UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
UNIDAD ACADEMICA DE LOS CICLOS PROFESIONAL Y
DE POSGRADO DEL C.C.H.

USO DE VEHICULOS MOLECULARES DE CLONACION Y DE SINTESIS QUIMICA DE POLINUCLEOTIDOS PARA EL ESTUDIO DE LA EXPRESION GENICA

Tesis para obtener el Título de
MAESTRO EN INVESTIGACION BIOMEDICA BASICA
PRESENTA

Francisco Xavier Soberón Mainero

Abril de 1982.

UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

El trabajo que se reporta en esta Tesis fue realizado en el Departamento de Biología Molecular del Instituto de Investigaciones Biomédicas de la Universidad Nacional Autónoma de México y en el Departamento de Genética Molecular de City of Hope Research Institute, Duarte California.

El trabajo fue supervisado por los Doctores Francisco Bolivar y Keiichi Itakura, respectivamente.

1) INTRODUCCION

1.1) Objetivos.

Los objetivos del trabajo de esta tesis se encuentran relacionados con dos desarrollos metodológicos recientes: la recombinación *in vitro* de ácidos nucléicos y la síntesis química de polinucleótidos.

Se intentó desarrollar tres aspectos básicos con consecuencias en el mejoramiento metodológico y aplicación al conocimiento básico de sistemas de expresión genética.

a) Construir vehículos de clonación molecular particularmente útiles para la clonación y utilización de segmentos sintéticos de DNA y para el estudio de la expresión de promotores procariontes.

b) Investigar aspectos importantes para simplificar la producción de fragmentos sintéticos de DNA de doble cadena.

c) Obtener un sistema útil para la investigación de la estructura y función de promotores de *E. coli*.

1.2.1) Vehículos de clonación molecular.

El vehículo de clonación molecular constituye un factor fundamental en las técnicas de recombinación in vitro de ácidos nucléicos. No es solamente el medio utilizado - para mantener el DNA recombinante establemente dentro de la célula sino que se estructura determina en muchos casos la estrategia y las diversas alternativas que se tienen para - un experimento de clonación (Bolivar et al, 1977; Bolivar y Backman, 1979).

Dos vehículos moleculares que han sido utiliza-- dos en un gran número de proyectos son los plásmidos pBR322 y pBR325 (Bolivar et al, 1977; Bolivar, F., 1978). Las características de estos vehículos pueden ser modificadas en forma ventajosa, eliminando material dispensable presente - en ellos. La región comprendida entre el origen de replicación y el extremo 3' del gene que codifica para resistencia a tetraciclina contiene algunas secuencias que son dispensables para las funciones de replicación y resistencia a antibióticos del plásmido; estas secuencias, sin embargo, con-- fieren al el plásmido la posibilidad de ser transferido en ciertas condiciones (Young and Poulis, 1975) y codifican por lo menos para dos polipéptidos pequeños (menos de 10 Kd) dispensables para las funciones esenciales del plásmido (Cova

rrubias et al, 1981). La eliminación de este segmento produce, por lo tanto, varias ventajas al plásmido como vehículo de clonación:

1) Eliminación de bandas en patrones de restricción. Como consecuencia se facilita el análisis del mapa físico de segmentos de DNA clonados en él.

2) Eliminación de bandas en geles de análisis de proteínas en sistemas de minicélulas o de transcripción y traducción in vitro, dependiente de DNA.

3) Mayor confinamiento biológico, dada la disminución en la posibilidad de transferencia del plásmido por conjugación.

A pesar de la existencia de vehículos de clonación muy versátiles y de un buen número de ellos, (Collins, J. 1979), proyectos específicos siguen requiriendo de la construcción de vehículos con características diferentes a las ya existentes.

Por otra parte, el desarrollo de la síntesis química de polinucleótidos (ver capítulo 1,2.2) ha generado

una demanda de vehículos moleculares especialmente útiles - para la aplicación de esta metodología.

Dos características particularmente importantes en los vehículos de clonación en este contexto son su tamaño y el número de sitios de restricción únicos que contengan.

Dada la limitación presente en la síntesis química de polinucleótidos en cuanto a la longitud de las cadenas que se pueden sintetizar directamente (no más de 40 nucleótidos), (Dembeck e Itakura, 1981), un vehículo molecular pequeño y 100% caracterizado a nivel de secuencia nucleotídica resulta importante en la manipulación de estos fragmentos.

Asimismo, la presencia de buen número de sitios de restricción únicos facilita el diseño de las piezas intermedias para construir secuencias más largas de DNA de doble cadena.

Los vehículos pBR327 y pBR328 reúnen estas características mejoradas y deben resultar particularmente útiles a este respecto.

1.2.2) Síntesis Química de Polinucleótidos.

1.2.2.1) Aplicaciones.

La convergencia del desarrollo de las técnicas de recombinación in vitro de ácidos nucléicos y la síntesis - química de polinucleótidos ha conferido a esta última una - relevancia particular como herramienta para el estudio básico de la expresión genética y en el diseño de sistemas para la expresión de DNA heterólogo en microorganismos.

Algunos de los aspectos en que destaca la aplicación de esta metodología son los siguientes:

a) Construcción y expresión de genes sintéticos.

Algunos genes que codifican para polipéptidos de interés en la medicina han sido sintetizados total o parcialmente e introducidos en células bacterianas o de levadura y empiezan a ser producidos por fermentación (Goedel et al, - 1979, a y b).

b) Estudio de regiones regulatorias.

El estudio de regiones de DNA que interaccionan - con proteínas regulatorias específicas ha sido facilitado - mediante la síntesis de estas regiones y análogos de ellas

(Kawashima et al, 1977, Goedel, et al, 1978). Este enfoque parece muy prometedor en el esclarecimiento de la relación estructura-función de estas regiones del DNA.

c) Conformación de ácidos nucléicos.

La posibilidad de sintetizar cantidades de miligramos de secuencias específicas y cortas de DNA en forma altamente pura, ha permitido la obtención de cristales de estas moléculas y su estudio por cristalografía de rayos X hasta una resolución atómica. Estos estudios han confirmado la validez de algunos de los modelos estructurales propuestos para el DNA (Wing, et al, 1980; Conner, et al, 1981) y han demostrado la existencia de estructuras alternativas nunca antes propuestas, altamente sugestivas en cuanto a su posible relevancia biológica (Wang, et al, 1977).

d) Mutagénesis dirigida.

Los polinucleótidos sintetizados químicamente proveen también de la posibilidad de realizar alteraciones totalmente específicas en sitios determinados de una secuencia de DNA. Este procedimiento ha sido utilizado con éxito en la producción de mutaciones puntuales y deleciones (Wallace, et al, 1980, Gillam and Smith, 1979), en varios sistemas.

Cabe destacar que el mismo oligonucleótido que se usa para producir una mutación es útil también para la detección de las clonas que la adquirieron al usarse como rastreador de hibridización. Así pues, la obtención y estudio de mutaciones específicas no requiere de un fenotipo seleccionable.

e) Aislamiento de genes específicos.

A la fecha se han usado con éxito rastreadores de hibridización sintéticos para la detección de secuencias específicas de RNA representadas en porcentajes del orden del 1% en la población de mensajeros (o un banco de cDNA correspondiente.) Este procedimiento consiste en el diseño y síntesis de un segmento de DNA cuya secuencia sea complementaria a la de una región del mensajero que se busca. Esto se hace basándose en la secuencia del polipéptido codificado por el mensajero (la degeneración del código genético hace que pueda ser necesaria la síntesis de varios segmentos o mezclas de ellos para asegurarse de que se tiene la secuencia requerida). Este enfoque debe resultar altamente fructífero en el aislamiento de genes cuyos mensajeros se encuentran en baja proporción. (Agarwal, et al, 1981).

1.2.2.2) Metodología: el enfoque del triester.

El objetivo buscado en la síntesis química de polinucleótidos es la formación sucesiva de enlaces entre los cuatro diferentes nucleótidos: desoxiadenilato, desoxicitidilato, desoxiguanilato y Timidilato. La formación de estos enlaces debe ser específica, tanto en lo que se refiere a las posiciones del nucleótido que reaccionan (amino 5', 3', etc) como, desde luego, en lo que se refiere al orden o secuencia en la formación de estos enlaces.

La materia prima para la síntesis consiste en los cuatro desoxinucleosidos que constituyen el DNA. Estos nucleósidos deben ser protegidos selectivamente en sus posiciones 5' y 3' y en los grupos amino Número 6, 2 y 4 de las bases adenina, guanina y citosina, respectivamente. La protección de los grupos amino es necesaria para evitar que reaccionen durante la síntesis de las cadenas DNA y produzcan cadenas ramificadas. Los grupos 3' y 5' son protegidos para poder controlar el sentido del crecimiento de las cadenas, es decir para producir siempre enlaces 5' $\longrightarrow$ 3', evitando los compuestos simétricos 3' - 3' y 5' - 5'.

La elección de los grupos protectores presentó un problema importante y fue fundamentalmente solucionado por el clásico trabajo de Khorana y colaboradores al final de los años '50 y la década de los '60. (Khorana et al, 1956; Agarwal et al, 1972).

La selección de los grupos protectores se basó en los siguientes criterios:

1) Relativamente fácil de introducir.
2) Estable durante las condiciones de reacción.
3) Fácil de remover al final de la síntesis, bajo condiciones en las cuales el producto es completamente estable.
4) La remoción de grupos protectores transitorios debe ser selectiva; esto es, el resto de los grupos protectores deberá ser estable durante este proceso.

Un buen cuerpo de investigación apoya el uso generalizado en la actualidad, de los siguientes grupos protectores (Reese, 1978):

1) Grupos 6 amino de adenina y 4 amino de citosina: benzoilo.
2) Grupo 2 amino de Guanina: isobutirilo.
3) Grupo 5' oxhidrilo: Dimetoxitrilo (4-4'-dimetoxi-trifenil metilo).

Los grupos protectores de las funciones amino - son estables bajo las condiciones de reacción de síntesis y son removidos por amoniólisis en condiciones bajo las cua-- les el DNA es completamente estable.

El grupo dimetoxitrilo es estable al alcali pero se hidroliza fácilmente en condiciones ácidas suaves.

La fosforilación de los nucleótidos protegidos - en sus grupos amino y 5' OH ha sido objeto de intensa inves tigación. Los trabajos iniciales se realizaron utilizando fosfato inorgánico y un reactivo acoplador (dicidohexilcar- bodiimida) y se desarrollaron para dar origen al método - llamado del fosfodiester (Argawal et al, 1972). Este meto- do de síntesis fue posteriormente sustituído por los del - fosfotriester y del fosfito, que son los usados en la actua lidad (Itakura, K., 1980).

El método del fosfotriester fue usado en el presente trabajo y será descrito, a grandes rasgos, a continua ción. (Broka et al, 1980).

Como se describe en la figura 1, la fosforila-- ción de los nucleosidos protegidos en amino y 5' se realiza

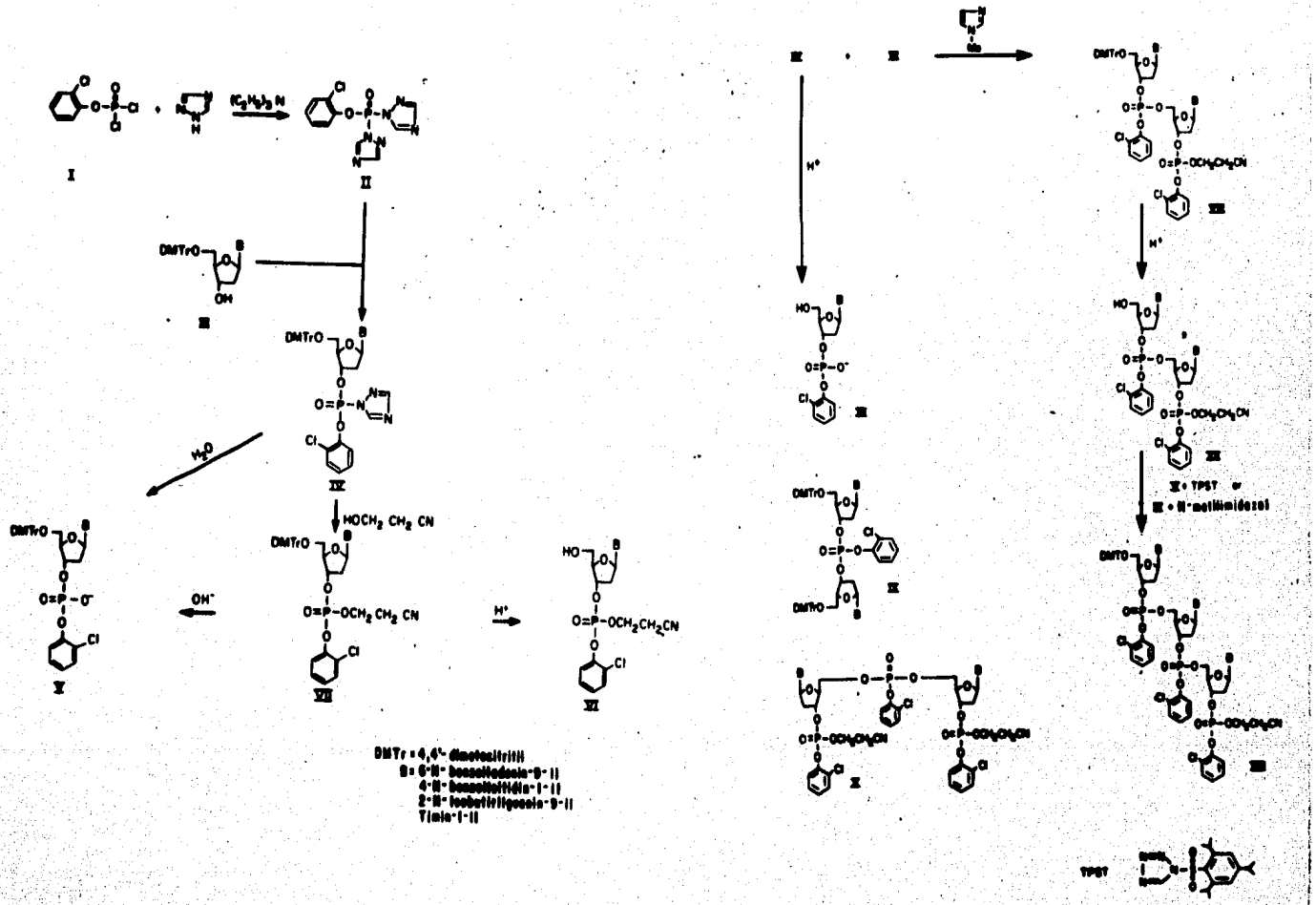
sobre el oxhidrilo libre 3' mediante el uso de 2-clorofenil fosforoditriazólido (II). Este reactivo fosforilador presenta la ventaja de ser esencialmente monofuncional en ausencia de catalizadores y resulta en la producción del monotriazólido (IV) del nucleósido correspondiente (III).

El monotriazólido es un intermediario versátil y puede ser activado para reaccionar en su extremo 3' fosfato al adicionar un nucleófilo y un catalizador o puede ser hidrolizado al correspondiente fosfato (V).

La reacción de IV con el nucleófilo 3-hidroxipropionitrilo, en presencia de 1-metilimidazol como catalizador, genera el nucleósido completamente protegido (VII). Este intermediario es muy estable (meses a -20°C) y puede ser activado selectivamente en sus extremos 5' ó 3' en presencia de ácido (destritilación, para dar VI) o en condiciones alcalinas suaves (descianoetilación, para dar V).

A partir de estos intermediarios, la síntesis se lleva a cabo mediante adiciones sucesivas de los nucleósidos específicos requeridos para una secuencia en particular.

Figura 1. Descripción esquemática de los principales intermediarios en la síntesis de bloques de triester completamente protegidos. Intermediarios clave: IV: fosfomonotriazol, monómero activado en 3'; VII: monómero completamente protegido, puede activarse en 5' ó 3'; VIII: dímero completamente protegido; XIII: trímero completamente protegido.


DMTr = 4,4'- dimetoxitritil
B = 6-N- benzoiladenin-9- il
4-N- benzoilcitidin-1- il
2-N- isobutirilguanin-9- il
Timin-1- il
TPST

Si se hacen reaccionar los intermediarios (IV) - y (VI) en presencia de N-metilimidazol, se obtiene el dinucleótido correspondiente. A partir de este, y después de - activarlo en su posición 5' puede repetirse la reacción con IV y obtener un trinucleótido. Utilizando variantes de esta estrategia se pueden ir adicionando nucleótidos o bloques (di y trinucleótidos) hasta construír el oligonucleótido requerido (Crea et al, 1978).

El uso de reactivos aclopadores como trisisopropio bencesulfonil tetrazólido (TPST) sobre los intermediarios 5' OH y 3' fosfato es también una ruta adecuada para - la obtención de los enlaces fosfotriester.

Las reacciones anteriores son llevadas a cabo en solución en piridina y en condiciones anhidras. Los intermediarios de la síntesis (oligómeros) son purificados cromatográficamente (sílica gel) después de cada reacción de acoplamiento.

El método de síntesis en solución descrito constituye un sistema adecuado para la preparación de cantidades grandes de DNA (100 mg de producto final o más).

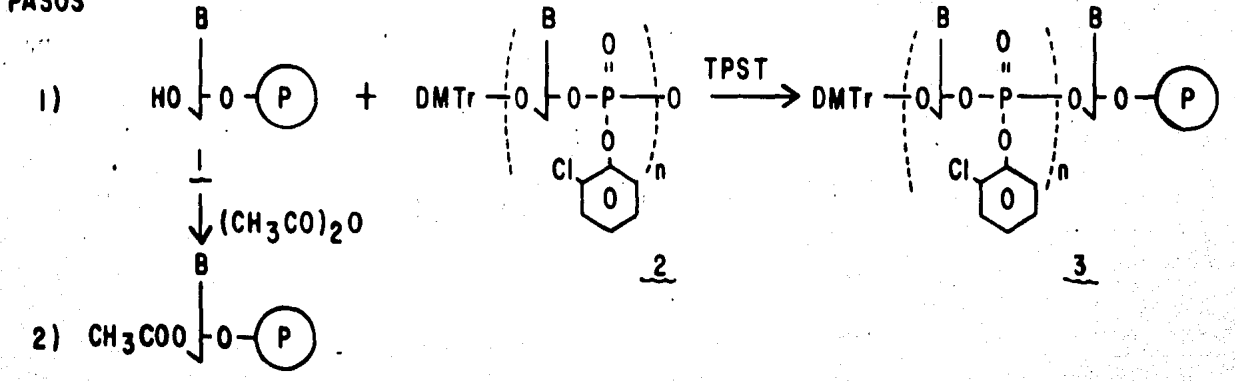
Alternativamente, para la preparación de cantidades menores de polinucleótidos, suficientes para la gran mayoría de las aplicaciones ya mencionadas, se utiliza un soporte sólido y se substituye el paso de purificación después de cada adición. En este caso los subproductos y los reactivos no reaccionados son lavados por filtración de la resina sobre la que la cadena va creciendo, (ver figura 2).

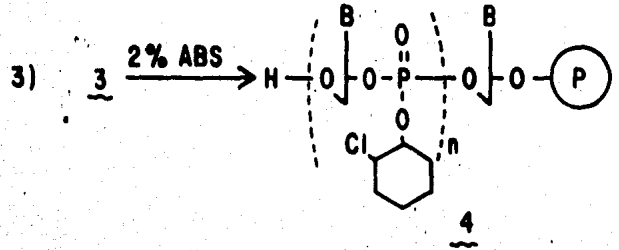
Como último paso, se remueven los grupos protectores y se purifica el producto terminado por cromatografía líquida de alta performancia y/o electroforesis en gel de poliacrilamida (Miyoshi et al 1980).

1.2.3) Elementos para el diseño de un promotor "consenso" de E. coli y su clonación en un vehículo molecular.

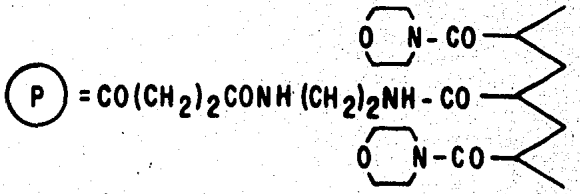
Figura 2. Síntesis en soporte sólido. Los pasos principales son 1) acoplamiento, 2) bloqueo de cadenas no reaccionadas 3) remoción del grupo protector 5'. Entre cada paso la resina se lava varias veces. Antes del primer paso. la resina se seca por coevaporación al vacío con piridina.

PASOS





n = 2 y 3



DMTr = 4', 4' - dimetoxitritilo

TPST = 2, 4, 6 - triisopropilbencensulfonil Tetrazolido

ABS = Acido bencensulfonico

B = Timina, N- Benzoil adenina
N-Benzoil citosina, N-iso Butiril guanina

En los últimos años, la disponibilidad de un buen número de reportes de regiones regulatorias del genoma de E. coli, caracterizadas a nivel de secuencia nucleótida, ha permitido un estudio comparativo de estas regiones y la generación de hipótesis respecto a la estructura general que constituye el promotor.

En E. coli, una sola enzima, la RNA polimerasa, es responsable de la transcripción de todos los genes y operones. Este hecho hace suponer que debe encontrarse un conjunto de características estructurales específicas, comúnes, en todas las secuencias responsables de la iniciación de la transcripción en este organismo.

El estudio comparativo de diferentes secuencias con actividad de promotor ha destacado interesantes semejanzas en cuanto a la secuencia nucleotídica de ellas.

Desde el principio resultó aparente la similitud encontrada en la región correspondiente a los nucleótidos -7 a -12 (+1 es el primer nucleótido que se transcribe (Pribnow, D. 1975)). Posteriormente estas semejanzas fueron extendidas a otra región. entre los nucleótidos -30 y -35. (Schal

ler et al, 1975). Cabe destacar que para que exista una máxima homología en las dos regiones mencionadas se debe permitir una variación en la distancia entre las dos de más o menos dos pares de bases.

En algunos trabajos de revisión, se han destacado estas similitudes después de comparar más de 50 promotores de E. coli. Existen tres aspectos fundamentales que destacan la importancia de estas regiones: Por una parte, se observa que cada una de las posiciones mencionadas está ocupada en la mayoría de los casos pur una base determinada (esto es estadísticamente significativo). En segundo lugar, la gran mayoría de las mutaciones encontradas hasta la fecha y que producen un aumento o disminución de la fuerza del promotor correspondiente, se encuentran alterando alguna de estas dos regiones de homología. Más aún, por lo regular, las mutaciones que desvían al promotor de la secuencia "consenso" disminuyen la actividad del promotor y viceversa. Por último, experimentos de protección y de bloqueo con reactivos químicos que modifican el DNA confirman la importancia de estas regiones (Siebeneist et al, 1980).

A pesar de estas observaciones, no se ha llegado aún al punto en el que a partir de la secuencia nucleotídica se pueda predecir la eficiencia de un promotor o siquie-

ra si un segmento de DNA debe tener actividad de promotor. Tampoco existe una idea clara respecto a la relación que guardan las regiones centradas en -10 y -35 con la función del promotor.

Tomando estos antecedentes en consideración se decidió realizar la síntesis química de un segmento de DNA que incorpora las características siguientes:

1) Secuencias "consenso" en las regiones -10 y -35. (ningún promotor natural contiene estas dos regiones consenso simultáneamente).

2) Un sitio de restricción en posición intermedia entre estas dos regiones. Esto se decidió con el objeto de poder recombinar en forma sencilla ambas regiones con otros segmentos de DNA, así como variantes de ellas unas con otras.

Adicionalmente, se introdujo en el diseño un sitio de restricción en un extremo de la molécula a sintetizar para facilitar la subsecuente clonación del segmento.

El resto de la secuencia fue diseñada considerando la simplificación del esquema de síntesis química y evitando secuencias palindrómicas que pudieran estorbar en manipulaciones subsecuentes.

En la figura 3 se muestra la secuencia "consenso" derivada de la revisión de 54 promotores (Siebenlist et al - 1980) y el diseño para el promotor sintético.

Para el diseño del promotor se consideró también la posibilidad de obtener un sistema eficiente para la expresión de genes. Cabría suponer que una secuencia consenso - que contiene las bases más conservadas de acuerdo con el estudio de un buen número de promotores, es un candidato para constituír un promotor fuerte. Por otra parte, resulta interesante como primera aproximación para un estudio sistemático de la relación estructura-función de estas regiones regulatorias.

2) RESULTADOS.

2.1) Obtención de deleciones de los plásmidos - pBR322 6 pBR325.

Como se detallan en la publicación anexa (Soberón, et al, 1980), se construyeron dos vehículos moleculares con - características mejoradas respecto a sus predecedores pBR322

Figura 3. a) Comparación de la secuencia "consenso" obtenida a partir de la revisión de 54 promotores y la secuencia diseñada por el promotor sintético.

b) El promotor sintetizado y sus principales características.

A.

Compilación de 54 promotores (Siebenlist et al, 1980):

AA(A/T)ATAATT(C/G)TTGACATTTTTT-TACTATT(T/G)GGTATAATGC---CCATCAATAGAT

Promotor sintético:

GAATTCTTGACAATTAGTTAACTATTTGTTATAATGTATTCCCAAGCTT

B.

EcoRI -35 HpaI -10

GGAATTCTTGACAATTAGTTAACTATTTGTTATAATGTATTC
CCTTAAGAACTGTTAATCAATTGATAAACAATATTACATAAG

y pBR325. Estas características (tamaño pequeño, menor número de transcritos (Covarrubias et al, 1981), y gran número de sitios de restricción únicos) son particularmente útiles para la manipulación de DNA sintético (ver Discusión General).

2.2) Síntesis química-enzimática de un promotor consenso de E. coli.

En el manuscrito anexo se describe la síntesis de un promotor consenso de E. coli. Los elementos considerados para su diseño han sido descritos en la introducción.

2.3) Construcción de un vehículo molecular para la detección y ensayo de promotores de E. coli.

El (los) gene(s) que confieren resistencia a tetraciclina en el plásmido pBR322 (y su derivado pBR327) son transcritos a partir de un promotor localizado en la región que se abarca desde la posición ∿8 hasta la posición ∿48 en el mapa del plásmido (Rodríguez et al, 1979).

En la posición 29 de estos plásmidos se localiza un sitio reconocido por la Endonucleasa HindIII, (Bolivar et al, 1977); este sitio ha sido usado en numerosas ocasiones para clonar frgamentos de DNA y la inserción de estos fragmentos inactiva, en la gran mayoría de los casos, la

resistencia a tetraciclina.

Tomando en consideración estos antecedentes, se decidió producir una deleción de aproximadamente 20 pares de bases (p.b.) alrededor del sitio de HindIII en el plásmido pBR327, como objeto de destruir la actividad del promotor localizado en esta región; esta operación debería conferir a las células que llevan este plásmido una sensibilidad a tetraciclina comparable a las células sin plásmido. La clonación sucesiva de un segmento con actividad de promotor en la orientación apropiada en este misma región debe restituir la resistencia a tetraciclina. Esto es, se debe obtener en vehículo apropiado para la selección de promotores (ver discusión general).

La estrategia seguida para la construcción de este vehículo se detalla en la figura 4. Las condiciones experimentales se describen en la leyenda de la figura.

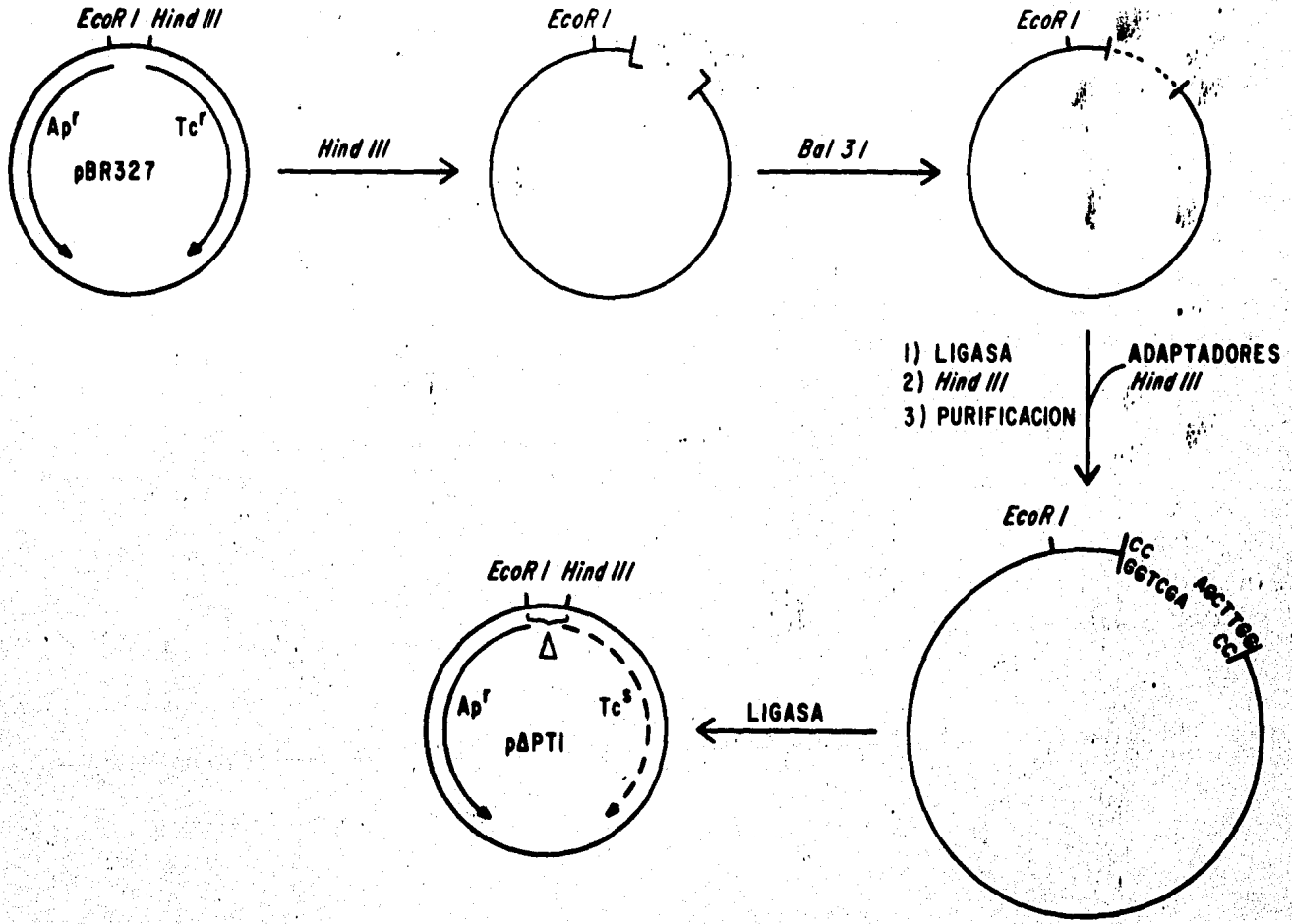
III) DISCUSION GENERAL Y PERSPECTIVAS.

3.1) Uso de los vehículos moleculares construídos.

En esta sección se intentarán destacar algunas estrategias metodológicas que se ven facilitadas por las características de los vehículos moleculares construídos.

Figura 4. Construcción de una deleción en el promotor de los genes de resistencia a tetraciclina. El vehículo pBR327 (3µg) fue digerido con la endonucleasa Hind III (tres unidades, 90 min), extraído con fenol y precipitado con 2.2 volúmenes de etanol. El DNA fue resuspendido en 50 µl de Tris-HCl 20mM pH8, $CaCl_2$ 12mM, $MgCl_2$ 12mM, NaCl 500mM y tratado con tres unidades de nucleasa Bal 31 (Bethesda Research Laboratories) durante 15 segundos a 20°C. La reacción fue terminada mediante la adición de un volumen de fenol saturado con agua. Después de precipitar el DNA se resuspendió em 20µl de Tris-HCl 66mM pH7.6, $MgCl_2$ 6.6mM, ditiotreitol 10mM, conteniendo además ATP o.5mM, 14 moles (exceso de 10 respecto al vehículo) del oligonucléotido d $^{5'}(CCAAGCTTGG)^{3'}$, y 0.5 unidades de polinucleótido ligasa de T4. El producto de esta reacción se sometió a extracción con fenol, se precipitó de etanol y se sometió a digestión con la nucleasa Hind III (10 unidades, e hrs). Después de extraer y precipitar se fraccionó el resultado de la digestión en un gel de agarosa al 1% y se seccionó el pedazo del gel que contenía

el DNA del tamaño del plásmido, éste se sometió a electroelución (Rossi et al 1979) y fue nuevamente tratado con ligasa en las misas condiciones usadas anteriormente excepto que el volumen final fue de 150 µl para favorecer la recircularización del plásmido. La mezcla de ligado se usó para transformar la cepa RRI y de las transformantes obtenidas en un medio selectivo para resistencia a ampicilina se buscaron aquéllas que fueron sensibles a tetraciclina. De entre éstas, se seleccionó una que mostró una delección de ∿20 p.b. al rededor del sitio original de la endonucleasa -- Hind III y se denominó pΔPT1.


EcoR I Hind III
Ap^r
Tc^r
pBR327
Hind III
EcoR I
Bal 31
EcoR I
1) LIGASA
2) Hind III
3) PURIFICACION
ADAPTADORES
Hind III
EcoR I
CC
GGTCGA
AGCTTGG
CC
LIGASA
EcoR I Hind III
Ap^r
Tc^s
pΔPTI

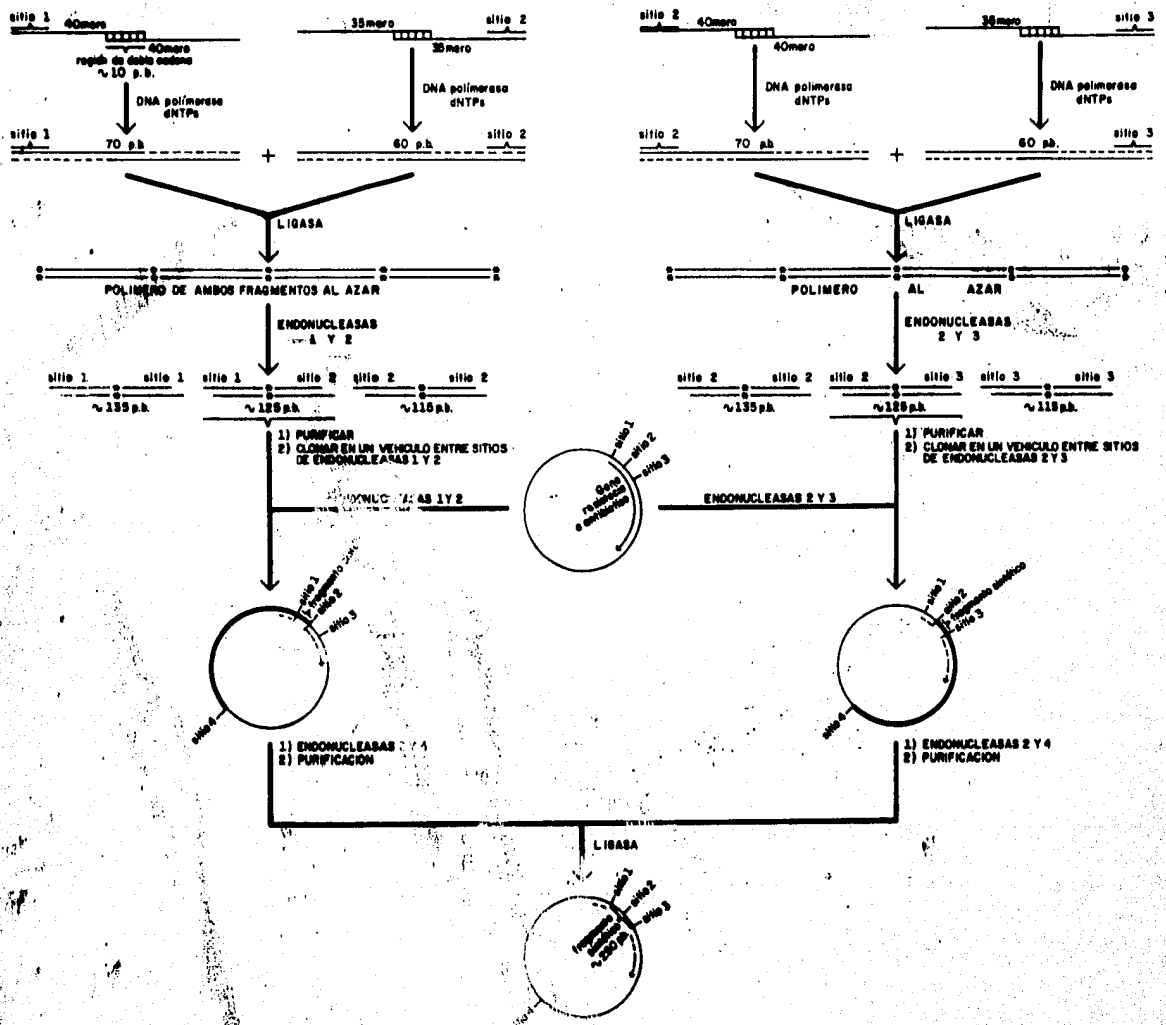
3.1.1) Clonación de segmentos de doble cadena.

Una de las aplicaciones importantes de la síntesis química de polinucleótidos es la preparación de segmentos de DNA de doble cadena que constituyen genes estructurales o regiones regulatorias. En ocasiones estos segmentos tienen una longitud de varias decenas de pares de bases. Se han utilizado diversos enfoques para la construcción de segmentos de esta naturaleza. (Goddel et al, 1979).

Basándose en los resultados obtenidos al sintetizar el promotor "consenso", se propone una alternativa para construcción de segmentos de doble cadena de más de 40 p.b. Este enfoque se ilustra en la figura 5.

A lo largo de un proceso como el propuesto, resulta de gran utilidad disponer en el vehículo de clonación tanto de un patrón de restricción sencillo donde se sigan con facilidad los cambios de mobilidad de las bandas específicas como consecuencia de la adición de los segmentos sintéticos, como la disponibilidad de una multiplicidad de sitios de restricción únicos que faciliten el diseño de la estrategia.

Figura 5. Estrategia propuesta para la producción de DNA biténico utilizando polinucléotodos sintéticos y DNA polimerasa. El enfoque que se sugiere - involucra la producción de intermediarios de - hasta 250 p.b. flanqueados por dos sitios de - restricción.


sitio 1
40mero
40mero
región de doble cadena
~10 p.b.
DNA polimerasa
dNTPs
sitio 1
70 p.b.
35mero
sitio 2
35mero
DNA polimerasa
dNTPs
60 p.b.
sitio 2
LIGASA
POLIMERO DE AMBOS FRAGMENTOS AL AZAR
ENDONUCLEASAS
1 Y 2
sitio 1
sitio 1
~135 p.b.
sitio 1
sitio 2
~125 p.b.
sitio 2
sitio 2
~115 p.b.
1) PURIFICAR
2) CLONAR EN UN VEHICULO ENTRE SITIOS
DE ENDONUCLEASAS 1 Y 2
ENDONUCLEASAS 1 Y 2
sitio 1
Gene
de resistencia
a antibiótico
sitio 2
sitio 3
sitio 1
fragmento
sitio 2
sitio 3
sitio 4
1) ENDONUCLEASAS
2) PURIFICACION
sitio 2
40mero
40mero
DNA polimerasa
dNTPs
sitio 2
70 p.b.
35mero
sitio 3
DNA polimerasa
dNTPs
60 pb.
sitio 3
LIGASA
POLIMERO AL AZAR
ENDONUCLEASAS
2 Y 3
sitio 2
sitio 2
~135 p.b.
sitio 2
sitio 3
~125 p.b.
sitio 3
sitio 3
~115 p.b.
1) PURIFICAR
2) CLONAR EN UN VEHICULO ENTRE SITIOS
DE ENDONUCLEASAS 2 Y 3
ENDONUCLEASAS 2 Y 3
sitio 1
sitio 2
fragmento
sitio 3
sitio 4
1) ENDONUCLEASAS 2 Y 4
2) PURIFICACION
LIGASA
sitio 1
sitio 2
fragmento
~250 pb
sitio 3
sitio 4

3.1.2) Mutagénesis dirigida a un sitio.

Como se mencionó en la introducción, a la fecha - se han usado con éxito diversos métodos para la obtención de mutantes en sitios específicos de segmentos de DNA. Los procedimientos que han utilizado tanto bacteriofagos (Miyada et al, 1982), como plásmidos (Wallace et al, 1980), como vehículos.

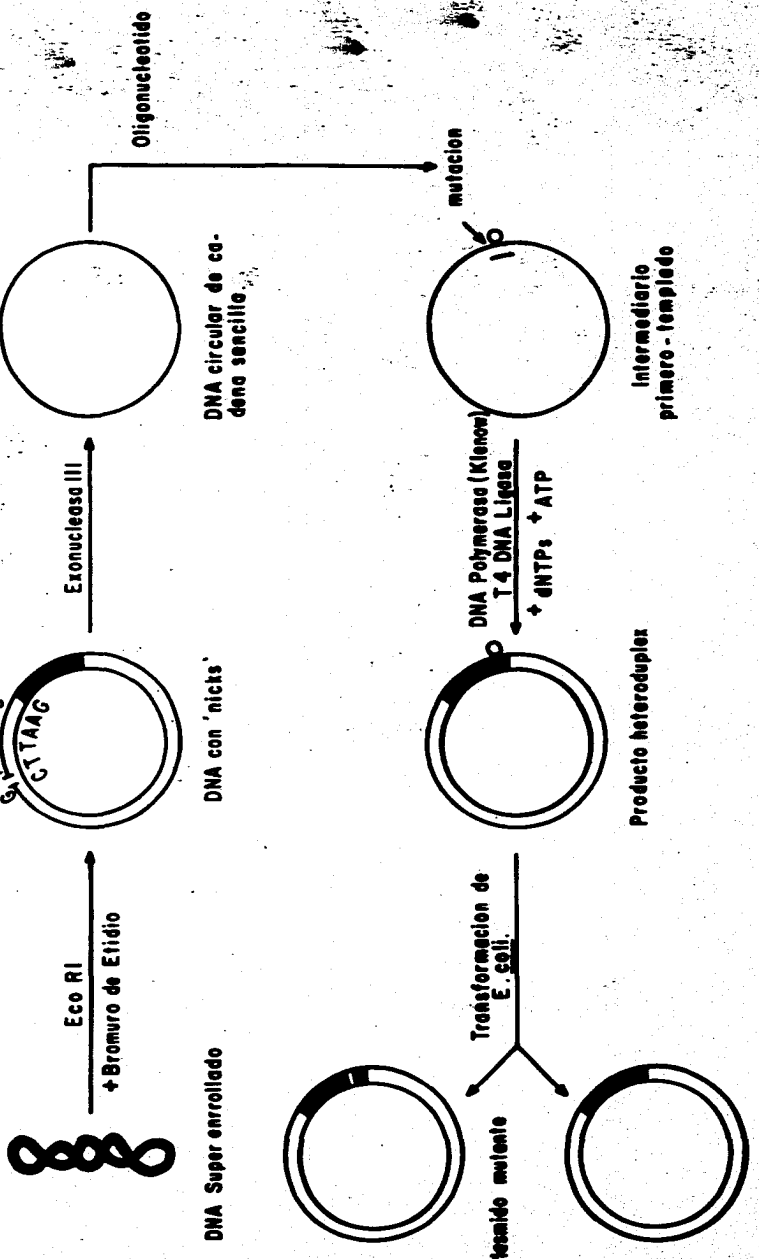
En la figura 6 se describe una estrategia que ha sido utilizada para el caso de mutagénesis dirigida sobre - DNA clonado en plásmido.

La disponibilidad de un vehículo de clonación pequeño puede resultar en un aumento de la eficiencia de la - reacción de reparación con DNA polimerasa 1. Además la probabilidad de encontrar secuencias nucleotídicas con homología parcial al primero que se utiliza, diferentes a la secuencia blanco, disminuye con la longitud del plásmido.

3.1.3) Selección de segmentos de DNA con actividad de promotor.

El plásmido pΔPT1 debe resultar útil como vehículo para la selección de segmentos de DNA con actividad de pro

Figura 6. Estrategia utilizada para la producción de mutaciones sobre plásmidos (Wallace et al., 1980).



motor. La presencia de dos sitios de restricción (EcoRI y Hind III) en posición 5' al gene estructural de resistencia a tetraciclina le confiere versatilidad a este respecto.

Deben hacerse notar, sin embargo, algunas limitaciones que el sistema tiene. El mecanismo de resistencia a tetraciclina involucra proteínas membranales y su cuantificación es poco confiable (Covarrubias et al, 1981), además la actividad del producto de estos genes no es susceptible a cuantificación in vitro. Estas características limitan la utilidad del vehículo pΔPT1 fundamentalmente a la selección de segmentos con actividad de promotor. Posteriormente deberán construírse vehículos diseñados con el propósito expreso de cuantificar la eficiencia de los promotores clonados.

3.2) Perspectivas para el uso del promotor sintético como modelo de estudio.

Existen algunas aplicaciones interesantes para usar el segmento de DNA sintetizado.

Por una parte es atractivo disponer de un promotor muy activo para la optimización de la expresión de genes heterólogos en E. coli. Cabe esperar que una secuencia

consenso o una variante sencilla de él tuviera esta cualidad, particularmente si se toma en consideración la observación - ya mencionada de que la gran mayoría de las mutaciones que - se acercan a la secuencia consenso aumentan la eficiencia del promotor y las que se alejan la disminuyen.*

Por otra parte, el diseño utilizado para la síntesis del promotor permite la obtención de variantes con una cantidad relativamente pequeña de trabajo adicional.

Una alternativa consiste en la utilización de intermediarios en la síntesis de los segmentos constituyentes del promotor. Aprovechando la región 3' de cada uno de los fragmentos se pueden producir cadenas qur varían en sus re-- giones 5' y someterse a reacciones con DNA polimerasa 1 diferentes combinaciones de estos análogos. Para este efecto se guardó el 50% de la resina en el momento en que la síntesis se había llevado a cabo hasta el nivel de 11 nucleótidos para el 27 mero y de 13 para el 25 mero. Porciones de estos intermediarios pueden ser reactivadas en sus extremos 5' y continuarse la síntesis hasta lograr los análogos deseados con variantes en una de las posiciones de la secuencia.

* Resultados preliminares obtenidos después de clonar el promotor en el sitio de HindIIId el plásmido pΔPT1, indican que la fuerza del mismo es similar a la de otros promotores naturales. Este resultado es inesperado, y requerirá de mayor investigación el arrojar algo de luz sobre el fenómeno. En cualquier caso, es posible que alguna variante - sencilla de este promotor sí resulte en un aumento considerable de su eficiencia.

Por otra parte, sobre la secuencia del promotor - ya incorporada a un plásmido se pueden producir mutaciones - utilizando oligonucleótidos sintéticos y después recombinar éstas, unas con otras, aprovechando el sitio de HpaI presente en medio del promotor. Estas posibilidades resultan interesantes como medio para esclarecer la relación estructura - función de promotores procariontes.

REFERENCIAS

Agarwal, A., Yamazaki, P.J., Cashion, P.J. and Khorana, H. G. Angew. Chem. Internest. Edit. 11, 451 (1972).

Agarwal, A., Brunstedt, J. and Noyes, B., J. Biol. Chem. 256: 1023-1028 (1981).

Bolivar, F., Rodríguez, R., Greene, P., Betlach, M., Heyneker, H. Boyer, H., Crosa, J. and Falkow, S.: (1977) Gene 2: 95-113.

Bolivar y Backman 1979 (Metods in enzimology, 68: 245),

Bolivar, F.,: Gene 4: 121-136 (1978).

Broka, C., Hozumi, T., Arentzen, R. and Itakura, K. Nucl. Acids Res. 8, 5461-5471, (1980).

Conner, B.; Takano, T.; Tanaka, S.; Itakura, K. and Dickerson, R. (1981) Nature, in press.

Crea, R., Kraszewski, A., Hirose, T. and Itakura, K., Proc. Natl. Acad. Sci. USA, 75: 5765-5769 (1978).

Collins, J., Curr. To. Microbiol and Imun. 78: 121-165 (1977).

Covarrubias, L., Cervantes, L., Covarrubias, A., Soberón, X., Vichido, I., Blanco, A., Kuperstoch-Portnoy, Y and Bolivar, F., Gene 13: 25-35 (1981).

Dembeck, P. and Itakura, K., JACS, 103: 706 (1981).

Gillam, S. and Smith M. (1979) Gene 8: 81-97.

Goeddel, A., Heyneker, H.; Hozumi, T.; Arentzen, R.; Itakura, K.; Yansura, D.; Ross, M; Hiozzari, G.; Crea, R., and Seeburg, P.H. (1979) Nature 282: 544-548.

Goeddel, A.; Kleid, A.; Bolivar, F.; Heyneker, H.; Yansura, A.; Crea, R.; Hirose, T.; Krasszewski, A.; Hakuta, K.B. and Riggs, A.A. (1979) Proc. Natl. Acad. Sci. USA 76: 106-110.

Goeddel, D.; Yansura, D.G. and Caruthers, M.H. (1978). Proc. Nat. Acad. Sci. USA 75: 3578-3582.

Itakura, K. (1980) TIBS, Mayo 1980. 114-116.

Kawashima, E.H.; Gadec, T. and Caruthers, M.H. (1977) Biochemistry 16: 4209-4217.

Khorana, H.G., Terer, J.G. Moffat, J.G. y Pol., E.H.: Chem. and Ind. 1523 (1956).

Miyada, G.; Soberón, X., Itakura, K. and Wilcox, G. (1982). Gene, in press.

Miyoshi, K.; Huang, T. and Itakura (1980) Nucl. Acids. Res 8: 5491.

Pribnow, D. (1975) Proc. Nat. Acad. Sci. USA 72: 784-788.

Reese, C. Tetrahedron, 34: 3143-3179 (1978).

Rodriguez, R.; West, R.; Heyneker, H.; Bolivar, F. and Boyer H. (1979) Nucl. Acids. Res. 6: 3267-3287.

Rossi, J.; Ross, W.; Egan, J.; Lipnan, A. and Landy, A. (1979) J. Mol. Biol. 128: 21-47.

Schaller, H.; Gray, C. and Herrman, K. (1975) Proc. Natl. Acad. Sci. USA 72: 737-741.

Siebenlist, U.; Simpson, R. and Gilbert, W. (1980) Cell 20: 269-281.

Young, I.G. and Poulis, M.I. Gene, 4, 175-179 (1978).

Wallace, B.; Johnson, P.; Tanaka, S.; Schold, M. Itakura, and Abelson, J. (1980) Science 209: 1396-1400.

Wang, A.; Quigley, G.; Kolpak, F.; Crawford, J.; Van Boom, J.; Vander Marel, G. and Rich. A. (1979) Nature 282: 680-686.

Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K. and Dickerson, R. (1980) Nature, 287: 755-758.

Gene, 9 (1980) 287–305
© Elsevier/North-Holland Biomedical Press

CONSTRUCTION AND CHARACTERIZATION OF NEW CLONING VEHICLES
IV. DELETION DERIVATIVES OF pBR322 AND pBR325

(Cloning vectors; *Bal*I, *Bam*HI, *Eco*RI, *Eco*RII, *Hin*dIII, *Pst*I, *Pvu*I, *Pvu*II and *Sal*I restriction endonucleases; nucleotide sequence; pBR327; pBR328)

X. SOBERON, L. COVARRUBIAS and F. BOLIVAR*

Departamento de Biología Molecular, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, México 20, D.F., Apdo. postal 70228 (México)

(Received September 17th, 1979)
(Revision received and accepted November 20th, 1979)

SUMMARY

In vitro recombinant DNA experiments involving restriction endonuclease fragments derived from the plasmids pBR322 and pBR325 resulted in the construction of two new cloning vehicles. One of these plasmids, designated pBR327, was obtained after an *Eco*RII partial digestion of pBR322. The plasmid pBR327 confers resistance to tetracycline and ampicillin, contains 3273 base pairs (bp) and therefore is 1089 bp smaller than pBR322. The other newly constructed vector, which has been designated pBR328, confers resistance to chloramphenicol as well as the two former antibiotics. This plasmid contains unique *Hin*dIII, *Bam*HI and *Sal*I sites in the tetracycline resistance gene, unique *Pvu*I and *Pst*I sites in the ampicillin resistance gene and unique *Eco*RI, *Pvu*II and *Bal*I sites in the chloramphenicol resistance gene. The pBR328 plasmid contains approx. 4900 bp.

INTRODUCTION

It is now clear that bacterial plasmids can serve as efficient vectors for the molecular cloning of various species of DNA fragments generated with a variety of restriction endonucleases. The *Escherichia coli* plasmid pBR322 (Bolivar et al., 1977a) and its derivative pBR325 (Bolivar, 1978), provide single cleavage sites for the *Hin*dIII, *Bam*HI, *Sal*I, *Pst*I, *Pvu*I and *Eco*RI endonucleases. The insertion of DNA fragments into any of the first three sites inactivates the tetracycline resistance (Tc^r) gene present in these plasmids.

*To whom reprint requests should be addressed.

Abbreviation: bp, base pairs.

The unique *Pst*I and *Pvu*I restriction sites present in these vectors are located in the ampicillin resistance (Ap^r) gene, while the *Eco*RI site in pBR325 is located in the chloramphenicol resistance (Cm^r) gene. Plasmid pBR322 also carries unique cleavage sites for the *Pvu*II and *Bal*I restriction endonucleases, which are located between the origin of replication and the Tc^r gene (Fig. 1). The vector pBR325 carries two *Pvu*II and two *Bal*I sites, of which one of them is located in the Cm^r gene (Alton and Vapnek, 1979).

A derivative of pBR325, designated pBR328, has been constructed in order to gain additional advantages as a cloning vector. The reason for constructing this vehicle was to eliminate the *Pvu*II and *Bal*I sites located between the origin of replication and the Tc^r gene. In this way, fragments of DNA generated by either restriction enzyme can be cloned in their respective unique sites, and cells carrying recombinant DNA molecules can be identified by their Cm^s, Tc^r, Ap^r phenotypes.

For the construction of pBR328, a Tc^r, Ap^r deletion of pBR322 was isolated. This cloning vector, designated pBR327, is exactly 1089 bp smaller than the parental plasmid pBR322.

MATERIALS AND METHODS

(a) Bacterial strains

E. coli K-12 strains RR1 (*pro leu thi lacY rpsL20 hsdR hsdM ara14 galK2 xyl5 mtl1 supE44*) and GM31 (*dcm6 thr1 leu6 thi1 his lacY galK2 galT22 ara14 tonA31 tsx78 supE44*) (obtained from T. Kadish) were used as recipient cells in all transformation experiments.

(b) Media and buffers

Bacterial strains were grown in either Luria Broth (LB) or M9 glucose minimal media. When needed, ampicillin, tetracycline and chloramphenicol were added to LB plates at final concentrations of 25 μg/ml.

(c) pBR322 DNA source

As properly suggested by Sutcliffe (1979), it is important that studies which rely substantially on the pBR322 sequence should be initiated with a pBR322 closely related to the one he originally sequenced.

According to this proposition, pBR322 DNA was obtained from the original *E. coli*[pBR322] lyophilized stab used to prepare the stab sent to J.G. Sutcliffe via W. Gilbert.

(d) Preparation of DNA

DNA from plasmids pBR322 and pBR327 was prepared by amplifying logarithmically growing cultures by the addition of 170 μg/ml of chloramphenicol (Clewell, 1972). Spectinomycin (300 μg/ml) was used for the amplification of cells carrying the Cm^r plasmids, pBR325 and pBR328.

Extraction and purification of plasmid DNA was performed using the cleared

lysate technique previously described by Betlach et al. (1976). *Neurospora crassa* DNA was a gift from R. Palacios.

(e) Enzymes

With the exceptions of *Bal*I, *Pvu*II, *Pvu*I and *Ava*I (P.L. Biochemicals), all other restriction enzymes used in this work were purified according to the procedure of Greene et al. (1978). T4 DNA ligase, a gift from H. Heyneker, was purified according to the procedure described by Panet et al. (1973). S1 nuclease, a gift from P. Seeburg, was used as described elsewhere (Ullrich et al., 1977).

(f) Ligation of DNA

Ligations were carried out in 66 mM Tris · HCl pH 7.6, 6.6 mM $MgCl_2$, 10 mM dithiothreitol and 0.5 mM ATP at 12°C for 6—12 h. The concentration of T4 DNA ligase and of DNA termini were varied to promote polymerization or circularization. When blunt-ended DNA fragments were ligated, the concentration of ends was at least 0.2 μM and approx. 5 units of T4 DNA ligase per ml were added to the reaction mixture (Heyneker et al., 1976). For cohesive-ended DNA fragments, 1 unit of T4 DNA ligase per ml was sufficient and the concentration of ends was adjusted in such a way that linear molecules were favored (Dugaiczyk et al., 1975).

(g) Agarose and acrylamide gel electrophoresis

The conditions for agarose and acrylamide electrophoresis have been described previously (Bolivar et al., 1977b).

(h) Isolation of DNA fragments from agarose gels

DNA fragments were isolated from 1% agarose gels using agarase from Calbiochem (Cat. No. 121811), as described by Finkelstein and Rownd (1978).

(i) Transformation of E. coli *K-12*

E. coli cells of strains RR1 and GM31 were prepared for transformation by the method described by Cohen et al. (1972).

(j) Nucleotide sequence determinations

Nucleotide sequence determinations were performed as described by Maxam and Gilbert (1977).

RESULTS

(a) Construction of pBR327

An in vitro DNA recombinant experiment using pBR322, was designed to delete a DNA fragment which is unessential for DNA replication or antibiotic resistance. As can be seen in Fig. 1, the *Eco*RII DNA fragment located between

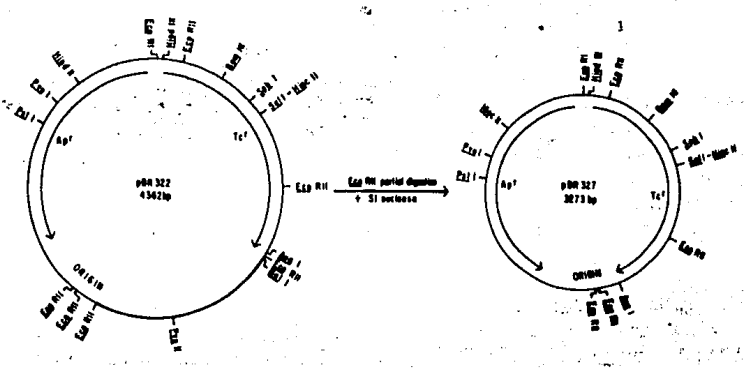


Fig. 1. Diagrammatic representation of the construction of pBR327. Plasmid pBR322 DNA was partially digested with the restriction endonuclease *Eco*RII and the resulting DNA was treated with S1 nuclease to generate blunt-ended DNA fragments. This DNA mixture was treated with T4 DNA ligase and the ligated mixture was used to transform *E. coli* strain GM31.

the origin of replication and the Tc^r gene in pBR322, is such a fragment (Bolivar et al., 1977a). In order to delete this region of pBR322, 20 μg of plasmid DNA were partially digested with *Eco*RII endonuclease. The resulting products were then digested with the S1 nuclease to eliminate the non-complementary *Eco*RII ends generated at positions 1440 (CCTGG) and 2500 (CCAGG) in pBR322 (Sutcliffe, 1979). The digested DNA was run on a 1% agarose gel and DNA fragments with the appropriate molecular weights (approx. 3300 bp) were isolated from the gel after agarase treatment. The DNA was ethanol-precipitated, phenol-extracted and ligated with T4 DNA ligase. The ligated mixture was used to transform *E. coli* strain GM31. The Ap^r, Tc^r transformants obtained were screened for plasmid DNA which gave a linear DNA molecule, of approx. 3300 bp upon digestion with *Eco*RI endonuclease.

(b) Restriction endonuclease characterization of pBR327

The restriction map of one of the resulting plasmids, pBR327, was determined using various restriction endonucleases and the known restriction map of pBR322 as a reference (Bolivar et al., 1977a; Sutcliffe, 1979). The plasmid pBR327 was found to carry unique substrate sites for the *Eco*RI, *Hin*dIII, *Bam*HI, *Sal*I, *Pst*I and *Ava*I restriction endonucleases. Double and triple digestions of the plasmid DNA (data not shown) showed that the relative positions of these sites were identical to those mapped in pBR322. However, no *Bal*I or *Pvu*II sites were present in pBR327, which would be expected for a plasmid lacking the 1060 bp *Eco*RII fragment.

*Eco*RII endonuclease was used to further characterize pBR327. As shown

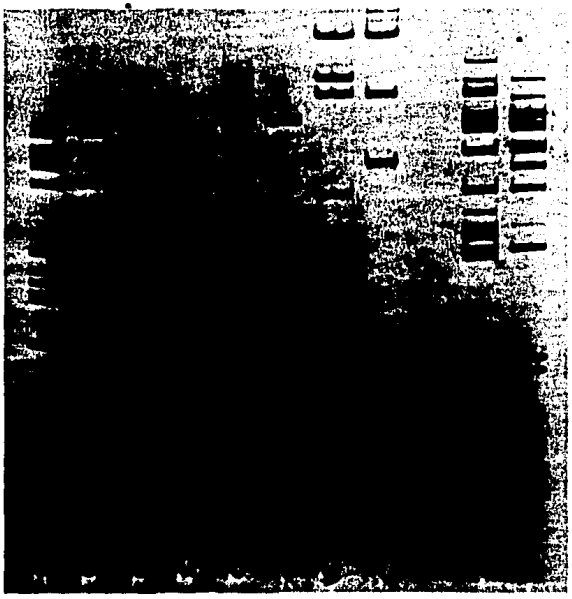


Fig. 2. Acrylamide slab gel electrophoresis of pBR322 and pBR327 plasmid DNA cleaved with various restriction endonucleases. Slots a and b show the *Hae*III pattern of plasmid pBR322 and pBR327, respectively. Notice that bands 2, 3, 10 and 18 (the last one not shown in the gel) from pBR322 are not present in the pBR327 pattern. Instead, a new band of 146 bp is present in pBR327. Slots c and d show the *Hha*I pattern of pBR322 and pBR327 respectively. Bands 2, 5, 8, 11, 14, 17, 24, 25 and 26 disappeared from pBR322 and a new band, running approximately in the same positions as bands 10 in pBR322, has appeared in pBR327. The *Hpa*II pattern of pBR322 and pBR327 are shown in slots e and f, respectively. Notice that six bands disappeared in the pBR327 pattern when compared with pBR322. A new band, running in the same position as band 4 in pBR322, is present in pBR327. Slots h and g show the *Eco*RII pattern of pBR327 and pBR322 respectively. Bands 2, 4 and 5 (the band above band 5 is a partial), present in pBR322 disappeared in pBR327. A new band of 475 bp, the result of the fusion of bands 4 and 5, is present in the new vector. The *Alu*I restriction patterns of pBR322 and pBR327 are shown in slots i and j, respectively.

in Fig. 2 (slot g), pBR322 has six *Eco*RII sites which yield fragments of 1857, 1060, 928, 383, 121 and 13 bp after digestion (Sutcliffe, 1979). By contrast, pBR327 has only 4 *Eco*RII sites (Fig. 2, slot h). As expected, the *Eco*RII fragment No. 2 (1060 bp) of pBR322 was missing in pBR327, and fragments Nos. 4 and 5 were fused in a new *Eco*RII segment (Figs. 3 and 4). The *Hpa*II fragments Nos. 2, 4, 8, 10, 12A and 18B from pBR322 were also found to be missing in pBR327. However, a new *Hpa*II fragment was observed which resulted from a fusion between the original *Hpa*II fragments No. 2 and No. 8

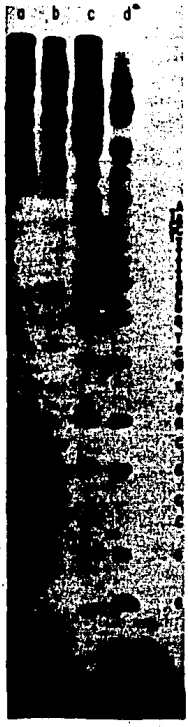
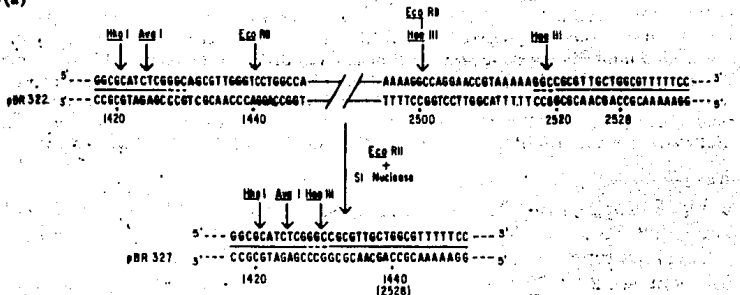


Fig. 3. The nucleotide sequence at the region where the deletion occurred during the construction of plasmid pBR327. (a) Plasmid pBR327 was obtained from pBR322 after a partial *Eco*RII digestion followed by an S1 nuclease treatment. To determine the nucleotide sequence at the junction in pBR327, DNA from this plasmid was digested with *Ava*I endonuclease. The linearised DNA was labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase and finally digested with *Hin*fI endonuclease. The resulting nucleotide sequence from the *Ava*I site (in the clockwise direction) into the junction was determined and is shown in the figure. (b) The continuous underlined regions in the pBR322 sequence were conserved in pBR327. The sequence GGC from positions 1428 to 1430, in pBR327, could arise from either side (as shown by the discontinuous lines) on the junction. The sequence shows that pBR327 is 1089 bp smaller than pBR322.

(a)



(b)

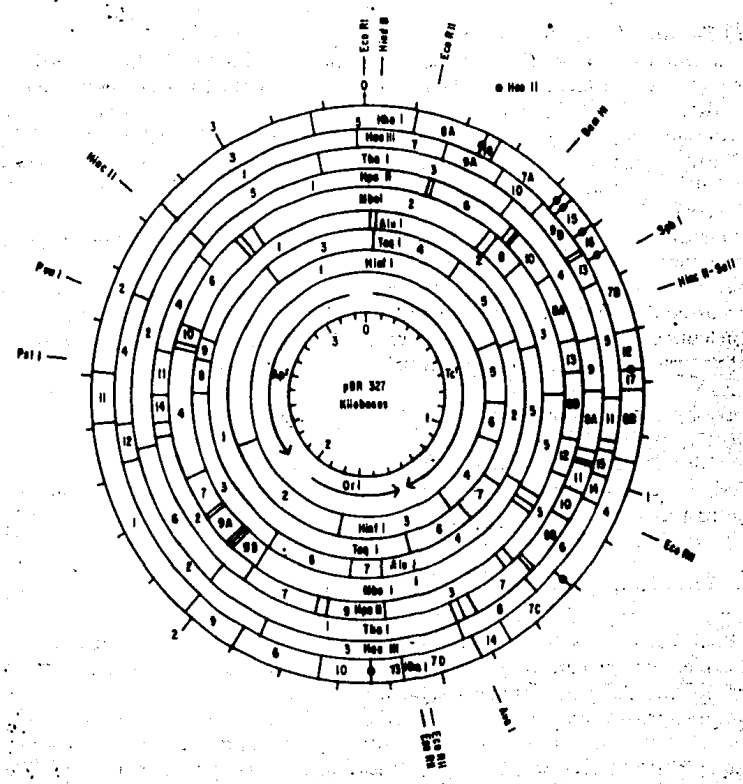


Fig. 4. Restriction map of pBR327 DNA. Plasmid pBR327 fragment map was determined from the primary sequence data from pBR322 (Sutcliffe, 1979), and the nucleotide sequence at the junction where the deletion occurred during the construction of pBR327 (see Fig. 3). The total length is 3273 nucleotide pairs. Position of "0" (= zero), has been arbitrarily chosen at the middle of the single *Eco*RI cleavage site. The figure shows the restriction cleavage maps for the endonucleases *Hae*III, *Alu*I, *Hpa*II, *Hha*I, *Hin*fI, *Tha*I, *Mbo*I, which are indicated in the concentric rings. The fragments are numbered by sizes and those sizes are listed in Table I. The precise coordinates of the restriction cuts are listed in Table II. Around the perimeter of the circular map appear the sites for enzymes which cut infrequently including those with unique restriction sites.

from pBR322 (Fig. 2, slots e and f). Similarly, *Hae*III segments Nos. 2, 3, 10 and 19 from pBR322, were not observed in pBR327, but a new *Hae*III fragment appeared (Fig. 2 slots a and b). Restriction maps for several endonucleases have been determined for pBR327 using this approach and all of them are in agreement with a deletion of approx. 1090 bp.

(c) The nucleotide sequence of pBR327

In order to completely characterize this new cloning vehicle, the nucleotide sequence was determined for the DNA region where the S1 nuclease-treated *Eco*RII sites of pBR322 were fused.

Plasmid pBR327 was digested with *Ava*I restriction endonuclease, the resulting fragment was labeled with [γ-^{32}P] ATP and T4 polynucleotide kinase and redigested with *Hin*fI. Figs. 3a and 3b show the nucleotide sequence from the *Ava*I site in the clockwise direction. From these data, and the reported nucleotide sequence for pBR322 (Sutcliffe, 1979), it can be concluded that the deletion that generated pBR327 comprises exactly 1089 bp. Since the restriction site analysis of pBR327 DNA indicates that no other changes occurred during the construction of this plasmid, we can assume that the nucleotide sequence of this cloning vehicle is completely known. On the basis of this assumption, we derived a restriction map of pBR327 that is shown in Fig. 4.

(d) Construction of pBR328

Using pBR327 and pBR325, an in vitro recombinant DNA experiment was designed to construct a cloning vehicle with unique *Pvu*II and *Bal*I sites located in the Cm^r gene. The construction of this plasmid was accomplished by the digestion of pBR325 DNA with *Pst*I and *Bam*HI restriction endonucleases which resulted in two pieces of DNA. As shown in Fig. 5, the smaller DNA of the two fragments carries the Cm^r gene (Bolivar, 1978) and part of the Ap^r and Tc^r genes. This fragment was isolated from a 1% agarose gel. The plasmid pBR327 DNA was digested with the same two endonucleases and the larger fragment, which carries the replication DNA region and the remaining portion of Ap^r and Tc^r genes, was also isolated from agarose.

The purified DNAs were mixed, ligated in vitro, and transformed into *E. coli* RRI. Transformants were selected for Ap^r, Tc^r and Cm^r. This transformation yielded several Ap^r, Tc^r, Cm^r clones which carried plasmids (e.g. pBR328) with the expected molecular weight of approx. $3.1 \cdot 10^6$.

(e) The restriction map of pBR328

As determined from gel electrophoresis patterns of DNA digests, pBR328 was found to carry unique substrate sites for the *Eco*RI, *Bam*HI, *Hin*dIII, *Sal*I, *Pst*I, *Pvu*I, *Pvu*II, *Bal*I and *Ava*I. Double and triple digestions of plasmid DNA (Fig. 6) enable us to localize the relative positions of these sites on pBR328. From these data, and other restriction endonuclease patterns, we constructed a restriction map for several endonucleases that is shown in Fig. 7.

(f) Cloning properties of pBR327 and pBR328

(i) Cloning in Tc^r gene. Since the restriction sites *Hin*dIII, *Bam*HI, and *Sal*I are present in the same relative positions in pBR327 and pBR328 as they are in pBR322, it was assumed that they were also in the Tc^r gene. To verify this assumption, *N. crassa* DNA fragments generated by *Hin*dIII, *Bam*HI or *Sal*I

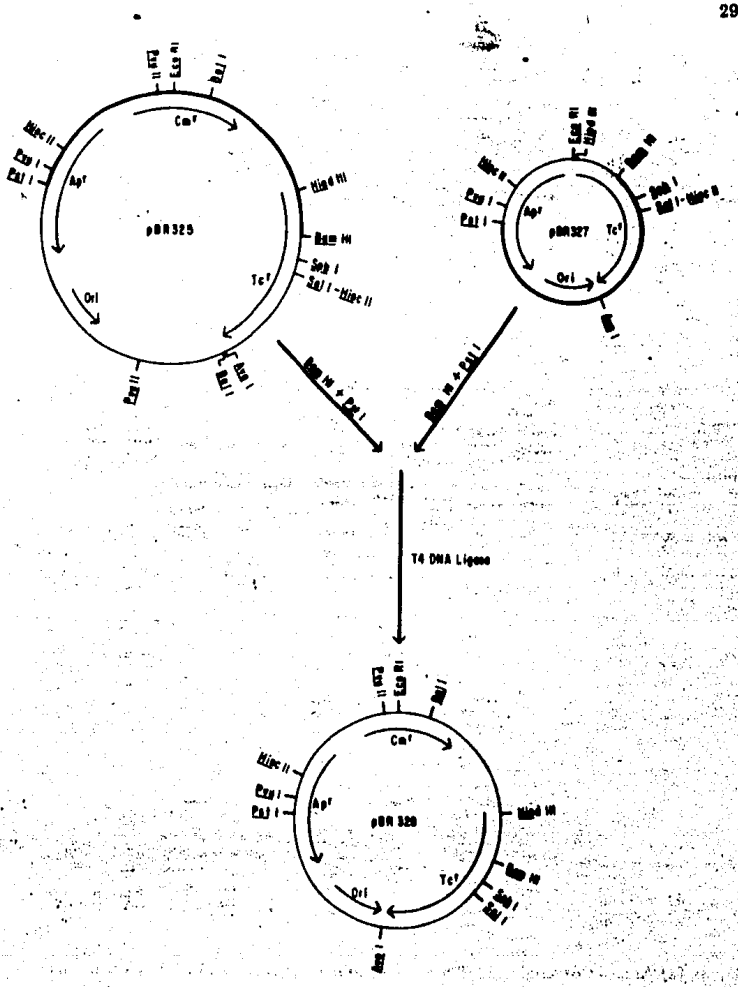


Fig. 5. Diagrammatic representation of the construction of pBR328. Plasmid pBR325 and pBR327 were separately digested with the restriction endonucleases *Pst*I and *Bam*HI. The generated fragments were isolated from a 1% agarose gel. The heavy-lined regions represent the two DNA fragments that were mixed and ligated to each other to generate pBR328.

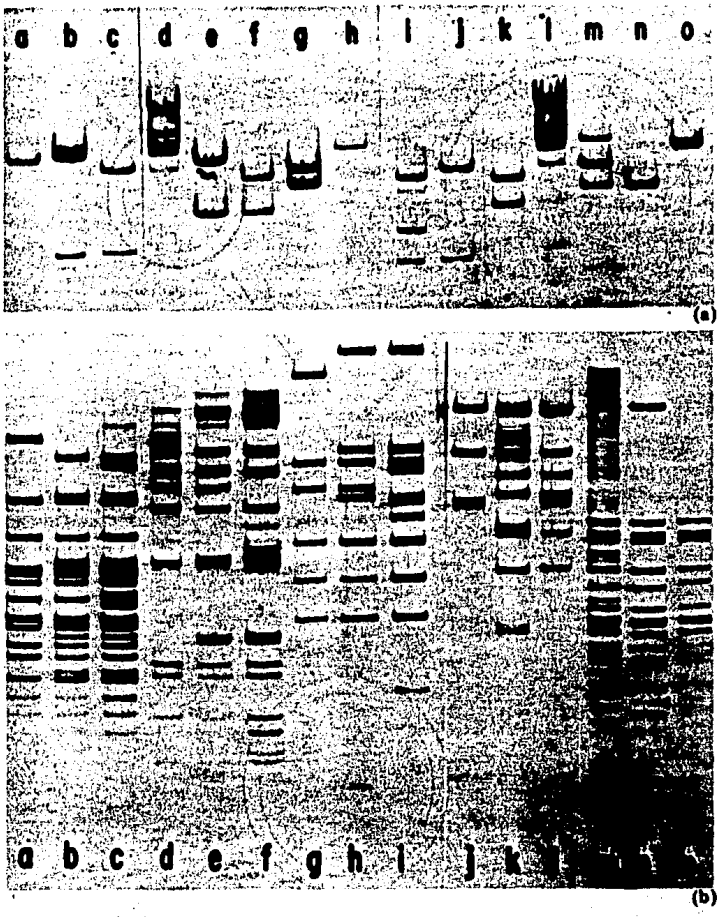


Fig. 6. (a) Analysis of single and double digestions of pBR328 and pBR325 using agarose gel electrophoresis. Molecular weight estimates were based on the six λ fragments generated by *Eco*RI (slots d and l). Slot a shows the double digestion *Eco*RI - *Bal*I of pBR328; two fragments are generated, the smallest is not shown in the picture. Slots b and c show the double digestions *Eco*RI - *Pst*I of pBR325 and pBR328, respectively. The *Sal*I - *Eco*RI double digestions of pBR325 and pBR328 are shown in slots e and f respectively. Slots g and h show the *Pvu*II digestion of pBR325 and pBR328, respectively. The double digestions *Pst*I - *Pvu*II of plasmids pBR325 and pBR328 are shown in i and j, respectively. Slot k shows the *Pvu*II - *Sal*I double digestion of pBR328. Slots m and n show respectively the *Hin*dIII - *Pst*I double digestion of pBR325 (partial) and pBR328 in which the two generated bands

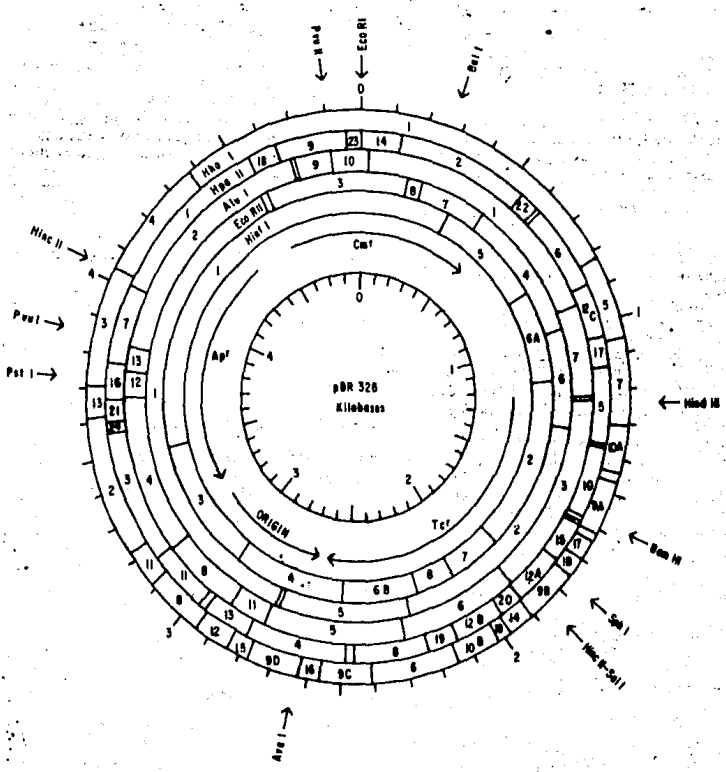


Fig. 7. Restriction map of pBR328. The restriction map of pBR328 was determined from the primary sequence data from pBR327 (positions 1200 to 4480), nucleotide sequence determinations around the *Eco*RI site (positions 4800 to 200 in the clockwise direction) and restriction endonuclease partial digestions (positions 4480 to 1200 in the clockwise direction).

(legend Fig. 6 continued)
run at the same position. Slot o shows the linear pBR328 DNA molecule generated after *Bal*I digestion. (b) Acrylamide gel electrophoresis of plasmid pBR328, pBR327 and pBR325 DNA fragments obtained by digestion with *Hpa*II, *Alu*I, *Hin*fI, *Eco*RII and *Hha*I endonucleases. The restriction enzyme digestions were as follows: Slots a, b and c show the *Hpa*II digestion of pBR327, pBR328 and pBR325, respectively. Slots d, e and f show the *Alu*I digestions of the three plasmid vectors in the same relative order. The *Hin*fI digestions of pBR327, pBR328 and pBR325 are shown in slots g, h and i, respectively. The *Eco*RII digestions of pBR327, pBR325 and pBR328 appear in slots j, k and l, respectively. Slot m shows a partial *Hha*I digestion of pBR325 and slots n and o show the *Hha*I endonuclease pattern of pBR328 and pBR327, respectively.

TABLE I

THE SIZES OF THE RESTRICTION ENDONUCLEASE GENERATED DNA FRAGMENTS FROM pBR322 AND pBR327

Notice that for each restriction enzyme those deleted fragments, originally present in pBR322 (framed, Sutcliffe, 1978; 1979), are not present in pBR327. However, a new fragment is present in pBR327 (underlined) that is not in pBR322. In the case of the *Hpa*II endonuclease, a fragment of 309 bp disappeared and a new one of exactly the same size appeared

*Hae*III Size		*Hpa*II Size		*Alu*I Size	
pBR322 fragment No.	pBR327 fragment No.	pBR322 fragment No.	pBR327 fragment No.	pBR322 fragment No.	pBR327 fragment No.
1 587	1	1 622	1	1 910	
2 540		2 527		2 680	1
3 504		3 404	2	3 655	2
4 458	2	4 309	3	4 521	3
5 434	3	5 242	4	444	4
6 351	4	6 238	5	5 403	5
7 234	5	7 217	6	6 351	
8 213	6	8 201		7 257	6
9 192	7	9 190	7	8 [illegible]	
10 184		10 180		9 126	7
188	8	11A 160	8A	10 100	8
11A 124	9A	11B 160	8B	11 83	9
11B 123	9B	12A 147		12 [illegible]	
12 104	10	12B 147	9	13 [illegible]	
13 89	11	13 122	10	14 [illegible]	
14 80	12	14 110	11	15 15	10
15 64	13	15 90	12	16 11	
16 57	14	16 76	13		
17 51	15	17 67	14		
18 21	16	18A 34	15		
19 16		18B 24			
20 11		19A 26	16A		
21 7	17	19B 26	16B		
		20 15	17		
		21A 9	18A		
		21B 9	18B		

*Hin*fI Size		TaqI Size	
pBR322 fragment No.	pBR327 fragment No.	pBR322 fragment No.	pBR327 fragment No.
1 1631	1	1 1444	1
2 517	2	2 1307	
3 506		3 475	2
452	3	4 368	3
4 396		5 315	4
5 344		6 312	5
6 298	4	218	6
7A 221	5	7 141	7
7B 220			
8 154	6		
9 75			

*Tha*I Size	
pBR322 fragment No.	pBR327 fragment No.
1 581	1
2 403	2
3 452	3
4 372	
5 285	4
6 [illegible]	
7 252	5
8 220	6
9 145	7
10A 120	8A
10B 120	8B
11 [illegible]	
12 110	9
13 [illegible]	
14 [illegible]	
15 [illegible]	
16 66	10
17 61	11
18 27	12
19 26	13
17	14
20 10	15
21 6	16
22 2	

TABLE I (continued)

*Hha*I Size			*Hae*II Size			*Mbo*I Size		
pBR322 fragment No.		pBR327 fragment No.	pBR322 fragment No.		pBR327 fragment No.	pBR322 fragment No.		pBR327 fragment No.
1	393	1	1	1676	1	1	1374	
2	347		2	622			900	1
3	337	2	3	429		2	665	2
4	332	3	4	430	2A	3	358	3
5	270			425	2B	4	341	4
6	259	4	5	370		5	317	
7	206	5	6	227	3	6	272	5
8	190		7	181	4	7	258	6
9	174	6	8	89		8	207	
10A	153	7A	9	60	5	9	105	7
10B	152	7B	10	53	6	10	91	8
10C	151	7C	11	21	7	11	78	9
	145	7D				12	75	10
11	121					13	46	11
12A	132	8A				14	36	12
12B	131	8B				15	31	13
13	109	9				16	27	14
14	104					17	18	15
15	100	10				18	17	16
16	92	11				19	15	17
17	83					20	12	18
18	75	12				21	11	19
19	67	13				22	8	20
20	62	14						
21	60	15						
22	53	16						
23	40	17						
24	36							
25	34							
26	30							
27	26	18						
28	21	19						

endonucleases were cloned into their respective sites in the new vectors. The isolated transformants carrying recombinant plasmids were Tc^s. By restriction endonuclease digestion, it was shown that all recombinant plasmids contained DNA fragments of varying sizes inserted into the *Hin*dIII, *Bam*HI or *Sal*I sites (data not shown).

(ii) Cloning in the Ap^r gene. As in the previous case, it was assumed that the unique *Pvu*I and *Pst*I endonuclease recognition sites were associated with the Ap^r gene carried by these plasmids. Cloning DNA into the *Pst*I or *Pvu*I sites of these vectors inactivated the Ap^r gene, and as above all recombinant plasmids contained DNA fragments of varying sizes inserted in these two sites (data not shown).

(iii) Cloning into the Cm^r gene of pBR328. *Eco*RI generated DNA fragments from *N. crassa* were ligated to linear pBR328 previously digested with the same endonuclease. Similarly, *Pvu*II and *Bal*I endonucleases generated *N. crassa*

DNA fragments were also ligated to *Pvu*II or *Bal*I linearized pBR328 DNA. In all cases, upon transformation, the resulting recombinant plasmids gave Ap^r, Tc^r, Cm^s phenotypic transformants (data not shown).

DISCUSSION

In our effort to construct better plasmid cloning vectors we have sought to eliminate unessential genetic information from pBR322 (Bolivar et al., 1977a–c). To this end, we have constructed a cloning vector, pBR327, which has been reduced in size precisely by 1089 bp, but retains all the cloning properties of pBR322. Being smaller than pBR322, and thus having a simpler restriction pattern for various endonucleases, we believe that will be a more useful vehicle in terms of the restriction site analysis of cloned DNA fragments.

Nucleotide sequence analysis enabled us to determine the base pair sequence in the region where the deletion occurred during the construction of pBR327. Since it appears that no other changes in the DNA occurred during the construction of this vehicle, and due to the fact that the pBR322 DNA source that we used for the construction of pBR327 is closely related to that used by Sutcliffe for the pBR322 original base pair sequencing, we assume that the nucleotide sequence of pBR327 is identical to that of pBR322 (Sutcliffe, 1979), less the 1089 bp deletion.

Using the total nucleotide sequences of pBR322 and pBR327 we have derived a complete restriction map of pBR327. The restriction cleavage sites have been determined either experimentally or from the nucleotide sequence of pBR322. The exact sizes of all the restriction fragments and their relative positions in the pBR327 genome as well as the complete nucleotide sequence, have been presented (Tables I and II, Fig. 8) to facilitate the use of this vehicle in future molecular cloning experiments.

Moreover, because pBR327 is a deletion product of pBR322, this plasmid has been submitted to the U.S. National Institutes of Health for certification as the vector component of an EK2 host vector recombinant DNA system.

Furthermore, it has been described (Bolivar et al., 1977a) that pBR322 can not be mobilized by R64drd11; however, Young and Poulis (1978) have reported that this vehicle can be mobilized by R64drd11 when a third plasmid, ColK, is present in the same cell. These authors proposed that ColK can provide a transacting product that allows the mobilization of pBR322. We would like to propose that such a transacting product, that should interact with a specific pBR322 DNA region, could be a "mobility-like" protein(s) (relaxation enzyme) such as the one described in ColE1 sought to be involved in the mobilization of this plasmid by conjugative elements (Warren et al., 1978, 1979). We know that pBR322 does not carry the gene, originally present in the parental plasmid pMB1, that codes for a similar "mobility" protein (Betlach et al., 1976; Bolivar et al., 1977a; Sutcliffe, 1979). However, there is still a DNA region in pBR322 that comprises a nucleotide sequence that could be the substrate site, (relaxation site or "*bom*" site, Warren et al., 1978) for

TABLE II

THE CLEAVAGE COORDINATES FOR ENZYMES WHICH CUT pBR327

The numbers correspond to the base where to the right of it the restriction enzyme nicks that strand which is 5′ to 3′ in the clockwise direction

*Hae*III	*Hpa*II	*Alu*I	*Hin*fI	*Taq*I	*Tha*I	*Hha*I	*Hae*II	*Mbo*I
174	161	16	631	23	347	103	235	348
297	170	31	851	338	702	235	416	375
401	387	686	1005	650	817	263	437	466
525	402	1089	1303	1125	946	416	497	824
532	411	1553	1756	1266	973	437	550	1096
596	533	1689	2272	2928	978	497	777	1127
830	693	1946			1039	550	1207	1142
919	769	2467			1105	702	1632	1951
940	929	2567			1234	777		2026
991	1019	2630			1244	817		2037
1048	1257				1389	948		2045
1261	1283				1415	1207		2123
1429	1592	*Eco*RI		*Sal*I	1432	1358	*Eco*RII	2135
1863	1739	3271=(−2)		650	2013	1420	129	2240
2321	1765				2343	1565	1057	2581
2401	1955	*Hin*dIII		*Ava*I	2835	1632	1532	2599
2668	2359	29		1424	3168	1732	1545	2645
3255	2393					1906		2903
	2460					2015		2920
	2570					2408		2956
	2812					2501		
*Sph*I		*Bam*HI		*Pst*I		2838	*Pvu*I	
565		375		2523		3170	2726	

this type of enzyme. This DNA region, localized between positions 2207 and 2263 in the pBR322 DNA sequence

5′-GCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGG
CGCGTCGGTACTGGGTCAGTGCATCGCTATCGCCTCACATATGACCGAATTGATACGCC

is very similar to a nucleotide sequence described by Bastia (1978), in which the relaxation site of ColE1 has been located. Therefore, if this assumption is correct we propose that the plasmids pBR327 and pBR328 should not be mobilized by R64drd11, in the presence of ColK, because they lack this specific DNA region. In order to verify this assumption we are performing those mobilization experiments that will confirm it or not. If the proposition is proven to be correct, we believe that pBR327, within *E. coli* strain χ1776, constitutes an important improvement in safety compared to the EK2 pBR322-χ1776 system.

Using plasmid pBR327, we were able to isolate a pBR325 like plasmid, which we have designated pBR328. This vector has two advantages over pBR325. The first is its size, which is 1089 bp smaller than pBR325 and the

Fig. 8. The complete nucleotide sequence of pBR327. Since the restriction site analysis of pBR327 indicates that except for the deletion, no other changes in DNA occurred during the construction of this vector from pBR322, and because the pBR322 DNA source that was used is closely related to that used by Sutcliffe for the pBR322 original nucleotide sequence determination, we assume that the nucleotide sequence of pBR327 is identical to that of pBR322 (Sutcliffe, 1979) less the 1089 bp deletion. The sequence is presented in double-stranded form. The top strand is 5′ → 3′. The sequence is numbered such that 0 (= zero bp) is the middle of the unique *Eco*RI site and the count increases in the clockwise direction first through the Tc^r gene, then the replication region and finally through the Ap^r gene.

second is the fact that it allows for molecular cloning, by insertional inactivation (Timmis et al., 1974) at the *Pvu*II and *Bal*I restriction cleavage sites of the Cm^r gene.

The use of these plasmid vectors now allows for the direct cloning, easy recovery and selection by insertional inactivation of recombinant DNA molecules generated by *Pst*I, *Pvu*I, *Pvu*II, *Bam*HI, *Sal*I, *Hin*dIII, *Eco*RI, *Bal*I, and *Sph*I (Fuchs et al., 1980) restriction endonucleases. Experiments are now being performed in order to apply for the approval of pBR328 in *E. coli* strain χ1776 as an EK2 system.

ACKNOWLEDGEMENTS

We wish to thank Fernando Bastarrachea, Arthur Riggs and specially Ray

Rodriguez for critically reading the manuscript and helpful discussions. We are also grateful to Rosa Maria Cuevas for the typing of this manuscript.

This research was supported by Grant No. 1649 from Proyecto Nacional de Ciencias Básicas (PNCB) del Consejo Nacional de Ciencia y Tecnología, México.

REFERENCES

Alton, K.N. and Vapnek, D., Nucleotide sequence analysis of the chloramphenicol resistance transposon Tn9, Nature, (1979) in press.

Bastia, D., Determination of restriction sites and the nucleotide sequence surrounding the relaxation site of ColE1, J. Mol. Biol., 124 (1978) 601—639.

Betlach, M.C., Hershfield, V., Chow, L., Brown, W., Goodman, H.M. and Boyer, H.W., A restriction endonuclease analysis of the bacterial plasmid controlling the *Eco*RI restriction and modification of DNA, Fed. Proc., 35 (1976) 2037—2043.

Bolivar, F., Rodriguez, R., Betlach, M. and Boyer, H.W., Construction and characterization of new cloning vehicles, I. Ampicillin resistant derivatives of the plasmid pMB9, Gene, 2 (1977b) 75—93.

Bolivar, F., Rodriguez, R., Greene, P.J., Betlach, M., Heyneker, H.L., Boyer, H.W., Crosa, J. and Falkow, S., II. Construction and characterization of new cloning vehicles, A multipurpose cloning system, Gene, 2 (1977a) 95—113.

Bolivar, F., Betlach, M., Heyneker, H.L., Shine, L., Rodriguez, R. and Boyer, H.W., Origin of replication of pBR345 plasmid DNA, Proc. Natl. Acad. Sci. USA, 74 (1977c) 5265—5269.

Bolivar, F., Construction and characterization of new cloning vehicles, III. Derivatives of plasmid pBR322 carrying unique *Eco*RI sites for selection of *Eco*RI generated recombinant molecules, Gene, 4 (1978) 121—136.

Clewell, D.B., Nature of ColE1 plasmid replication in the presence of chloramphenicol, J. Bacteriol., 110 (1972) 667—676.

Cohen, S.N., Chang, A.C.Y. and Hsu, C.L., Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *Escherichia coli* by R factor DNA, Proc. Natl. Acad. Sci. USA, 69 (1972) 2110—2114.

Dugaiczyk, A., Boyer, H.W. and Goodman, H.M., Ligation of *Eco*RI endonuclease generated DNA fragments into linear and circular structures, J. Mol. Biol., 96 (1975) 171—195.

Finkelstein, M. and Rownd, R.H., A rapid method for extracting DNA from agarose gels, Plasmid, 1 (1978) 557—562.

Fuchs, L.Y., Covarrubias, L., Escalante, L., Sanchez, S. and Bolivar, F., Characterization of a site-specific restriction endonuclease *Sph*I from *Streptomyces phaeochromogenes*, Gene, 10 (1980) in press.

Greene, P.J., Heyneker, H.L., Betlach, M., Bolivar, F., Rodriguez, R., Covarrubias, A., Fodor, I., Backman, K. and Boyer, H.W., General method for restriction endonuclease purification, Nucl. Acids Res., 5 (1978) 2773—2780.

Heyneker, H.L., Shine, J., Goodman, H.M., Boyer, H.W., Rosenberg, J., Dickerson, R.E., Narang, S.A., Itakura, K., Liu, S. and Riggs, A.D., Synthetic *lac* operator is functional in vivo, Nature, 263 (1976) 748—752.

Maxam, A. and Gilbert, W., A new method for sequencing DNA, Proc. Natl. Acad. Sci. USA, 74 (1977) 560—564.
Panet, A., Van de Sande, J.H., Loewen, P.C., Khorana, H.G., Raae, A.J., Lillehang, J.R. and Kleppe, K., Physical characterization and simultaneous purification of bacteriophage T4 induced kinase, polynucleotide ligase and deoxyribonucleic acid polymerase, Biochemistry, 12 (1973) 5045—5050.
Sutcliffe, J.G., pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long, Nucl. Acids Res., 5 (1978) 2721—2728.
Sutcliffe, J.G., Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322, Cold Spring Harbor Symp. Quant. Biol., 43 (1979) pp. 77—90.
Timmis, K., Cabello, F. and Cohen, S.N., Utilization of two distinct modes of replication by a hybrid plasmid constructed in vitro from separate replicons, Proc. Natl. Acad. Sci. USA, 71 (1974) 4556—4560.
Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W. and Goodman, H., Rat insulin gene: construction of plasmids containing the coding sequences, Science, 196 (1977) 1313—1321.
Warren, G.J., Twigg, A. and Sherratt, D.J., ColE1 plasmid mobility and relaxation complex, Nature, 274 (1978) 259—261.
Warren, G.J., Saul, M.W. and Sherratt, D.J., ColE1 plasmid mobility: Essential and conditional functions., Mol. Gen. Genet., 170 (1979) 103—107.
Young, I.G. and Poulis, M.I., Conjugal transfer of cloning vectors derived from ColE1, Gene, 4 (1978) 175—179.

Communicated by W. Szybalski.

NOTE ADDED IN PROOF

Since we submitted the manuscript, the plasmid pBR327 has been certified by the NIH as an EK2 vector, on December 16, 1979.

SOLID-PHASE SYNTHESIS OF POLYNUCLEOTIDES. VII. Synthesis of a "Putative Promoter" DNA Fragment using Chemically Synthesized Polynucleotides and DNA Polymerase I.

Xavier Soberón, John J. Rossi, Garrett Larson and Keiichi Itakura

City of Hope Research Institute, Division of Biology, Molecular Genetics Section, Duarte, California 91010 USA.

ABSTRACT

A 42-base pair deoxyribonucleic acid fragment, comprising an *E. coli* "putative promoter" was produced by a combination of chemical and enzymatic reactions. Two single-stranded polynucleotides, 27 and 25 bases in length, respectively, were chemically synthesized by the phosphotriester method using di, tri and tetranucleotide blocks as coupling units on a solid-phase polyacrylmorpholide support. The sequences of the fragments were designed to hybridize with each other through their 3' ends, forming a 10-base pair stretch. A reaction with DNA polymerase I, using this hybrid as substrate, generated a 42-base pair product. These experiments constitute a novel approach that is rapid and convenient for the synthesis of double-stranded DNA fragments.

INTRODUCTION

In previous papers we have reported recent advances for the chemical synthesis of polydeoxyribonucleotides (1,2). These methods have greatly simplified the preparation of DNA fragments of defined sequences. In many applications of this methodology there has been a requirement for double-stranded DNA molecules. For example, the synthesis of whole structural genes (3,4) and "connector" sequences (5) for the bacterial expression of mammalian hormones provide useful alternatives for the production of medically important polypeptides. Also, some biological properties of regulatory sequences have already been investigated with the aid of synthetic double-stranded DNA segments (6-8). In this paper we wish to report experiments that show a simple and accurate way to produce these types of molecules.

A previous approach for making double-stranded DNA segments utilized the ligation of single-stranded fragments aligned by complementary overlaps (9). This approach was used when the fragments were twenty nucleotides long or less. We propose a

new method to prepare double-stranded DNA fragments comprised of longer polynucleotides, now available through our improved synthetic approach. Using the recently developed techniques for the solid-phase synthesis of polydeoxyribonucleotides with an acrylmorpholide support and the block-coupling approach, we synthesized a heptaicosamer (27-mer) and a pentaicosamer (25-mer) DNA fragment. The sequences of the fragments were designed to form a ten-base pair (b.p.) complementary stretch through their 3' ends. When these fragments were hybridized to each other and used as substrates for a DNA polymerase reaction, a 42 b.p. double-stranded DNA molecule was obtained.

MATERIALS AND METHODS

Except for the preparation of dimers (see below), the synthesis of the blocks (trimers and tetramers) and the oligonucleotides was performed according to the published procedures (1,2). Polyacrylmorpholide resin (500 mg), with the first nucleoside T (0.13 mmole/gram), was used for the 27-mer and with the nucleoside C (0.14 mmole/gram) for the 25-mer.

Synthesis of Dinucleotides

Some of the dinucleotides used in this work were synthesized by an alternative route to the one previously reported (1). The procedure is schematically presented in Figure 1. 5'-Dimethoxytrityl deoxynucleosides (10 mmole) were phosphorylated as described (1) to produce the monotriazolide derivative II. The dioxane solution of monotriazolide was then added to 1.5 equivalents of an amino-protected 2'-deoxynucleoside V, previously dried by repeated co-evaporation with pyridine. The reaction was performed in the presence of 5 equivalents of 1-methylimidazole as nucleophilic catalyst. Under these conditions, the 5'-dimethoxytrityl dinucleotide III was obtained as the main product. The amount of by-products is variable, depending upon the sequence of the dimer prepared. The product was purified by column chromatography on silica gel (10 g/mmole) using a stepwise gradient of methanol (0-5%) in chloroform. The procedure described was used only in those cases in which the difference between the main and the by-products (such as the

3'-3' dinucleotide, IV) was large enough to make the chromatographic separation feasible and with a reasonable recovery.

Dinucleotides were phosphorylated in the same manner as the monomers with 2 equivalents of ditriazolide I and the resulting monotriazolides hydrolyzed by the addition of aqueous pyridine (50%). The mixtures were evaporated, taken into chloroform and extracted with 0.2 M triethylammonium bicarbonate (TEAB) buffer, pH7.5. The chloroform layer was collected, concentrated to a gum, dissolved in a small amount of pyridine or chloroform and added to a stirred mixture of ether-petroleum ether (1:1 v/v, 100 ml/mmole). The resulting precipitate was collected by centrifugation, washed with ether and the residual solvent immediately removed *in vacuo*. Dinucleotides prepared in this manner have a charged phosphodiester at the 3' position and are stable for months at -20°C in the desiccator. The yields of the final products were 40-60%.

Deblocking and Purification of Oligonucleotides

The following procedure was used for both fragments. A sample of the resin (10% of the total amount) was treated with 100 µl 1M tetramethylguanidinium p-nitrobenzenealdoximate in dioxane-water (1:1 v/v) (10) and allowed to react for 12-16 hours at room temperature. 5 µl of Tetramethylguanidine was added and the reaction continued for 8 additional hours. To this mixture, 0.5 ml of pyridine and 2 ml of ammonium hydroxide were added and the tightly capped reaction tube shaken for an overnight period at 50°C. The resin was filtered off and the solution evaporated to dryness and redissolved in 1 ml 0.01 M TEAB buffer, pH8. The whole sample was applied to a Sephadex G50-80 column (2.5 X 100 cm) and eluted with 0.01 M TEAB buffer, pH8. The excluded material was collected, evaporated to dryness and redissolved in 500 µl of the same buffer. High performance liquid chromatography (HPLC) was performed using a µBondapak C_{18} column (Waters) on a SP3500 liquid chromatography (Spectra-Physics). The mixture was eluted by applying a linear acetonitrile gradient from 10-25% at pH7 (0.01 M ethylenediammonium acetate buffer); the flow rate was 2 ml min^{-1} with a sweep time of 20 minutes.

The fraction eluted at 8 minutes (Figure 3) was collected, evaporated to dryness, redissolved in 1 ml 80% acetic acid in water and allowed to stand for 15 minutes at room temperature. The reaction mixture was concentrated, redissolved in 1 ml 0.1 M TEAB buffer and extracted with ether three times. The sample was lyophilized twice and part of it (~2.0 O.D. units) electrophoresed in a 20% polyacrylamide, 7 M urea gel. Electro-elution of the DNA oligomers from the gel slices and their final purification over a BND cellulose column was performed as previously described (11) with the exception that the column buffer contained 0.1 M NaCl instead of 0.3 M NaCl.

DNA Polymerase Reactions

The 5'-ends of the polynucleotides were phosphorylated using T4 polynucleotide kinase (Miles Labs) and labeled with [γ-^{32}P]ATP. For the DNA polymerase reactions, appropriate amounts of each oligonucleotide (50 pmols for preparative procedures) were combined and the mixture brought to 40 µl of 60 mM NaCl, 50 mM Tris·HCl pH7.6, 7 mM $Mg(OAc)_2$, 7 mM DTT. The mixture was boiled for three minutes, chilled in ice and added to 10 µl of a cocktail containing 5 mM of each deoxynucleotide triphosphate and 1-5 units DNA polymerase I (Klenow fragment, Boehringer-Mannheim) in the same buffer. The reaction was allowed to proceed at room temperature for 30 minutes. The product was phenol-extracted and ethanol-precipitated. Electrophoresis of the reaction products was performed on 10% acrylamide slab gels.

DNA Sequencing

The DNA polymerase reactions were performed as above, with only one of the two fragments labeled. The double-stranded products were gel purified as described previously (11) and subjected to each of the base-specific cleavage reactions described by Maxam and Gilbert (12). The partial cleavage products were resolved on slab gels (0.05 X 20 X 50 cm), 12% polyacrylamide, 8 M urea. Autoradiography was performed at -80°C for 1-5 days with Kodak XRP-1 X-ray film and DuPont Cronex intensifying screens (13).

RESULTS AND DISCUSSION

Synthesis of Heptaicosamer and Pentaicosamer

The fragments were designed to produce a "putative *E. coli* promoter". The rationale of this design will be published elsewhere (manuscript in preparation). The sequences of both polynucleotides and their complementary regions are depicted in Figure 2A. The synthesis was carried out using published procedures (2). The dimer, trimer and tetramer blocks used (written in the 5'-3' direction) were as follows: (i) For the heptaicosamer, AT, ACT, TTA, AG, AATT, GAC, CTT, ATT and GGA were added to the nucleoside T bound to the polyacrylmorpholide resin, in the above order, along nine coupling cycles with an average yield of 78%. (ii) To produce the pentaicosamer, AA, GTT, ATA, ACAA, ATA, ATT, TAC and GAA were sequentially added to the nucleotide C, bound to the polyacrylmorpholide resin, in eight coupling cycles. The average yield was 76%.

It should be mentioned that the yield of the coupling reaction did not decrease with the growing of the polynucleotide chain. In another experiment (14) a 31-mer was synthesized using the same strategy with a similar average coupling yield. It is very difficult to synthesize polynucleotides longer than an icosamer with a defined sequence by a solution method because there is no efficient method for the isolation of the product (15). Thus, the capability for the synthesis of long polynucleotides with defined sequences may be another advantage of the solid-phase method.

Purification of Polynucleotides

In previous papers, we reported the use of ionic exchange HPLC on Permaphase AAX for the purification of the final reaction mixture (2). However, the resolution of this column is practically limited to polynucleotides smaller than an icosamer. To overcome this problem, reverse-phase HPLC was used to purify the partially deprotected polymers without removing the dimethoxytrityl (DMT) group.

Since each of the coupling cycles involved the complete removal of the DMT group of the growing chain prior to the addition of the new block and the masking of unreacted 5'-hydroxyl

groups after it, only the full-length polynucleotide product should contain the DMT function. This should allow the separation of the product from the rest of the mixture, since 5'-hydroxyl compounds are eluted faster on the reverse-phase column. The HPLC profiles showed two major fractions (Figure 3). The more retained fraction represents polynucleotides containing DMT function. The presence of more than one peak in this position of the chromatograms (seen as shoulders) reveals polynucleotide fragments that are shorter than the full-length product, yet they contain the DMT moiety. This could be due to incomplete removal of the DMT group during the synthesis and/or degradation during deblocking reactions (10). In any event, the more retained fraction (Figure 3) was collected and analyzed by gel electrophoresis after removal of the DMT group. This experiment showed that a major part of this peak was the desired length, with a small percentage of shorter polynucleotides. Preparative gel electrophoresis was, therefore, used for the final purification of the polynucleotides. Reverse-phase HPLC could be a method of general applicability for the purification of synthetic oligomers using the DMT group as a handle. Additional experiments (data not shown) showed that oligonucleotides up to 16 nucleotides in length could be resolved using this method without the requirement for any further purification step.

Preparation of Double-Stranded Fragments

The sequences of the two polynucleotides were designed to form a 10 b.p. complementary stretch along their 3'-ends. A reaction with *E. coli* DNA polymerase I (Klenow) should fill in the 5'-protruding ends, using the fragments as templates and primers for each other. Besides the structure described in Figure 2A, at least two more double-stranded combinations (depicted in Figures 2B and 2C) could be expected to form in detectable amounts. Indeed, one of these structures (Figure 2B) was formed. In order to favor the formation of the correct double-stranded product (Figure 2A), a three-fold molar excess of the 25-mer over the 27-mer was used and the DNA polymerase reaction was carried out at room temperature. Under these conditions, the formation of the desired duplex I should be preferred over

duplexes II and III (Figure 2).

Figure 4 shows an electrophoretic analysis on a polyacrylamide gel of the reaction mentioned above, along with the products of DNA polymerase reactions with 27-mer or 25-mer alone under otherwise identical conditions. The reaction is quite efficient for the duplex of the 27-mer with 25-mer, as well as for the 27-mer alone. This is not the case when the 25-mer is the only polynucleotide present. Since the desired product of the polymerase reaction differs in size from the product of duplex II by only 2 b.p., it was difficult to estimate the relative proportion of the two products by simple inspection of the corresponding bands in a gel. However, the nucleotide sequence analyses of the products mentioned (see next sub-section) was informative in this regard.

Nucleotide Sequence of the Double-Stranded Product

To obtain the nucleotide sequence of the 42 b.p. fragment, two separate reactions with DNA polymerase were performed in which either the 25-mer or the 27-mer was labeled at their 5'-ends. The full-length polymerase products were purified by gel electrophoresis and their nucleotide sequences were obtained according to the procedure of Maxam and Gilbert (13).

The sequences (Figure 5) were as expected for the chemical synthesis of the two polynucleotides and the subsequent enzymatic reaction. However, as mentioned above, it was necessary to use a three-fold molar excess of 25-mer to minimize contamination with the polymerase product of duplex II (27-mer with itself). The identity of the DNA polymerase product of 27-mer alone was confirmed by the determination of its nucleotide sequence, which is shown in Figure 6. The simplest interpretation for the appearance of a band in the A lane at position 26 of the sequence is a "correction" of the mismatched T by the 3'-5' exonuclease activity of DNA polymerase I (Klenow fragment).

From these experiments we concluded that it is desirable to avoid extensive self-complementarity at the ends of the polynucleotides. It should be noted, however, that the design of these sequences included an *Eco*R1 site, which is produced after the polymerase reaction near the 5'-end of the 27-mer fragment.

After digestion of the DNA polymerase products with the endonuclease *Eco*R1, the desired fragment should contain an *Eco*R1 "sticky" end and a blunt end, while the products resulting from structures II and III should contain two *Eco*R1 ends or two blunt ends, respectively. Thus, the proper ligation into a plasmid vehicle containing a single blunt end and an *Eco*R1 end should allow some selectivity in the recombinant products. Further experiments (to be published elsewhere) showed that the sequence of the fragment after cloning was as expected for the proper polymerase reaction and incorporation into the cloning vehicle.

These results demonstrate that the strategy combining the chemical synthesis of polynucleotides and an enzymatic DNA polymerase reaction is simple and accurate for obtaining long double-stranded DNA fragments. There are at least two major advantages in using this approach as compared to the ligation procedure (9). It requires a smaller amount of synthetic work; for a 42 b.p. fragment, only 52 (27 + 25) nucleotides were chemically synthesized, 84 would be needed for the ligation approach. Furthermore, this system does not require an extensive review of possible wrong structures and self-complementary sequences during the design for the synthesis as is the case for the ligation method. It is enough to avoid undesired potential hybrids at the 3'-end of the polynucleotides.

NOTE: A similar sequence, corresponding to an *E. coli* "putative" promoter, has been synthesized independently and preliminary results published (16) during the preparation of the manuscript for this paper.

ACKNOWLEDGEMENTS

This work was supported by USPHS grant GM26408. K. Itakura is a member of the Cancer Research Center at the City of Hope Research Institute supported by USPHS grant CA16434.

REFERENCES

1. Broka, C.; Hozumi, T.; Arentzen, R. and Itakura, K. (1980) Nucl. Acids Res. 8:5461-5471.
2. Miyoshi, K.; Huang, T. and Itakura, K. (1980) Nucl. Acids Res. 8:5491-5505.
3. Itakura, K.; Hirose, T.; Crea, R.; Riggs, A.D.; Heyneker, H.L.; Bolivar, F. and Boyer, H.W. (1977) Science 198:1056-1063.

4. Crea, R.; Kraszewski, A.; Hirose, T. and Itakura, K. (1978) Proc. Nat. Acad. Sci. USA 75:5765-5769.
5. Goeddel, D.; Heyneker, H.; Hozumi, T.; Arentzen, R.; Itakura, K.; Yansura, D.; Ross, M.; Miozzari, G.; Crea, R. and Seeburg, P.H. (1979) Nature 281:544-548.
6. Kawashima, E.H.; Gadec, T. and Caruthers, M.H. (1977) Biochemistry 16:4209-4217.
7. Goeddel, D.V.; Yansura, D.G. and Caruthers, M.H. (1978) Proc. Nat. Acad. Sci. USA 75:3578-3582.
8. Heyneker, H.; Shive, J.; Goodwen, H.; Boyer, H.; Rosenberg, R.; Dickerson, R.; Narang, S.; Itakura, K.; Lin, S. and Riggs, A.D. (1976) Nature 263:748-752.
9. Khorana, H.G. (1979) Science 203:614-625.
10. Reese, C.B.; Titmas, T.C. and Yan, L. (1978) Tetrahedron Lett. 2727.
11. Rossi, J.J.; Ross, W.; Egan, J.; Lipman, D. and Landy, A. (1979) J. Mol. Biol. 128:21-47.
12. Maxam, A. and Gilbert, W. (1980) in Methods in Enzymology 65:499-560.
13. Swanstrom, R. and Shank, P.R. (1978) Anal. Biochem. 86: 184-192.
14. Dembek, P.; Miyoshi, K. and Itakura, K. (1981) J. Amer. Chem. Soc., in press.
15. For the isolation of the product in a solution method, a column chromatography on silica gel is used but the resolution of oligonucleotide phosphotriesters longer than decanucleotide is not satisfactory.
16. Dobrynin, V.; Korobko, V.; Severtsoba, I.; Bystrob, N.; Chuvpilo, S. and Kolosov, M. (1980) Nucl. Acids Res. Symp. Series No. 7:365-376.

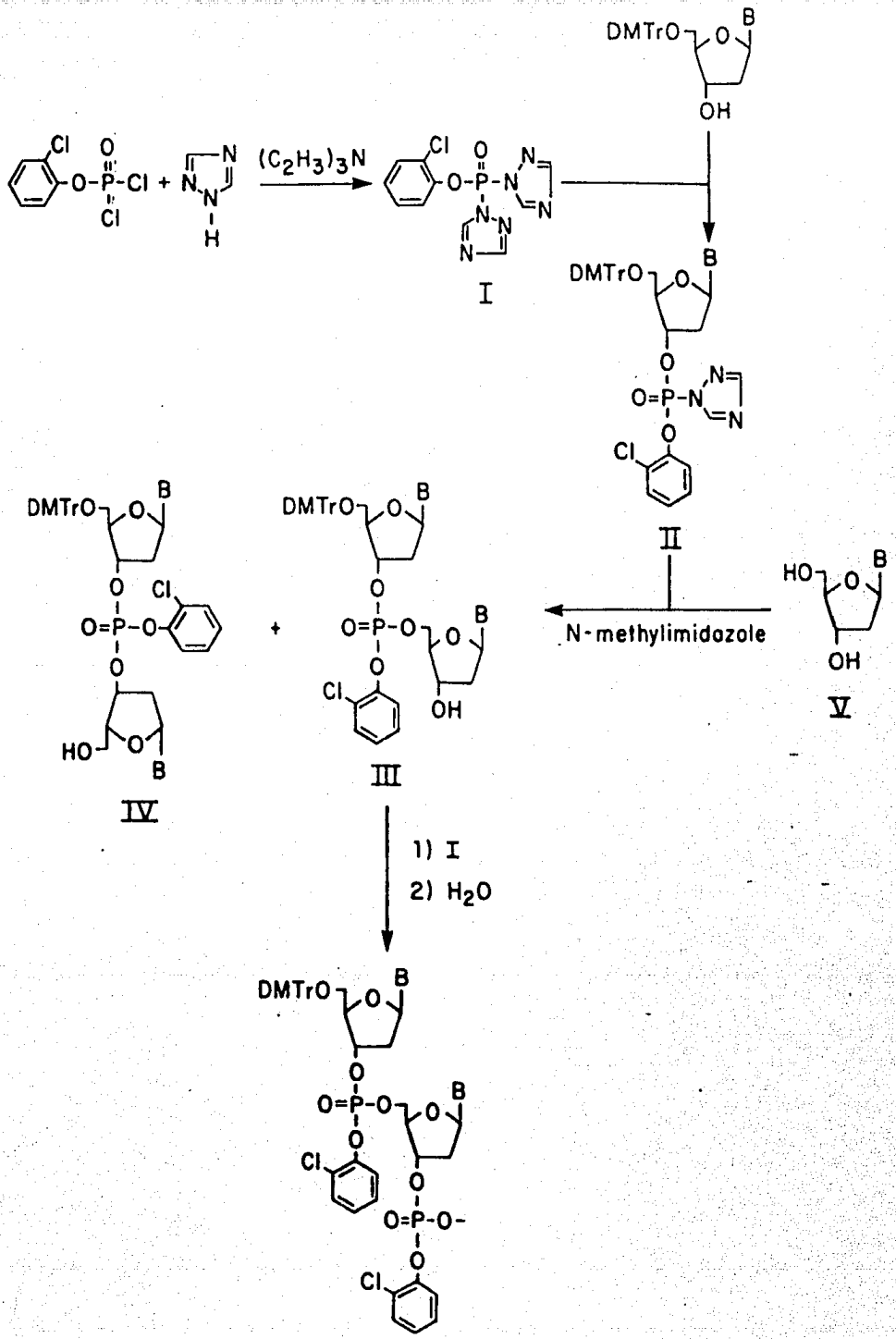
FIGURE 1. Schematic outline of the procedure used for the synthesis of dinucleotides.

FIGURE 3. HPLC profiles of the 27-mer (A) and 25-mer (B) mixtures. The DMT group was not removed prior to injection. Chromatography was performed as described in Materials and Methods.

FIGURE 4. Autoradiography of polyacrylamide electrophoresis gel of DNA polymerase reaction products. Lane B shows the product of reaction with 25-mer; lane D with 27-mer; lane E with the mixture of labeled 27-mer and a three-fold molar excess of unlabeled 25-mer. Lanes A and C show the migration of 25-mer and 27-mer, respectively, without enzyme.

FIGURE 5. Nucleotide sequence of the double-stranded product. Panel A shows the sequence of the strand primed by the 27-mer (which was labeled at its 5' end). The first band at the bottom of the autoradiogram is the C at position 7 in the sequence. Panel B shows the strand primed by the 25-mer. The sequence is read from the A at position 12 from the 5' end. The sequences of the first nucleotides, not shown in these gels, were confirmed by additional experiments (data not shown).

FIGURE 6. Nucleotide sequence of the DNA polymerase I product of Duplex II (27-mer with itself). The first G detected in the gel corresponds with position one from the 5' end. Note the A in position 26 (see text).


DMTrO
B
O
OH
Cl
O
O-P-Cl
Cl
+
N
N
N
H
(C2H3)3N
Cl
O
O-P-N
N
N
N
N
I
DMTrO
B
O
O=P-N
N
N
O
Cl
II
N-methylimidozole
HO
B
O
OH
V
DMTrO
B
O
O
Cl
O=P-O
O
HO
B
O
B
IV
+
DMTrO
B
O
O
O=P-O
B
O
Cl
OH
III
1) I
2) H2O
DMTrO
B
O
O
O=P-O
B
O
Cl
O
O=P-O-
O
Cl

A.

27-mer

5' GGAATTCTTGACAATTAGTTAACTATT 3'

3' CAATTGATAAACAATATTACATAAG 5'

25-mer

DUPLEX I

B.

27-mer

5' GGAATTCTTGACAA^{T}TAGTTAACTA^{T}T 3'

3' T_{T}ATCAATTGAT$_{T}$AACAGTTCTTAAGG 5'

27-mer

DUPLEX II

C.

25-mer

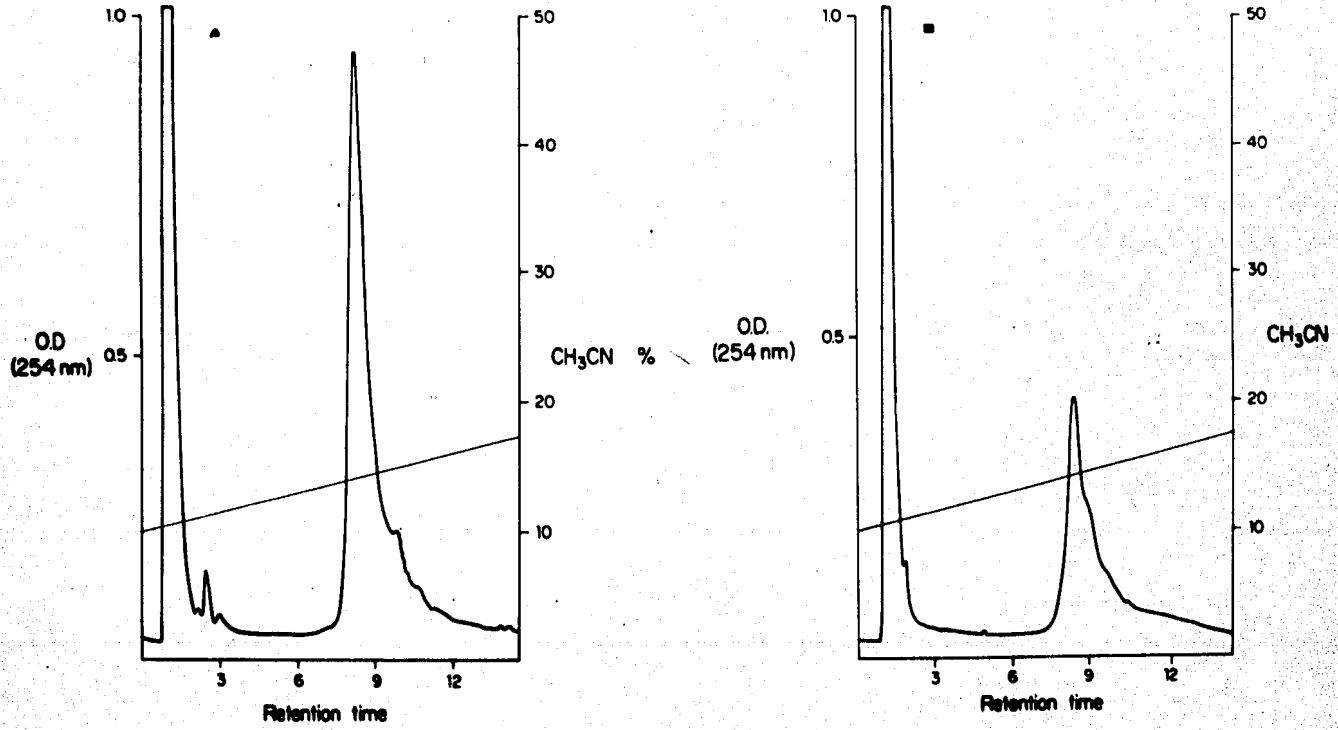
5' GAATACATTATAACAAATAGTTAAC 3'

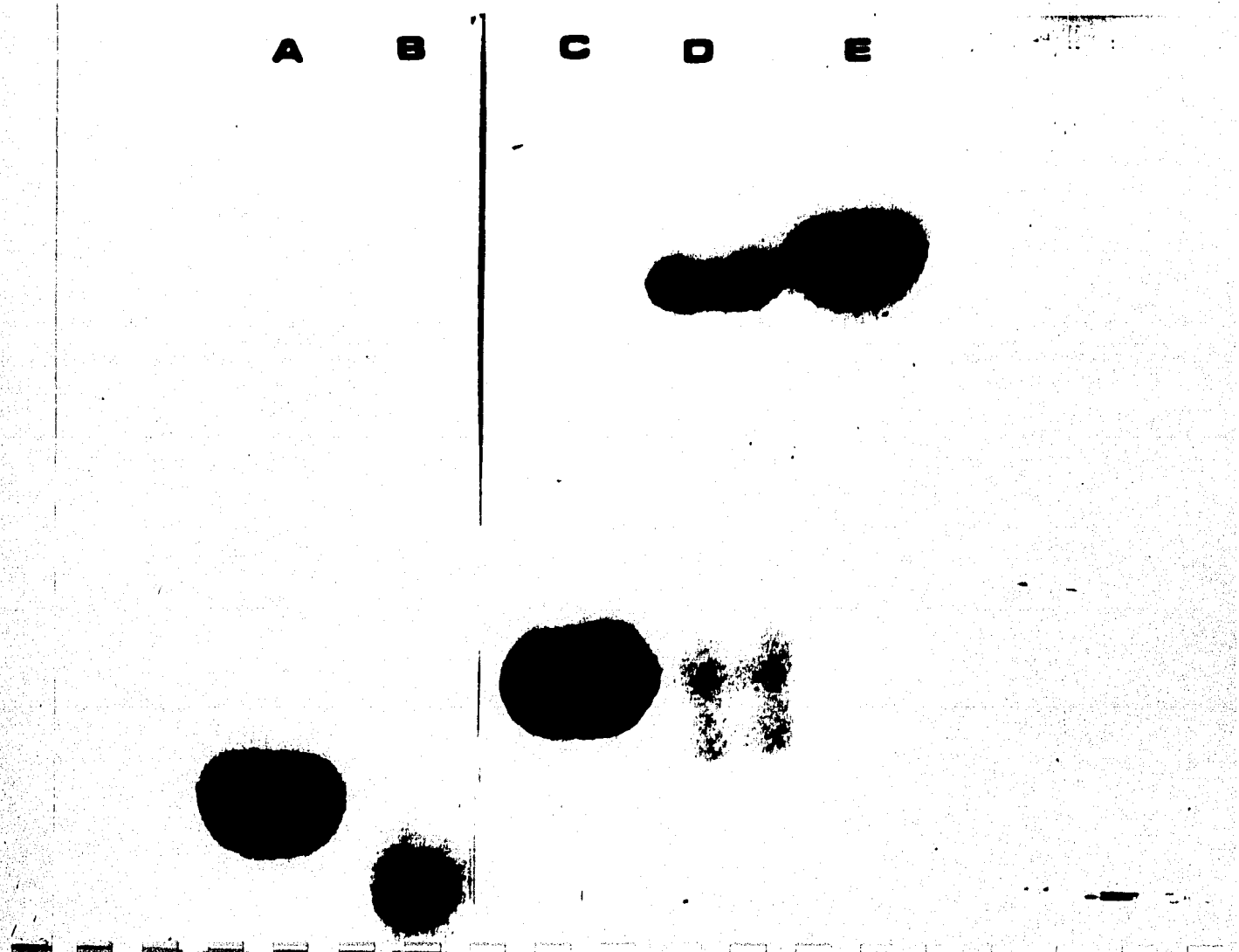
3' CAATTGATAAACAATATTACATAAG 5'

25-mer

DUPLEX III

FIGURE 2. Nucleotide sequences of single-stranded fragments and alternative duplex structures they can form.


1.0
0.5
O.D
(254 nm)
50
40
30
20
10
CH3CN %
3
6
9
12
Retention time
1.0
0.5
O.D.
(254 nm)
50
40
30
20
10
CH3CN
3
6
9
12
Retention time


A
B
C
D
E

A
G A T C
T G A A T T G T T A T C A T T G A T T A C A G T T C
B
G A T C A
G A C T G T T A A T C A T T G A T T A A A C A A

G
A
T
C
E
A
A
C
G
T
T
T
T
(A)
T
A
T
C
A
T
A
T
T
A
A
C
A
G
T
T
C
T
T
A
A
G
G