (i.e., 5' or 3' overhangs, and dephosphorylated 5' ends) that would not tail well in preparation for ligation, the sheared DNA was divided into 2 aliquots. One aliquot was treated with T4 polynucleotide kinase to phosphorylate any unphosphorylated 5' ends, and was then ethanol precipitated. The other aliquot was incubated with T4 polynucleotide kinase, and the products were ethanol precipitated. The Klenow fragment of DNA polymerase I was then used to repair ragged ends, and the products were again ethanol precipitated. Both aliquots were then dCTP tailed and subsequently phenol extracted and ethanol precipitated. In preparation for cloning, M13 RF was linearized with SmaI, phenol extracted and ethanol precipitated, and then tailed with dGTP, and again phenol extracted and ethanol precipitated.

Ligation reactions were performed between the complementarily tailed M13 and nebulized DNA. In the first ligation, the vector concentration was

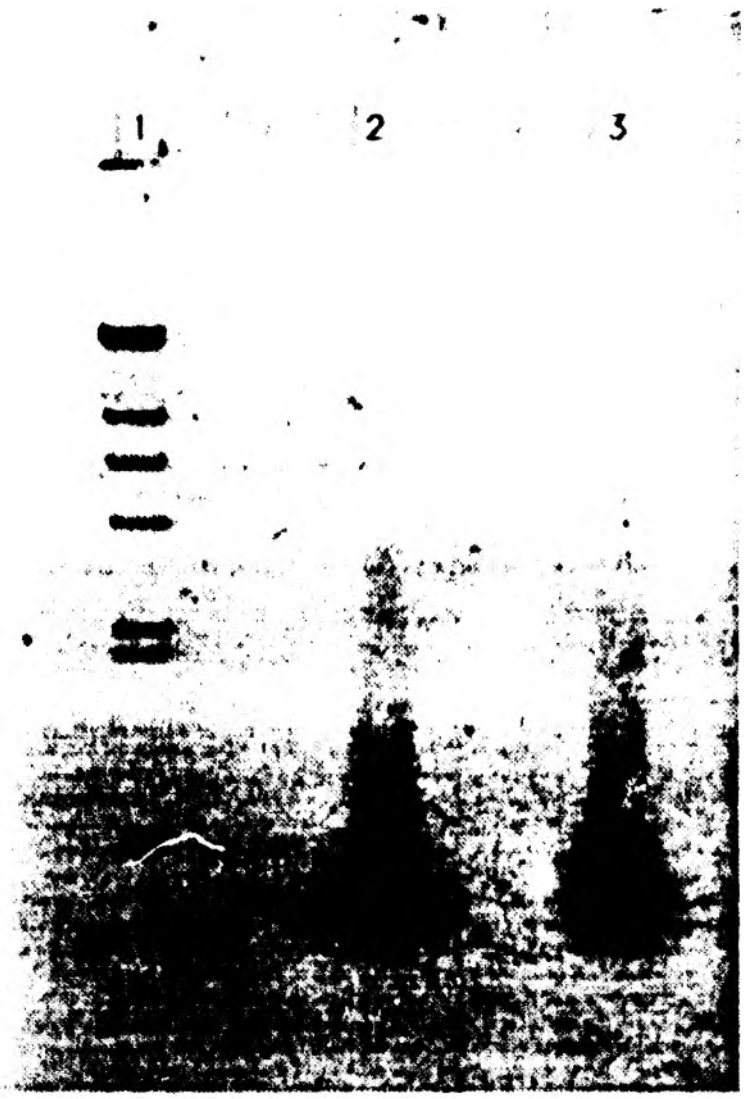


Figure 4

Nebulized phage λ DNA. Lane 1 contains 355 ng λ HindIII fragments used as size markers. Lane 2 contains about 350 ng DNA from a λ clone after aspiration through a nebulizer for 2 minutes at 30 psi N_2. Lane 3 contains about 200 ng of the same DNA after aspiration through a nebulizer for 4 minutes at 30 psi N_2. Note that there is no change in size distribution of fragment sizes with increasing times of nebulizing at constant pressure.

inadvertently ten times higher than it should have been, and observation of the ligation products with agarose gel electrophoresis indicated that the predominant ligation product was self-ligated M13 RF. The ligation reactions were redone using an adjusted amount of M13, and an agarose gel of the ligation products (figure 5) indicated that the ligation products probably included recombinant M13 molecules. A transfection was carried out using these ligation products, but the ratio of white plaques to blue plaques was not very high (about 10%), and was about the same regardless as to whether the nebulized DNA was repaired using T4 polynucleotide kinase only or both T4 polynucleotide kinase and the Klenow fragment. Nevertheless, phage stocks were grown and screened for inserts. Sixteen clones that seemed likely to possess inserts were purified into templates suitable for sequencing, and two separate runs on the ABI 370A DNA Sequencer were done. On the first run the gel did not polymerize properly and the sequence obtained was rather erratic and ambiguous. However, on the second run the data obtained was very clean and gave unambiguous sequence.

The sequences obtained were computer analyzed as described in MATERIALS AND METHODS, and it was determined that four of the clones were M13 without insert DNA that formed white plaques because of small, frame-shifting deletions. However, ten of the clones were true recombinants and possessed inserts, ranging in size from 40-420 bases, that yielded sequence data (see table 1). A significant aspect of the data was that in the clones that possessed inserts, the cloning site of the M13 showed no dGTP tailing, and thus that the the only ligation occurring was blunt-ended.

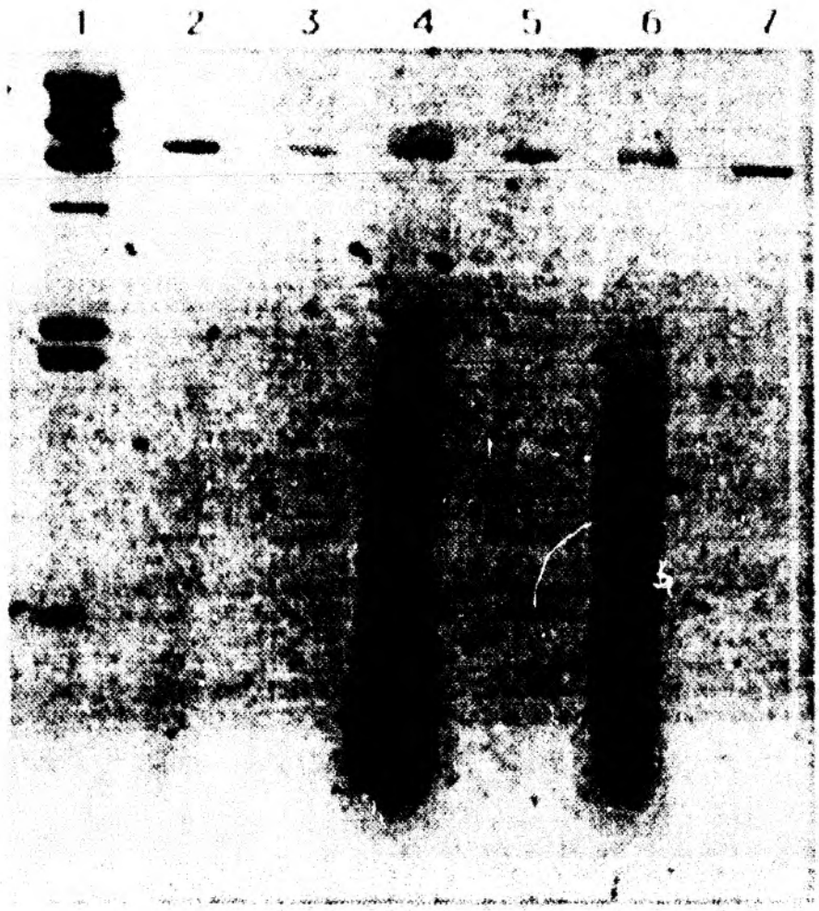


Figure 5

DNA products of ligations between M13 RF and nebulized λ DNA. Lane 1 contains 355 ng λ HindIII fragments as size markers. Lanes 2 and 7 contain 25 ng of linearized, unligated M13. Lanes 3 and 4 contain 25 ng M13 ligated with 10 ng and 100 ng of kinased and tailed nebulized λ DNA. Lanes 5 and 6 consist of 25 ng M13 ligated with 10 ng and 100 ng kinased, Klenow treated, and tailed nebulized λ DNA. Note the sharpness of the bands in the unligated M13 lanes, and compare to the bands at the same positions in lanes 3-6. The fact that the ligation reactions produced smeared M13 bands indicates that ligation must have been occurring.

Table 1. Sequence data obtained from clones of DNA sheared with nebulizer.

Template	Status of cloning site
MBL1-1	21 base pair deletion
MBL1-3	single base pair deletion
MBL1-5	320 base pair insert
MBL1-6	single base pair deletion
MBL1-7	single base pair deletion
MBL1-8	40 base pair insert
MBL1-9	360 base pair insert
MBL1-10	90 base pair insert
MBL1-11	400 base pair insert
MBL1-12	single base pair deletion
MBL1-13	400 base pair insert
MBL1-14	400 base pair insert
MBL1-15	420 base pair insert
MBL1-16	350 base pair insert
MBL1-17	54 base pair insert
MBL1-18	no sequence obtained

DISCUSSION

Some of the data recorded in RESULTS may seem ambiguous, but this is mainly a result of problems encountered in tailing and subsequent ligation of the vector and insert DNAs. Still, there are a number of conclusions that may be drawn from these investigations of DNA fragmentation using a nebulizer and fragmentation using DNase I .

The advantages of fragmentation using a nebulizer over digestion with DNase I are plainly evident. First, the only specific condition the nebulizer requires is that the target DNA be in aqueous solution at a reasonable concentration. In contrast, DNase I digestion is dependent on pH, ionic conditions, DNA concentration, enzyme concentration, and digestion time.

The amount of fragmentation of a given DNA using a nebulizer is easily regulated; varying the pressure of the gas used to aspirate the target DNA is all that is needed to produce fragments of desired sizes (Stefan Surzycki, unpublished results). In contrast, the only way to achieve the same effects using DNase I is to follow the unwieldy procedure of digesting the same target DNA with DNase I for varying amounts of time, examining the products from each timepoint by gel electrophoresis, and then pooling the DNA products which seem most useful for further manipulations. The DNA from timepoints which have produced inappropriately sized fragments of DNA is discarded, reducing the yield of DNA for cloning. Additionally, the relative activity of DNase I can vary from batch to batch, requiring that each new batch be individually standardized.

The fragmentation caused by a nebulizer is based strictly on random physical shearing, and thus shows no specificity of action with regard to the

base composition of the target DNA. DNase I, however, has been found to display some specificity with respect to sequence recognition. Anderson (1981) noted that in 92 sequenced M13 clones, 33% of the cut sites occurred one nucleotide to the 5' side of the dinucleotide sequence G-T, which is three times the frequency expected if the cuts were located randomly with respect to G-T dinucleotides. Such specificity of cutting means that libraries constructed using DNase I would not be truly random, which could cause problems in obtaining complete representation of the target DNA with a minimum number of M13 clones.

DNA fragments produced by a nebulizer are extremely uniform in size as compared to fragments produced during a DNase I digestion. Figures 6 and 7 (3-dimensional plots of figures 4 and 1 respectively) graphically demonstrate this fact; while the fragments produced by the nebulizer are concentrated in the range of 300-500 base pairs with a little smearing at larger sizes, the fragments from the DNase I digestion are present as broad smears of DNA ranging from about 300-1500 base pairs, depending on which timepoints are chosen. There does seem to be an area of localization of fragment sizes at around 500 base pairs, but it's hardly comparable to the control given with the nebulizer.

Finally, use of the nebulizer does not require purification of its products, because no enzymes or special buffers are used. In contrast, the products of a DNase I digestion must be rigorously purified, because any remaining enzyme could be reactivated by divalent cations added later in the manipulation of the digestion products.

Use of a nebulizer in DNA fragmentation is not without drawbacks, though. Of immediate concern is the production of aerosols containing DNA

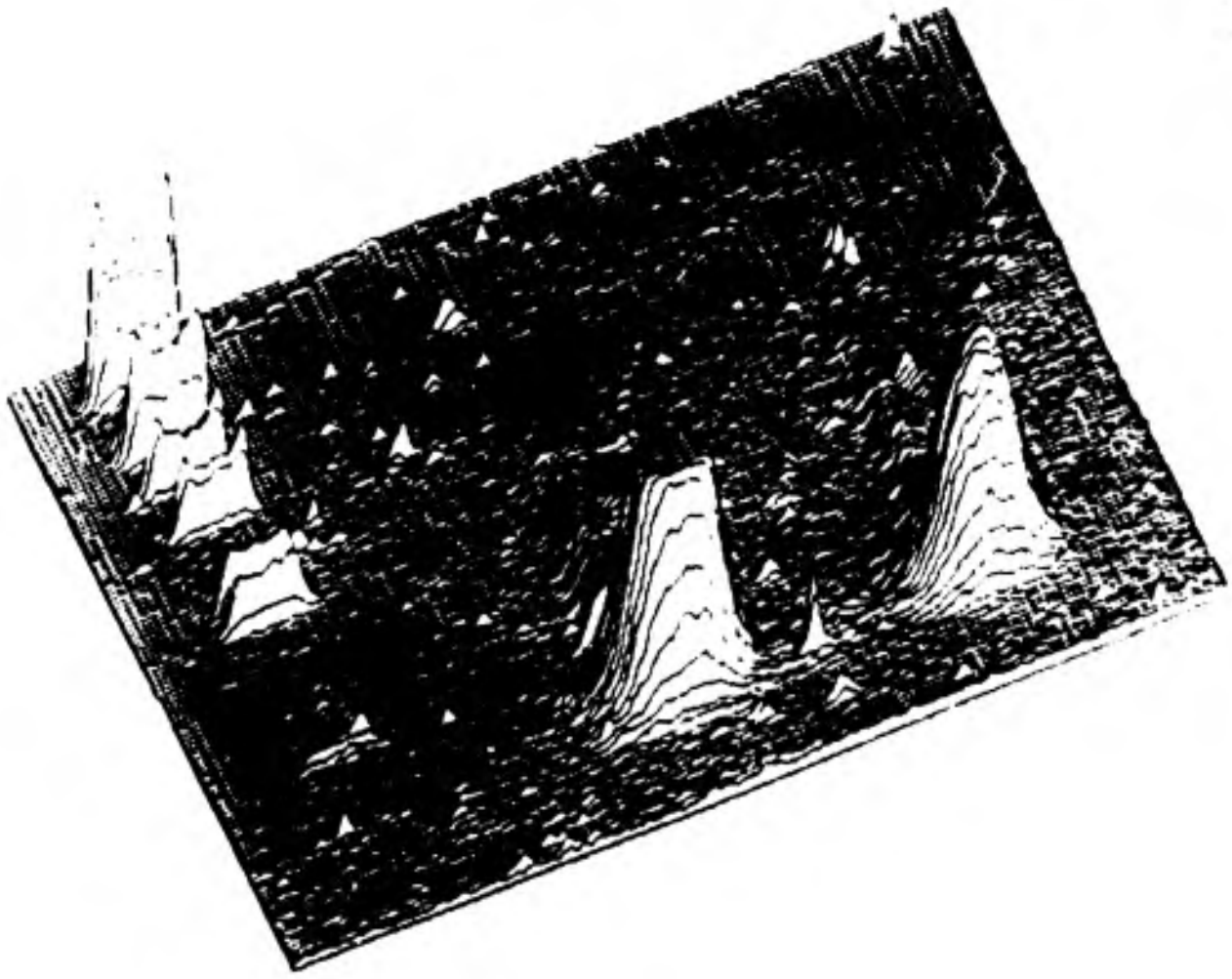


Figure 6

3-dimensional plot of nebulized DNA products. The digitized photograph of the nebulized λ clone DNA (figure 4) was 3-dimensionally plotted using the NIH *Image* program, version 1.17. Height in the plot is a function of intensity, which in turn is a function of amount of DNA in the gel. The first set of peaks is from the λ HindIII marker. The next two large peaks are from the nebulized DNA, and illustrate the narrow distribution of sizes of fragments produced during a nebulizing experiment.

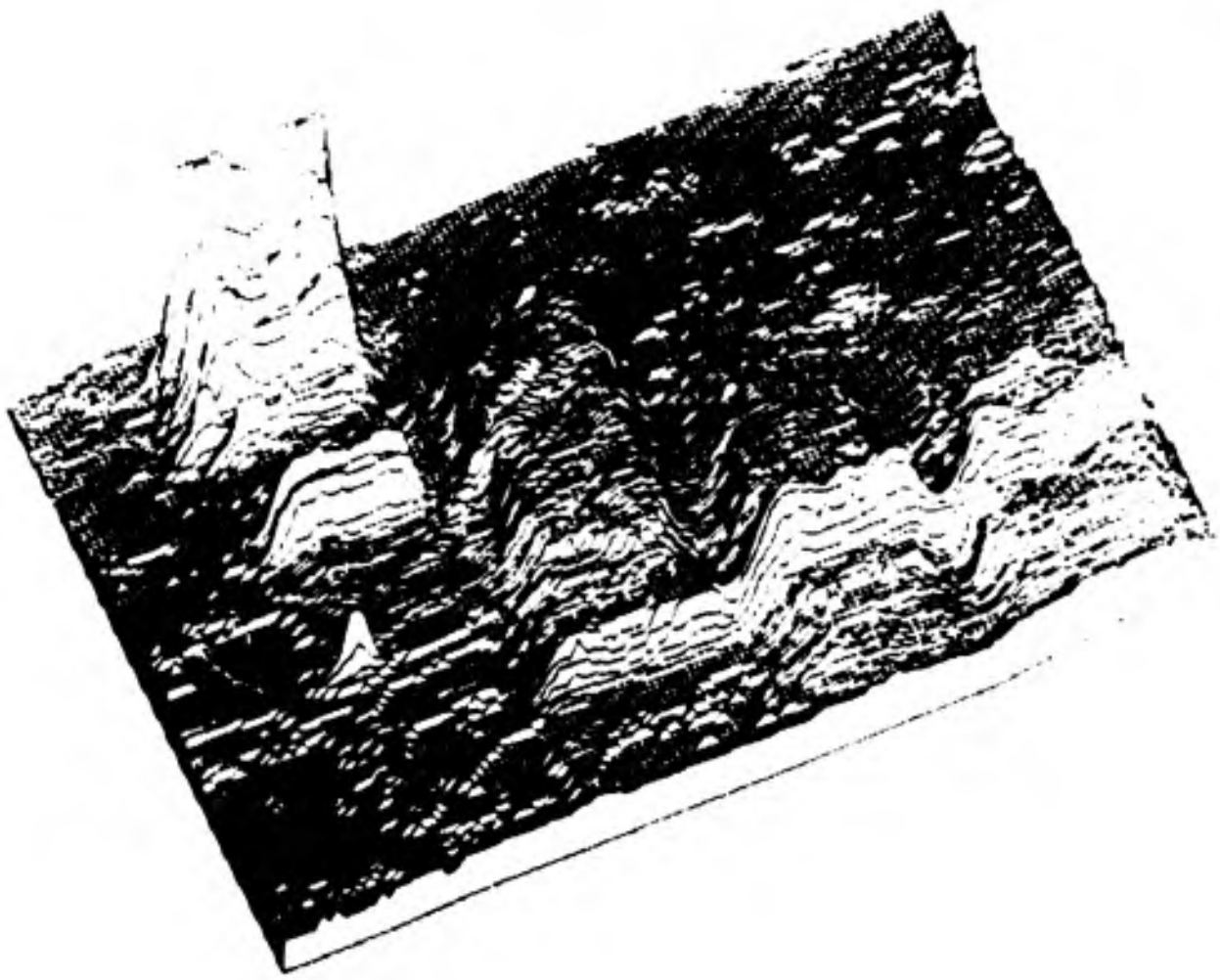


Figure 7

3-dimensional plot of DNase I partial digestion products. The plot was taken from lanes 6-9 of figure 1, with the peaks showing the progressive digestion of the target DNA with time. The relatively broad distribution of fragment sizes produced by DNase I digestion can be seen in the peaks produced, which are neither very sharp nor tightly localized.

that may contaminate laboratory surfaces, and that may also be harmful to laboratory personnel. Appropriate aerosol containment techniques must be utilized, and in any case it would not be advisable to use a nebulizer for fragmenting any kind of potentially harmful DNA, such as DNA from plasmids containing genes for drug resistance, DNA from any known bacterial or viral pathogen, or DNA from unhealthy organisms.

In summary, it is evident from the experiments carried out that fragmentation of large DNAs using a nebulizer for subcloning and sequencing is not only possible, but that it is also superior to the DNase I method of fragmentation. How fragmentation using a nebulizer compares to fragmentation by sonication (Bankier et al. 1987) has yet to be studied, but it is probable that these two methods are comparable because of the fact that both use mechanical forces for shearing the DNA.

An important consideration in the development of random sequencing techniques is their application to very large scale sequencing projects (i.e. the Human Genome Project). Clearly, such projects require that the basic techniques employed in generating data (such as subcloning of large DNA fragments) be absolutely standardized and their efficiency maximized, so that attention may be focused on the data itself rather than on problems associated with data production. Given the fact that DNA fragmentation using a nebulizer is dependent on so few variables as compared to fragmentation using DNase I, it is logical that the nebulizing technique would be chosen for a scale-up, if not for its relative simplicity, then for its easy standardizability.

LITERATURE CITED

1. Anderson, S. (1981). Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Research* **9**, 3015.
2. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (1989). "Current otocols in Molecular Biology." John Wiley and Sons, New York.
3. Bankier, A. T., Weston, K. M., Barrell, B. G. (1987). Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. *Methods Enzymol.* **155**, 51.
4. Cohen, S. N., Chang, A. C. Y., Hsu, L. (1972). Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of Escherichia coli by R-factor DNA. *Proc. Natl. Acad. Sci.* **69**, 2110.
5. International Biotechnologies, Inc. (1986/87). "Products for Molecular Biology." New Haven, Connecticut.
6. Melgar, E., and Goldthwait, D. A. (1968). Deoxyribonucleic acid nucleases. II. The effects of metals on the mechanism of action of deoxyribonuclease I. *J. Biol. Chem.* **243**, 4409.
7. Roychoudhury, R., Jay, E., Wu, R. (1976). Terminal labeling and addition of homopolymer tracts to duplex DNA fragments by terminal deoxynucleotidyl transferase. *Nucleic Acids Research* **3**, 101.
8. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). "Molecular Cloning", Second Edition. Cold Springs Harbor Laboratory, Cold Springs Harbor, New York.
9. Weiss, B., Jacquemin-Sablon, A., Live, T. R., Fareed, G. C., Richardson, C. C. (1968). Enzymatic breakage and joining of

deoxyribonucleic acid. VI. Further purification and properties of polynucleotide ligase from Escherichia coli infected with bacteriophage T4. *J. Biol. Chem.* **243**, 4543.

10. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 / pUC19 vectors. *Gene* **33**, 103.