Genetic Engineering

Principles and Methods

Volume 4

Edited by

Jane K. Setlow

and

Alexander Hollaender

Genetic Engineering

Principles and Methods

Volume 4

Edited by

Jane K. Setlow

Brookhaven National Laboratory

Upton, New York

and

Alexander Hollaender

Associated Universities, Inc.

Washington, D.C.

Plenum Press · New York and London

The Library of Congress cataloged the first volume of this title as follows:

Genetic engineering; principles and methods. v.1-
New York, Plenum Press [1979-
v. ill. 26 cm.
Editors: 1979- J. K. Setlow and A. Hollaender.
Key title: Genetic engineering, ISSN 0196-3716.

1. Genetic engineering—Collected works. I. Setlow, Jane K. II. Hollaender, Alexander, 1898-
QH442.G454 575.1 79-644807
MARC-S

Library of Congress Catalog Card Number 79-644807
ISBN 0-306-41113-X

PREFACE TO VOLUME 1

This volume is the first of a series concerning a new technology which is revolutionizing the study of biology, perhaps as profoundly as the discovery of the gene. As pointed out in the introductory chapter, we look forward to the future impact of the technology, but we cannot see where it might take us. The purpose of these volumes is to follow closely the explosion of new techniques and information that is occurring as a result of the newly-acquired ability to make particular kinds of precise cuts in DNA molecules. Thus we are particularly committed to rapid publication.

Jane K. Setlow

Alexander Hollaender

ACKNOWLEDGMENT

The Editors are very grateful to June Martino, who did all the processing of the manuscripts.

CONTENTS

NEW METHODS FOR SYNTHESIZING DEOXYOLIGONUCLEOTIDES

M.H. Caruthers, S.L. Beaucage, C. Becker, W. Efcavitch, E.F. Fisher, G. Galluppi, R. Goldman, P. deHaseth, F. Martin, M. Matteucci and Y. Stabinsky

Department of Chemistry
University of Colorado
Boulder, Colorado 80309

INTRODUCTION

Genes and gene control regions can now be routinely characterized and studied at the molecular level. This is possible because of several recent advances in the technology associated with manipulating and modifying DNA. Of particular importance have been advances in DNA sequencing (1,2), the isolation of a large number of host restriction modification enzymes (3,4) and the construction of vectors for cloning and amplifying defined DNA sequences (5,6). Because of these developments, a substantial interest in the synthesis of sequence defined deoxyoligonucleotides has emerged. For example, synthetic deoxyoligonucleotides can be used as probes for isolating natural genes (7,8) and for experiments involving site-directed mutagenesis (9). Moreover, totally or partially synthetic genes can now be prepared, cloned and expressed (10-16). Synthetic deoxyoligonucleotides have also proven to be quite useful for biochemical (17) and biophysical (18,19) studies on how DNA sequences in gene control regions relate to function. However, until recently the synthesis and isolation of deoxyoligonucleotides was a difficult and time-consuming task. Ideally chemical methods should be simple, rapid, versatile and accessible to the nonchemist. We have developed a synthetic methodology which appears to satisfy all these criteria.

The general synthetic strategy involves adding mononucleotides sequentially to a nucleoside covalently attached to an insoluble polymer support. Reagents, starting materials and side-products are then removed simply by filtration. After

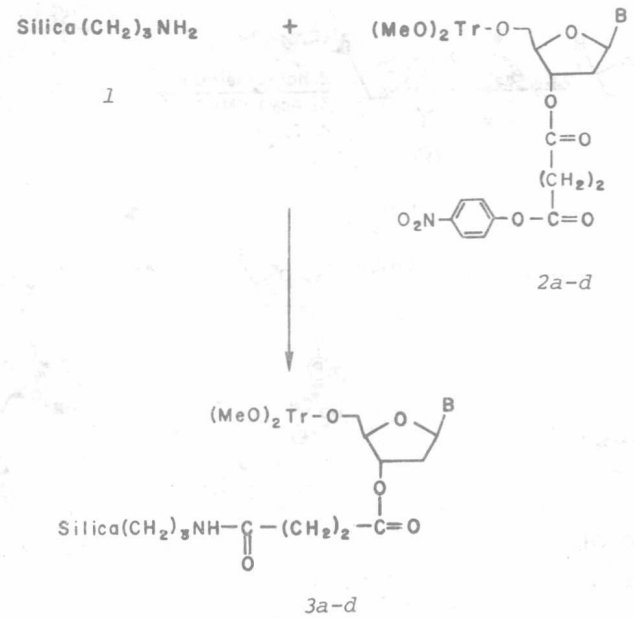
various additional chemical steps, the next mononucleotide can be joined to the growing, polymer-supported deoxyoligonucleotide. At the conclusion of synthesis, the deoxyoligonucleotide is chemically freed of blocking groups, hydrolyzed from the support and purified to homogeneity by polyacrylamide gel electrophoresis. This chapter outlines our current synthetic procedures.

OUTLINE OF THE PROCEDURE

The Polymer Support

The polymer support used throughout this research is derivatized silica gel originally developed for use in high performance liquid chromatography. These polymers are designed for efficient mass transfer. We therefore anticipated and have since shown that reactants can be removed rapidly from these derivatized silica gels after various synthesis steps. Additionally, silica gel is a rigid, nonswellable matrix in common organic solvents. It can be packed into a column and reactants can be pumped through. These features make silica gel an attractive support for deoxyoligonucleotide synthesis. Originally a silica gel, Vydac TP-20 (Separation Group) was used (20). The pore size (300 Å) and particle size (20 μm) of this material was considered optimal for deoxyoligonucleotide synthesis. However, we now use a silica gel called Fractosil 500 (Merck) because of its larger particle size (50 μm). Fractosil is more easily manipulated by the isolation procedures (centrifugation and filtration) most commonly used in the laboratory at the present time. Furthermore, the yields per condensation are similar to those obtained with Vydac TP-20.

The support is prepared as outlined schematically in Figure 1. A detailed description of this procedure has been published (21). The initial step involves forming compound 1 by refluxing 3-aminopropyl triethoxysilane with silica gel in dry toluene for 3 hr. The next step is synthesis of compounds 2a-d. The appropriate base protected 5'-O-di-*p*-anisylphenylmethylnucleoside is reacted initially with succinic anhydride. After an aqueous extraction against citric acid, the succinylated nucleoside is converted to the *p*-nitrophenyl ester with *p*-nitrophenol and dicyclohexylcarbodiimide. This nucleoside is then added to compound 1 in a mixture of dimethylformamide, dioxane and triethylamine. An intense yellow color forms rapidly, indicating the elimination of *p*-nitrophenol and the formation of compounds 3a, 3b, 3c or 3d. Usually the ratio of reagents is adjusted so that approximately 50 μmole nucleoside per gram silica gel is obtained. Thus 1 to 2 μmole of a decanucleotide can be produced from 100 to 200 mg of silica gel. Underivatized amino groups are blocked against further reaction by acylation with acetic



1

2a-d

3a-d

Figure 1. Synthesis of the polymer support. B refers to thymine in 2a and 3a; to N-benzoylcytosine in 2b and 3b; to N-benzoyladenine in 2c and 3c; to N-isobutyrylguanine in 2d and 3d. $(MeO)_2Tr$ designates the di-p-anisylphenylmethyl protecting group.

anhydride. This overall procedure is quite satisfactory for attaching any nucleoside to the polymeric support. Additionally, we observe more consistent, reproducible loading than when compounds 3a, 3b, 3c and 3d are prepared by an alternative procedure (20).

The Synthesis Cycle

The addition of one mononucleotide to compound 3a, 3b, 3c or 3d requires the following four steps: 1. removal of the trityl protecting group with $ZnBr_2$; 2. condensation with the appropriate 5'-O-di-p-anisylphenylmethyl-3'-methoxy-N,N-dimethylaminophosphine nucleoside; 3. acylation of unreactive 5'-hydroxyl groups; 4. oxidation of the phosphite triester to the phosphate triester. These steps are summarized in Figure 2. Thus the synthesis proceeds in a 3' to 5' direction by adding one nucleotide per cycle. The individual steps for one synthesis

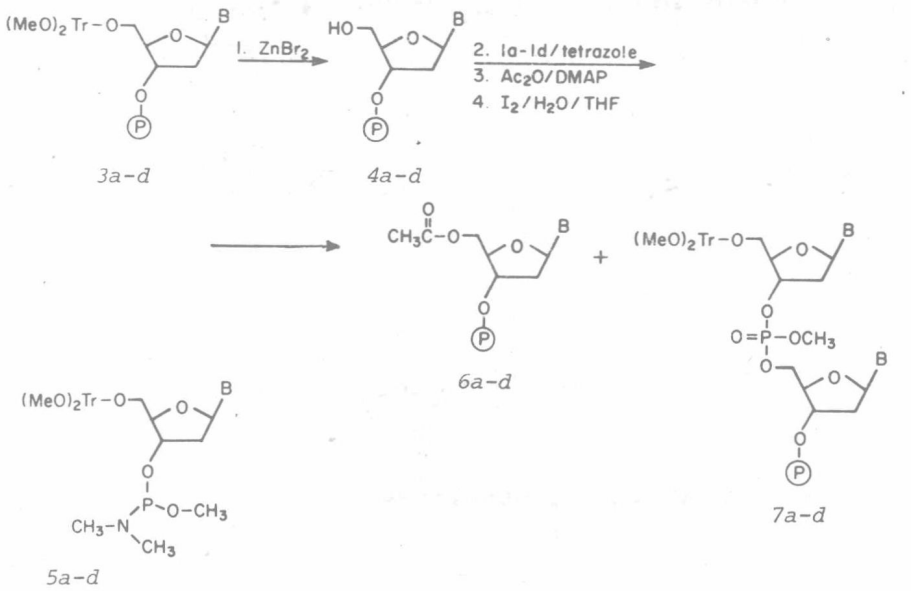


Figure 2. Steps in the synthesis of a dinucleotide. B refers to thymine in 3a, 4a, 5a, 6a and 7a; to N-benzoylcytosine in 3b, 4b, 5, 6b and 7b; to N-isobutyrylguanine in 3d, 4d, 5d, 6d and 7d. $(MeO)_2Tr$ refers to the di-p-anisylphenylmethyl group when 5d is used. For 5a, 5b and 5c, trityl groups as defined in Table 2 are used.

cycle are listed in Table 1 and outlined in the following paragraphs.

Step 1. Removal of the Trityl Ether Protecting Group. p-Toluenesulfonic acid in acetonitrile rapidly removes the 5'-O-di-p-anisylphenylmethyl group from either deoxypyrimidine nucleosides or deoxyoligopyrimidines attached covalently to silica gel (22). However, when purines were examined, this procedure was unsatisfactory (20). The results indicated that repetitive treatment with p-toluenesulfonic acid (as required for any polymer support procedure) would lead to complete depurination. We therefore examined alternative methods for removing trityl ethers and observed that several Lewis acids were satisfactory. Specifically, we discovered that $ZnBr_2$ in nitromethane completely removed trityl ethers from support-bound deoxyoligonucleotides (23) without any associated depurination (20). The

Table 1
Protocol for One Synthetic Cycle

Reagent or Solvent	Purpose	Time (min)
Saturated $ZnBr_2$ in 5% methanol/ CH_3NO_2	Detritylation	5
CH_3NO_2	Wash	3
Acetonitrile	Wash	1
Activated nucleotide in acetonitrile	Add one nucleotide	10
Tetrahydrofuran/H_2O/lutidine (2/2/1)	Wash	3
I_2 solution	Oxidation	5
Acetic anhydride/dimethylamino-pyridine	Cap	5
Methanol	Wash	2
Nitromethane	Wash	1

Table 2
Assignment of Colored 5'-Triarylmethyl Groups to the Protected Deoxynucleotides

Nucleotide	Color
5'-O-di-p-anisylphenylmethyl-N-isobutyryl-deoxyguanosine-3'-N,N-dimethylaminomethoxyphosphine	Orange
5'-O-p-anisyl-1-napthylphenylmethyldeoxy-thymidine-3'-N,N-dimethylaminomethoxyphosphine	Red
5'-O-di-o-anisyl-1-napthylmethyl-N-benzoyldeoxycytidine-3'-N,N-dimethylaminomethoxyphosphine	Blue
5'-O-p-tolyldiphenylmethyl-N-benzoyl-deoxyadenosine-3'-N,N-dimethylaminomethoxyphosphine	Yellow

maximum time required was 30 min. The reaction time can be decreased to 5 min. by adding 5% methanol (21) or 1% water (24).

A recent innovation in this step is to monitor the progress of the synthesis (25). As can be seen by examining the reaction scheme presented in Figure 2, a di-p-anisylphenylmethyl group is removed during each synthetic cycle. More importantly, this di-p-anisylphenylmethyl group is part of the mononucleotide (5a, 5b, 5c or 5d) most recently added to the growing oligonucleotide attached to silica gel. The di-p-anisylphenylmethyl group forms an orange color when removed from the nucleotide during the $ZnBr_2$-catalyzed detritylation step. If four triarylmethyl protecting groups that form distinct colors in $ZnBr_2$/nitromethane could be attached to the different nucleotides, then each condensation step could be individually monitored. For example, if the di-p-anisylphenylmethyl group were assigned only to 5d, then an orange color during the subsequent $ZnBr_2$/nitromethane cleavage step would indicate that N-isobutyryldeoxyguanosine had been added to the growing oligonucleotide during the preceding condensation step. By measuring the color intensity, an approximate quantitation of the condensation yield is also possible. A different color resulting from another triaryl protecting group would indicate that one of the other mononucleotides (5a, 5b or 5c) had been added during the condensation. By attaching different triarylmethyl protecting groups to the four mononucleotides, condensation reactions can be monitored with respect to yield and nucleotide sequence. The triarylmethyl protecting groups that we presently use and their associated colors in $ZnBr_2$/nitromethane solutions are listed in Table 2. Thus, the additions of thymidine, cytidine, adenosine and guanosine are monitored by the formation of red, blue, yellow and orange colors, respectively, during the detritylation step.

Step 2. Condensation of Activitated Nucleotides to the Polymer Support. Appropriately protected nucleoside phosphoramidites (5a-d) are used for the sequential addition of mononucleotides. These compounds have several attractive features. They can be synthesized using standard organo-chemical procedures (26) and can be stored as white powders. They are also resistant to aqueous hydrolysis and to oxidation. We therefore store, measure and transfer these reagents in a manner analogous to any other solid compound. Activation in dry acetonitrile is achieved by the addition of a mild acid. We have selected tetrazole as the activating agent since it is a nonhygroscopic, commercially available material. Other mild acids such as amine hydrochlorides can also be used, but are not recommended. Amine hydrochlorides are hygroscopic and would introduce water into the condensation step. Our synthesis procedure therefore involves dissolving the nucleoside phosphoramidite and tetrazole in acetonitrile, mixing and adding the solution to silica gel support. The condensation reaction is complete within a minute, but we usually allow the reaction to proceed for five minutes.

Step 3. Capping Unreacted 5'-Hydroxyl Groups. Based on detailed and careful analysis of various condensation reactions, approximately 1 to 5% of the deoxynucleoside or deoxyoligonucleotide bound to the support (compound 4a-d) does not react with the activated nucleotide to form 7a-d. These unreactive compounds must be blocked or capped in order to prevent the formation of several deoxyoligonucleotides with heterogeneous sequences. Moreover, in the absence of a capping step, most of these unwanted sequences will contain one or two less nucleotides than the expected product, making purification more difficult. If a capping step is included, the unwanted sequences will have diverse sizes with only 1 to 5% of the reaction mixture having one nucleotide shorter than the product. Currently this capping step can best be accomplished using a tetrahydrofuran solution of acetic anhydride and dimethylaminopyridine. The reaction is complete within 1 to 2 min. and does not lead to any detectable side products.

Step 4. Oxidation of the Phosphite to the Phosphate. The internucleotide phosphite triester is oxidized to the phosphate triester with the use of I_2 in water, 2,6-lutidine and tetrahydrofuran. Oxidation is extremely rapid (1 to 2 min.) and side products are not generated. We have attempted to postpone the oxidation until after all condensation steps, but the results have not been encouraging. Several uncharacterized side products are observed.

Synthesis Machines. We use several devices to aid in the synthesis of deoxyoligonucleotides. The schematic for a manually-operated apparatus used routinely in our laboratory is shown in Figure 3. Machine-assisted synthesis of deoxyoligonucleotides begins by loading the column with 0.20 to 0.30 g of derivatized silica gel. The protocol outlined in Table 1 is then followed to add one nucleotide. Solvents and reactants can be selected through a series of three-way valves and then either cycled once or recycled many times through the column. An injector port is used for the addition of activated nucleotides and the efficiency of various wash cycles is monitored with a spectrophotometer. We are presently designing a completely automatic, microprocessor-controlled machine that is based on this simple design. Other less complex devices are also used for synthesizing deoxyoligonucleotides. For example, we have successfully synthesized deoxyoligonucleotides in test tubes and on sintered glass funnels. In the former case, each synthesis step is followed by a low speed centrifugation and decantation in order to remove solvents and reagents. For the latter case, reagents and solvents are removed simply by filtration. When test tubes or sintered glass funnels are used, 6 to 8 deoxyoligonucleotides are synthesized simultaneously. The appropriate mononucleotide is added to each test tube during a synthesis cycle. These additions are monitored by observing the color during the subsequent detritylation step.

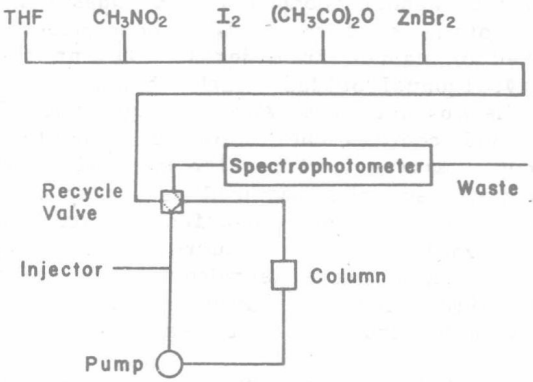
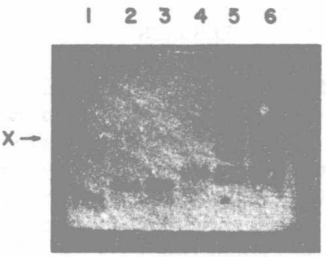


Figure 3. Schematic diagram of a simple, manual machine. Five flasks containing two solvents (tetrahydrofuran and nitromethane) and three reactants (iodine, acetic anhydride and zinc bromide) are represented in the top part of the diagram.

Isolation of Deoxyoligonucleotides

Once a synthesis has been completed, the deoxyoligonucleotide is freed of protecting groups and isolated by polyacrylamide gel electrophoresis. Silica gel containing the reaction product is first treated with triethylammonium thiophenoxide in dioxane (20) in order to remove the methyl groups from internucleotide phosphotriesters. This step is followed by treatment with concentrated ammonium hydroxide at 20°C for 3 hr. in order to hydrolyze the ester joining the deoxyoligonucleotide to the support. After centrifugation and recovery of the supernatant containing the deoxyoligonucleotide, the N-benzoyl groups from deoxycytosine and deoxyadenosine, and the N-isobutyryl group from deoxyguanosine are removed by warming the concentrated ammonium hydroxide solution at 50°C for 12 hr. Finally the trityl ether is hydrolyzed with 80% acetic acid. This is the preferred detritylating agent after the amino protecting groups have been removed. Depurination is not observed with completely unprotected

GEL ELECTROPHORESIS



1. d(G-T-T-C-G-T-C-A-G-A-A-G-A-G-T-C)
2. d(A-A-T-T-C-T-A-T-T-A-G-T-C-T-T-T-A)
3. d(C-G-T-A-A-T-A-A-A-G-A-C-T-A-A-T-A-G)
4. d(G-C-G-T-T-C-C-C-T-G-A-C-T-C-T-T-C-T-G-A)
5. d(C-G-A-A-C-C-T-C-C-A-G-G-A-A-C-G-C-C-T-G)
6. d(T-T-A-C-G-C-A-G-G-C-G-T-T-C-C-T-G-G-A-G)

Figure 4. Purification of deoxyoligonucleotides synthesized on silica gel. Lanes 1 to 6 show the results obtained from six syntheses completed on silica gel. Each lane represents a total hydrolysate from silica gel after removal of all deoxyoligonucleotide protecting groups. The sequence of each deoxyoligonucleotide is listed below the gel pattern. X represents xylene cyanol blue which was added to lane 1 as a dye marker. The dye absorbs ultraviolet light and appears as a dark band.

deoxyoligonucleotides and, unlike $ZnBr_2$, 80% acetic acid is volatile and easily removed. The reaction mixture containing unprotected deoxyoligonucleotides is then fractionated by polyacrylamide gel electrophoresis. When a standard slab gel device

is used, as many as eight compounds can be purified simultaneously. Usually these reaction mixtures are fractionated on 20% acrylamide gels with the standard Tris-borate buffer system (27). The appropriate deoxyoligonucleotide is visualized with an ultraviolet light. Elution from the gel is completed using standard procedures (1).

RESULTS

Deoxyoligonucleotides containing from 10 to 26 mononucleotides have been synthesized with the approach described. A typical result is shown in Figure 4. Deoxyoligonucleotides numbered 1 through 6 were simultaneously synthesized on sintered glass funnels. Following removal of all protecting groups, total hydrolysates from silica gel were applied to a polyacrylamide gel and fractionated by electrophoresis. (A photograph of the gel irradiated with ultraviolet light is reproduced in Figure 4.) Analysis of this gel showed that one major band corresponding to the expected deoxyoligonucleotide was observed for each synthesis. Occasionally a synthesis fails or several additional bands corresponding to significant amounts of various failure sequences are observed. These problems appear to be due to mistakes made during the many steps required to complete manually a large number of simultaneous syntheses. The synthetic deoxyoligonucleotides can be eluted and characterized by two-dimension sequence analysis. Examples are illustrated in Figure 5. Panel A shows the analysis of a deoxyoligonucleotide (segment 9_p) which is a part of a hybrid E. coli promoter shown in Figure 6. The main objective of this experiment is to analyze for homogeneity. Unlike any other sequencing technique, the total sample is displayed (including impurities and side-products). Therefore the chemical purity as well as the sequence may be determined. As shown by the presence of a single series of spots corresponding to the correct sequence, the deoxyoligonucleotide is homogeneous. Panel B shows a similar analysis for a deoxyoligonucleotide that forms part of a synthetic SV40 T-antigen binding site (28). Deoxyguanosine-rich deoxyoligonucleotides historically have been extremely difficult to prepare using previously developed chemical synthesis procedures. We have not observed similar problems with the chemistry outlined here. The compound whose analysis is reproduced in Panel B is very rich in deoxyguanosine residues (6 of 9) and yet it is quite homogeneous when analyzed by the two-dimension method.

Chemically-synthesized deoxyoligonucleotides form DNA duplexes that are recognized by various enzymes. A typical example is illustrated in Figures 6, 7 and 8. A major objective of our research effort is to develop an E. coli promoter that will maximally express a cloned gene under controlled conditions. One E. coli promoter presently being examined extensively has the

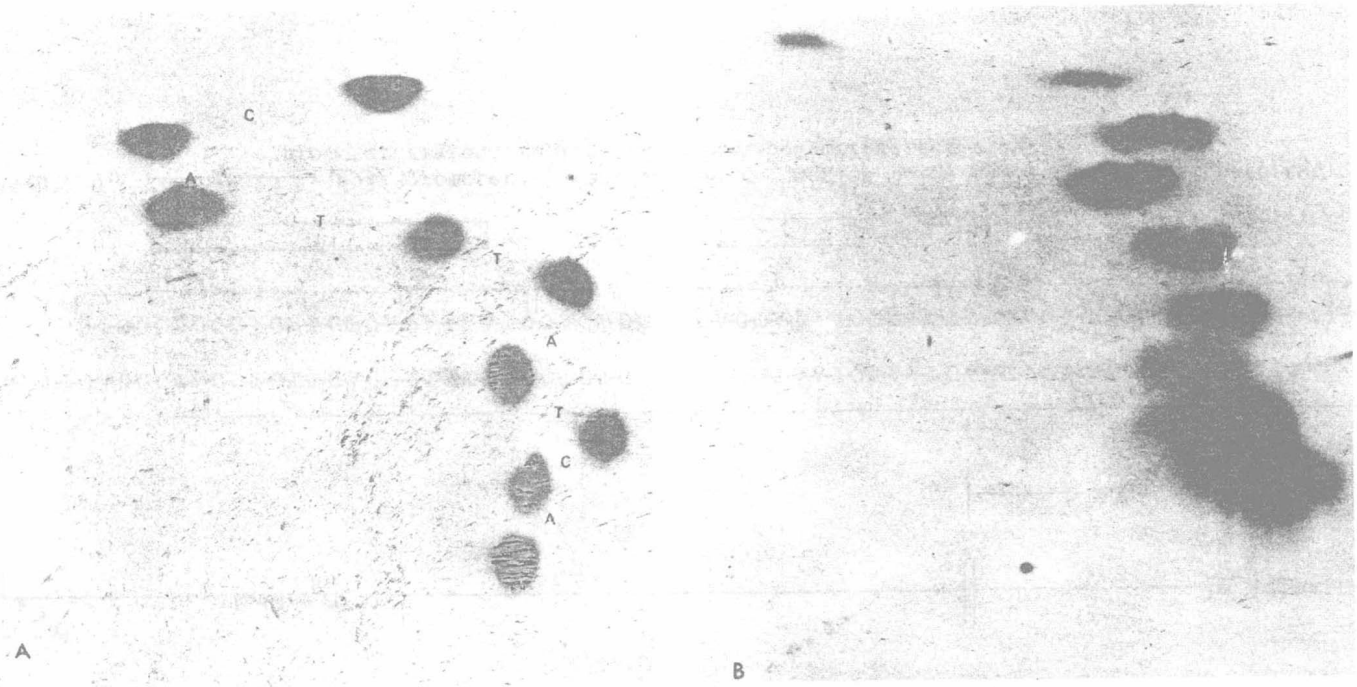


Figure 5. Two-dimension sequence analysis. Panels A and B show analyses of d(T-C-A-T-T-A-T-C-A) and d(G-T-G-A-G-G-A-G-G), respectively.