DTIC FILE COPY

AD ____________

AD-A199 075

STRUCTURE AND FUNCTIONAL STUDIES ON DENGUE-2 VIRUS GENOME

Annual Report

Radha Krishnan Padmanabhan, Ph.D.

March 1, 1986

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD17-82-C-2051

DTIC ELECTE SEP D

University of Kansas Medical Center
Department of Biochemistry
Kansas City, KS 66103

DOD DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

88 9 22 017

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: University of Kansas Medical Center; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code): Department of Biochemistry, 39th & Rainbow Blvd., Kansas City, KS 66103	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: U.S. Army Medical Research & Development Command; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: Contract No. DAMD17-82-C-2051
8c. ADDRESS (City, State, and ZIP Code): Fort Detrick, Frederick, Maryland 21701-5012	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61102A	3M16-1102BS10	AA	092

11. TITLE (Include Security Classification)

Structural and Functional Studies on Dengue-2 Virus Genome

12. PERSONAL AUTHOR(S)

Radha Krishnan Padmanabhan

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual Report	FROM 9/1/84 TO 8/31/85	1986 March 1	28

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
06	03	
06	01	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

RA 1; Characterization of three CDNA clones by hybridization; complete DNA sequence analysis of these clones totalling 4,586 nucleotides; Deduced amino acid sequences of dengue viral antigens.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

1. Out of eleven cDNA clones, we selected three for further characterization by DNA sequence analysis. Two of these clones overlapped by 470 base pairs.

2. For DNA sequence analysis of these clones, we used the strategy of Guo et. al. (1983). Complete DNA sequence of these three clones have been determined. The region sequenced represents about 42% of the dengue virus genome.

3. Examination of the primary nucleotide sequences using computer programs developed for translation in all possible reading frames revealed the presence of long open reading frames spanning the entire length of the cDNA clones. From the DNA sequences, amino acid sequences of segments of putative polyprotein precursor of dengue virus genome were derived. The putative polypeptides derived from the DNA sequence are 885 and 643 amino acids in length.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Mrs. Virginia M. Miller	22b. TELEPHONE (Include Area Code): 301-663-7325; 22c. OFFICE SYMBOL: SGRD-RMI-S

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

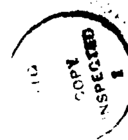
FOREWORD

The investigators have abided by the National Institutes of Health Guidelines for Research involving Recombinant DNA Molecules (April 1982) and the Administrative Practices Supplements, as indicated in the Memorandum of Understanding and Agreement, reviewed orginally by Dr. Larry H. Baker (late), Chairman of the Institutional Biosafety Committee and by Dr. Stanley Barban, Scientific Administrator, Office of Recombinant DNA Activity, N.I.H.

Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

TABLE OF CONTENTS

		Page No.
1.	Introduction	1
2.	Body of the report	
2.1	Characterization of cDNA clones by Southern hybridization	2-3
2.2	Strategy for subcloning pVV1 and pVV9 cDNA	4
2.3	Progressive shortening of the pVV1 and pVV9 after Bal 31 digestion and cloning into E. coli JM83	4-5
2.4	Complete DNA sequence analysis of the subclones of pVV1 and pVV9 generated by Bal-31 and the verification of DNA sequence for the cDNA clone, pVV17 by sequencing the complementary strand	5-6
2.5	Analysis of the DNA sequences of pVV17, pVV1 and pVV9 for the presence of open reading frame DNAs which are putative regions coding for the dengue viral antigens	6-7
2.6	Amino acid sequences of the dengue viral polypeptides deduced from the DNA sequence	7-8
3.	Appendixes	9-20
3.1	Figure Legends	9-11
3.2	Figures	12-20a
4.	Literature Cited	21
5.	Personnel Supported by the Contract	22



Accession For
NTIS GRA&I
DTIC TAB
Unannounced
Justification
Dist Special
A-1

INTRODUCTION

The progress made from September 1, 1984 to August 31, 1985 can be divided into the following categories.

1. Characterization of two cDNA clones, pVV1 and pVV9 by Southern hybridization.
2. Strategy for subcloning pVV1 and pVV9 DNA.
3. Progressive shortening of the pVV1 & pVV9 DNA after Bal31 digestion and cloning into E.coli JM83.
4. Complete DNA sequence analysis of the subclones of pVV1 and pVV9 generated by Bal-31 and the verification of DNA sequence for the cDNA clone, pVV17 by sequencing the complementary strand.
5. Analysis of the DNA sequences of pVV17, pVV1 and pVV9 for the presence of open reading frame DNAs which are putative regions coding for the dengue viral antigens.
6. Amino acid sequences of the dengue viral polypeptides deduced from the DNA sequence.

1. Characterization of cDNA clones by Southern hybridization.

a. Rationale

During the period, Sept. 1, 1983 - Aug. 31, 1984, we obtained eleven cDNA clones pVV1,2,4,7,8,9,11,14,15,17 and 18 which had cDNA inserts, ranging in sizes from 0.95 kilobases to 2.0 kilobases (see Fig. 7 of Report Number Three). Of these clones, only nine clones had cDNA inserts that could be released by PstI digestion. In the clones pVV14 and 18, one of the PstI sites were lost. The further characterization of these two clones is not complete and therefore is not reported here. However, the characterization of the other nine clones were carried out by Southern hybridization (1975). The results are described in the following section.

b. Experimental

In order to characterize these cDNA clones with respect to the overlapping regions present in them, several Southern hybridization (see Southern, 1975) were carried out. In these experiments plasmid DNAs were prepared from these nine clones. Each was digested with two different combinations of restriction enzymes, BglII + PstI and Pst I + EcoRI, except with cDNA clones, pVV7, pVV1 and pVV9. pVV7 was digested only with BglII + PstI, pVV1 with PstI + PvuII and pVV9 with KpnI + Pst I. (Fig. 1b). In Fig. 1a, these plasmid DNAs were digested with only PstI as a control to release the cDNA inserts. These restriction digests were electrophoresed on an agarose gel (1%) and the DNA fragments were transferred to a Gene Screen Plus membrane filter (NEN-DuPont). The DNA fragments immobilized on the membrane filter were hybridized with the probes prepared from the cDNA inserts of pVV17 (Fig. 2a), pVV4 (Fig. 2b), pVV7 (Fig. 2c) and pVV11 (Fig. 2d) by nick translation. The results, summarized in Table I indicate that the cDNAs from the plasmids pVV2,4,7,8,11,15 hybridized with the probe from pVV17 and therefore share common sequences with pVV17. In contrast, these cDNAs did not

hybridize with pVV1 and 9. In order to examine whether pVV1 and pVV9 shared any sequences, the pVV1 plasmid DNA was digested with either PstI alone (lane 2, Fig. 2e) or PstI + PvuII (lane 3, Fig. 2e). As controls, undigested pVV17 DNA and PstI digested pVV9 DNA were used. These DNAs were electrophoresed on an agarose gel, transferred to a Gene Screen Membrane filter and hybridized with the cDNA probe from pVV9 plasmid. The results shown in Fig. 2e indicated that pVV1 and pVV9 shared common sequences and the homologous sequences are located in the large fragment obtained from the digestion of pVV1 cDNA insert with PvuII. Further hybridization analysis of pVV1 and pVV9 cDNAs subsequent to double digestion with KpnI and PstI showed that pVV9 cDNA is located 5' to pVV1 cDNA and that these two cDNA shared approximately 450 base pairs (see Fig. 3).

Table I

Southern hybridization analysis of DEN-2 cDNA clones

Expt.	Probe	Hybridization signal to cDNA clones +	-
1	17	2,4,7,8*,11,15,17	1,9
2	11	2,4,11,15,17	1,9,7
3	4	2,4,7,11,15,17	1,9
4	7	2,4,7,17	1,9,11,15

*pvv8 was not tested in experiments 2-4.

2. Strategy for subcloning pVV1 and pVV9 cDNA.

a. Rationale.

In order to analyze the organization of a viral genome and localization of the gene coding for viral specific structural and non-structural proteins, it is important to determine the nucleotide sequence from which the amino acid sequence of the proteins can be deduced (see the following sections). Several strategies for rapid sequencing of long DNA fragments, which are over 2000 bp in length, have become available recently using either the phage M13 system (Barnes and Bevan, 1983) or the plasmid system (Guo et.al. 1983). The latter strategy was used for sequencing pVV17 cDNA clone (see Report Number Three for details).

The same strategy was used for sequencing pVV1 and pVV9 cDNA clones. In principle, a long target DNA is progressively shortened from one end, by digestion with Bal-31 nuclease, followed by cleaving off the shortened vector DNA. The family of the shortened target DNA molecule, is subsequently cloned between the PstI site and SmaI site, within the polylinker region of pUC13'-1 (Fig. 4). DNA fragments cloned into this plasmid are sequenced directly by the chemical method of Maxam and Gilbert (1977). A few ambiguities in our sequence analysis were clarified by sequencing the complementary strand by labeling the 3' or 5' terminus.

3. Progressive shortening of the pVV1 & pVV9 DNA after Bal31 digestion and cloning into E.coli JM83.

a. Background

Bal 31 nuclease is a processive exonuclease which is capable of digesting double stranded DNA to give rise to progressively shortened DNA molecules. The extent of digestion is controlled by varying the incubation time. The termini of the Bal 31

generated molecules are not exactly blunt-ended and therefore need to be treated with the DNA polymerase (large fragment) from E.coli to make them exactly blunt-ended.

b. Experimental

The conditions for Bal31 nuclease digestion of pVV1 and pVV9 were the same as described in the Report Number Three for pVV17 clone. The only differences in the protocol are the omission of SalI linkers to the blunt-ended termini (Step iv on page 18 of Annual Report #3) and the use of PstI-SmaI site of the vector for cloning (step vi on page 19 of Annual Report #3). Several transformants were screended and the mini-preparation of plasmid DNA was carried out. The size of the shortened target DNA was determined after digestion with two restriction enzymes (PstI + BamHI) and electrophoresed with appropriate size markers on an agarose gel (1%). Fig. 5 shows a stepwise ladder formation of shortened DNAs which were obtained from various subclons of pVV9 cDNA. Similar protocol was used for pVV1 DNA. The subclones of pVV1 and 9 were used for DNA sequence analysis.

4. Complete DNA sequence analysis of the subclones of pVV1 and pVV9 generated by Bal-31 and the verification of DNA sequence for the cDNA clone, pVV17 by sequencing the complementary strand.

a. Background

The shortened inserts in these subclones are such that the difference in cDNA insert size between any two of them is approximately 200-250 nucleotides (Fig. 5). The rationale for choosing such a progression of shortened cDNA is that if the DNA sequence analysis is carried out from a fixed site of these cDNA clones, and if we are able to obtain sequence of 250-300 nucleotides per sequencing gel, then each block of sequence data is expected to overlap with the next one (from a shortened cDNA) by about 50

nucleotides. This strategy will enable us to obtain the complete sequence of a cDNA, 2 kb in length, very rapidly by sequencing the subset of shortened cDNA clones from the fixed site of the plasmid. We used the chemical method of Maxam and Gilbert (1977) for sequence analysis.

b. Experimental

The plasmid DNA (20 μg) from the parent clone (pVV1 or pVV9) or subclone of either parent clone was digested with BamH1 (25 units) at 37°C for 60 min. The reaction mixture was then treated accordingly depending on whether the DNA was to be labeled at the 3' end or the 5' end. The end labeling methods of DNA are described in detail by Maniatis et. al. (1982) and were followed. After labeling, the DNA was digested with the appropriate restriction enzyme (PstI or HindIII), and the resultant fragments were separated by low melting point (LMP-) agarose gel electrophoresis. The labeled fragement of interest was recovered from the gel by phenol extraction. The purified fragment was used for sequence analysis.

The complete nucleotide sequence of the DEN-2 cDNA clone pVV1 and pVV9 which overlapped by 470 bp is given in Fig. 6a. In Fig. 6b, the complete sequence of pVV17 clone of DEN-2 RNA is given. This sequence is essentially the same as that reported in Annual Progress Report (#3) except that a few ambiquities which existed in the previous sequence have been corrected by sequencing the complementary strand of selected subclones of pVV17 cDNA.

5. Analysis of the DNA sequences of pVV17, pVV1 and pVV9 for the presence of open reading frame DNAs which are putative regions coding for the dengue viral antigens.

a. Background and Results

Once the primary sequences of a DNA segment has been determined, it is important to locate the protein coding region(s) within that segment in order to genetically engineer that segment for expression. The characteristic feature of a protein coding region within DNA segment is the presence of an open reading frame (ORF) with the occurrence of the initiation codon ATG at the beginning of the ORF and a termination codon at the end of the ORF. The distance between these two codons depends on the size of the protein encoded by this region. There are computer programs available (we have one set of such programs from International Biotechnologies, Inc., Connecticut) to scan the primary sequence of a DNA segment to locate protein coding regions. The DNA sequences of both strands of pVV91 (the composite sequence of pVV9 and pVV1 clones) and pVV17 were scanned in the 5'---3' orientation in all possible (six) reading frames [illegible] ORFs. As shown in Fig. 7, the third reading frame of pVV91 and pVV1 sequences contained an open reading frame which spanned the entire length of the sequence, indicating that this is the reading frame which is most likely used for translation into DEN-2 viral proteins and that the polarity of the DNA sequence in 5' -3' is same as that of DEN-2 RNA genome.

All the other five reading frames contained many termination codons and are therefore, not involved in translation. Presence of long ORF in all the three cDNA clones of DEN-2 RNA tend to support the notion that DEN-2 virus similar to yellow fever virus genome in having a single long ORF coding for a polyprotein precursor (Rice et al. 1985).

6. Amino acid sequences of the dengue viral polypeptides deduced from the DNA sequence.

a. Background and Results

The translation of the DNA sequence of DEN-2 cDNA clones, pVV91 (composite sequence) and pVV17 into the putative precursor polypeptide sequences of Dengue viral antigens using a computer program is shown under the DNA sequences in Fig. 6a and 6b. These are 885 and 643 amino acids in length, respectively.

LEGEND TO FIGURES

Figure 1 Restriction enzyme analysis of cDNA clones of DEN-2 RNA.

a. The plasmid DNAs including the control vector (lane 11) was digested with only PstI. The plasmid DNAs that were used for this experiment were: pUC 13'-1 (lane 11), pVV1 (lane 10), pVV2 (lane 9), pVV4 (lane 8), pVV7 (lane 7), pVV8 (lane 6), pVV9 (lane 5), pVV11 (lane 4), pVV15 (lane 3) and pVV17 (lane 2).

b. The plasmids containing the cDNA inserts of DEN-2 RNA were prepared and digested with a combination of two enzymes, BglII + PstI (lanes 2,4,6,7, 9,12), PstI + EcoRI (lanes 3,5,8, and 13), PstI + PvuII (lane 10) or PstI + KpnI (lane 11). The plasmid DNA used are: pVV2 (lanes 2 and 3), pVV4 (lanes 4 and 5), pVV7 (lane 6), pVV15 (lanes 7 and 8), pVV11 (lanes 9), pVV1 (lane 10), pVV9 (lane 11) and pVV17 (lanes 12 and 13). DNA digested with either Hind III alone (Fig. 1a, lane 12) or with a mixture of EcoRI + Hind III (Fig. 1a & 1b, lane 1) were used as DNA size markers. The digests were electrophoresed on an agarose gel (1%), stained with ethidium bromide and photographed.

Figure 2 Southern hybridization analysis of cDNA clones from DEN-2 RNA.

The restriction digests from Fig. 1 were electrophoresed and transferred to Gene Screen Plus Membrane Filter (NEN-DuPont). The DNAs immobilized onto the filter were then hybridized to the probes, prepared by nick-translation of cDNA inserts from the clones pVV17 (Fig. 2a), pVV4 (Fig. 2b), pVV7 (Fig. 2d) and pVV11 (Fig. 2c). The bands that hybridized were visualized by autoradiography. Fig. 2a, lane 1, λ HindIII + EcoRI;

lane 2, pUC13'-1; lanes 3-11, pvv1,2,4,7,8,9,11,15 and 17, respectively. The numbered lanes in Fig. 2 b-d corresponds to Fig. 1b. Fig. 2e. Southern hybridization of pVV1 and pVV9 DNAs to localize homologous sequences. pVV1 DNA was digested with either PstI alone (lane 2) or PstI + PvuII (lane 3). Lane 1 contained DNA from pVV17 (undigested); lane 4 contained a PstI digest of pVV9 DNA. The DNAs were electrophoresed on an agarose gel (1%), transferred to a Gene Screen membrane filter and hybridized with the nick-translated probe of pVV9-cDNA insert. The bands were detected by autoradiography.

Figure 3 Mapping of DEN-2 cDNA clones by restriction enzyme analysis and Southern hybridization.

The results from Fig. 1 and Fig. 2 is summarized in this graphic plot. The numbers refer to the number assigned to cDNA clones. The overlapping region between pVV1 and pVV9 is about 450 bp. E = EcoRI; S = Stu I; B = Bgl II; K = Kpn I.

Figure 4 Strategy for kilo-base sequencing of the target DNA. The strategy is the same as described by Guo et al. (1983).

Figure 5 Size analysis of pVV9 cDNA subclones by agarose gel electrophoresis.

pVV9 DNA (210 μg) was linearized by digesting with EcoRI. It was then digested with Bal31 nuclease (1.5 units) in a total volume of 482 μl under conditions described by Guo et al. (1983). Aliquots were taken at various time intervals and were analyzed by electrophoresis on an agarose gel. Progressively shortened DNA molecules were cloned between PstI and SmaI sites of PUC13'-1 vector and were used for transformation of E.coli JM83. Transformants were screened for the size of the cDNA inserts by double digestion with restriction enzymes, Bam H1 + Pst I. Lanes 1-5 contained DNA from subclones derived by Bal 31 digestion of pVV9 DNA for 2' (lane 1), 6'

(lanes 2 and 3) and 8' (lanes 4 and 5). Lanes 6 and 7 contained DNA molecular weight markers. Lanes 8-14 contained DNA clones derived from Bal 31 digested pVV9 DNA for 6' (lanes 8-10 and 14) 8' (lanes 11 and 12) and 10' (lane 13).

Figure 6 DNA sequence of DEN-2 cDNA clones.

a. The cDNA clones pVV1 and 9 were completely sequenced. We found that these two clones overlapped by 470 bp as expected from the hybridization results. The composite sequence of these two clones (call "pVV91" in the text) is shown.

b. The sequence of pVV17 clone is shown. This sequence was derived by sequencing about 75% of the complementary strand of pVV17 clone, the sequence of which was originally reported in Report Number Three. A few errors in the previous sequence were corrected.

Figure 7 Graphic analysis of open reading frame DNA.

The primary nucleotide sequence as shown in Fig. 6a & b and their complements were translated using computer program in all possible reading frames. The location of the stop codons are indicated by vertical lines.

Fig.1a

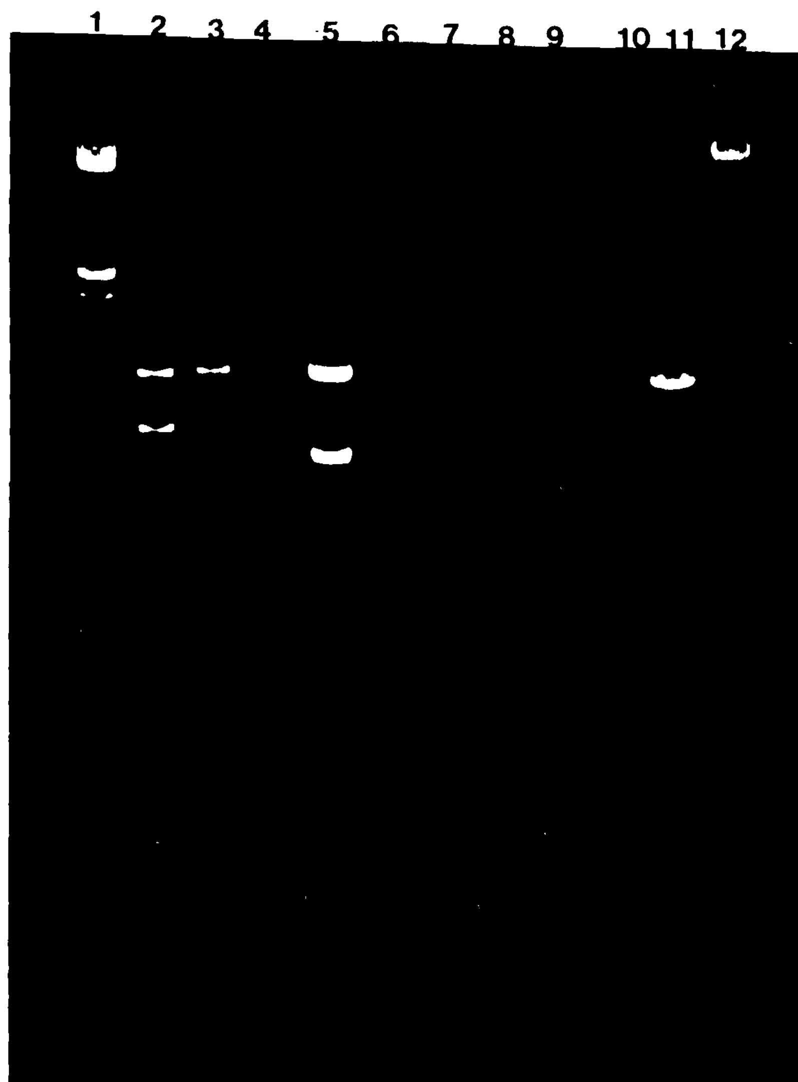


Fig.1b

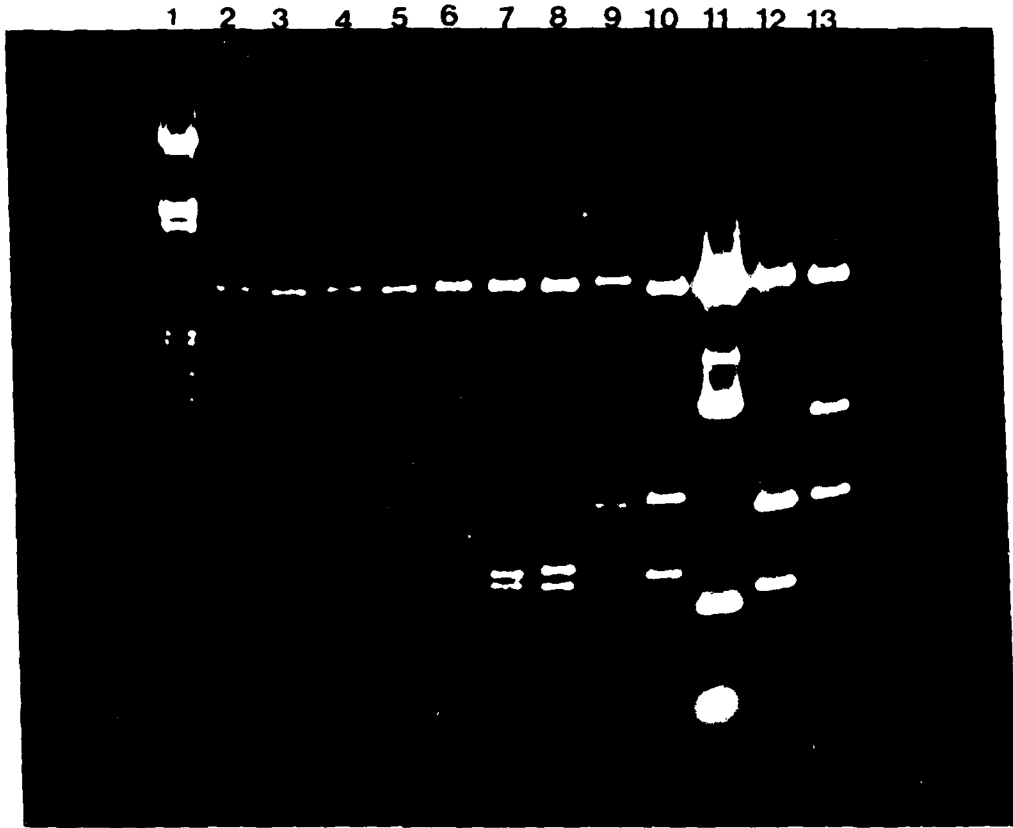


Fig.2a

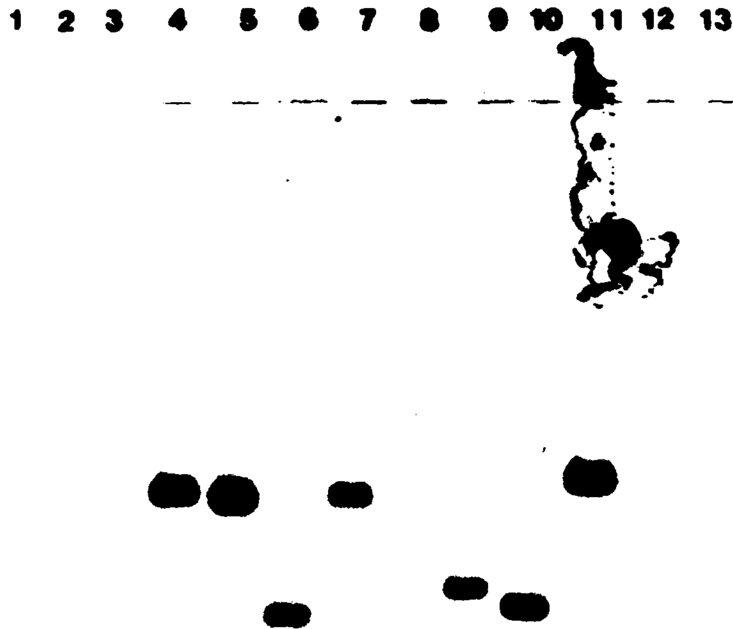


Fig.2b

1 2 3 4 5 6 7 8 9 10 11 12 13

Fig.2c

1 2 3 4 5 6 7 8 9 10 11 12 13

Fig.2d

1 2 3 4 5 6 7 8 9 10 11 12 13

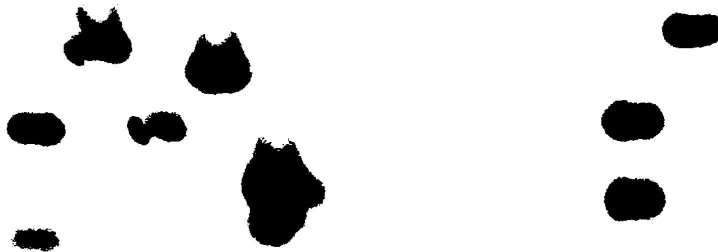
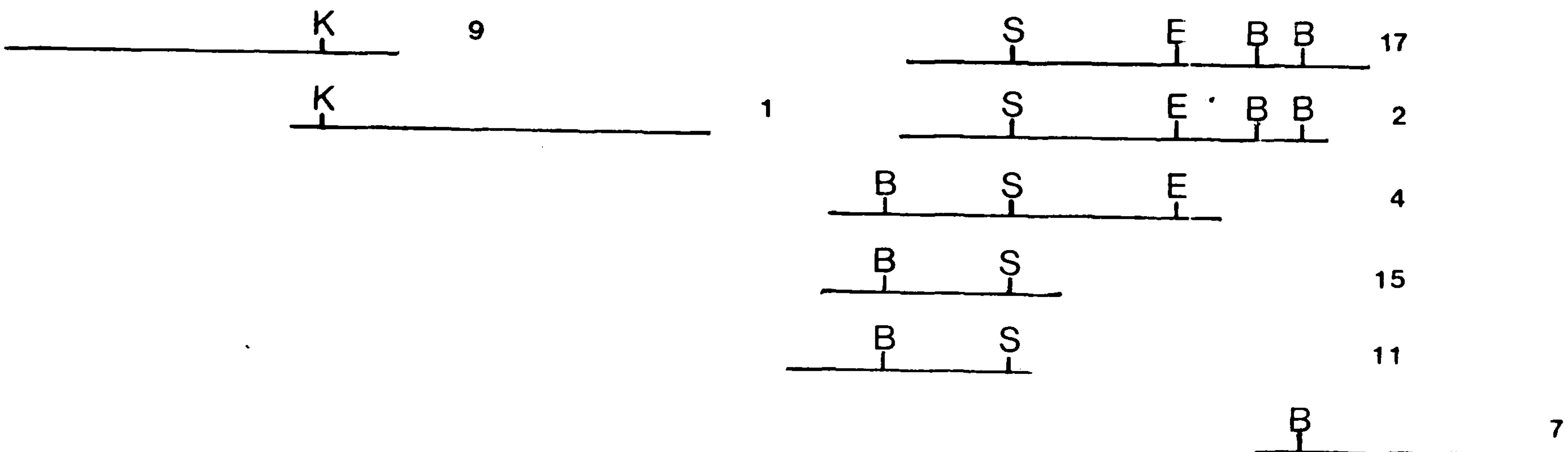


Fig.2e

1 2 3 4

Fig.3



target DNA

Bal 31 nuclease

Klenow fragment , dNTPs

Restriction enzyme linkers

Restriction enzymes digestion

+

1. Ligate

2. Transform

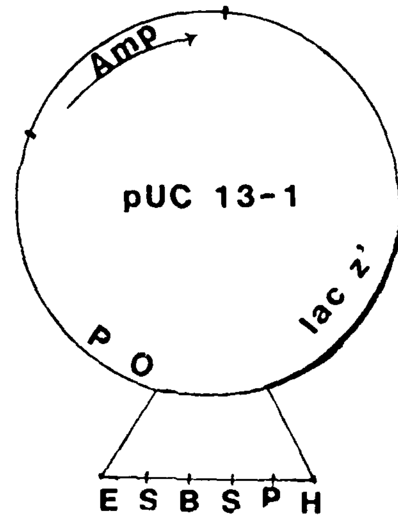


Fig.5

Fig.6a

[illegible]

Fig.6b

```
GAC AAC TTA GCA GTG CTG CAC ACG GCT GAA GCA GGT GGA AGG GCG TAC AAT CAT GCT CTC AGT GAA CTG CCG GAG ACC CTG GAG ACA TTG   90
 D   N   L   A   V   L   H   T   A   E   A   G   G   R   A   Y   N   H   A   L   S   E   L   P   E   T   L   E   T   L   30
CTT TTA CTG ACA CTT CTG GCT ACA GTC ACA GGA GGA ATC TTT TTA TTC TTG AAT AGC GGA AGG GGT ATA GGG AAG ATG ACC CTG GGA ATG  180
 L   L   L   T   L   L   A   T   V   T   G   G   I   F   L   F   L   N   S   G   R   G   I   G   K   M   T   L   G   M   60
TGC TGC ATA ATC ACG GCT AGT ATT CTC CTA TGG TAC GCA CAA ATA CAG CCA CAC TGG ATA GCA GCT TCA ATA ATA CTG GAG TTT TTT CTC  270
 C   C   I   I   T   A   S   I   L   L   W   Y   A   Q   I   Q   P   H   W   I   A   A   S   I   I   L   E   F   F   L   90
ATA GTT TTG CTT ATT CCA GAA CCA GAA AAG CAG AGA ACA CCC CAA GAT AAC CAA TTG ACC TAC GTT GTC ATA GCC ATC CTC ACA GTG GTG  360
 I   V   L   L   I   P   E   P   E   K   Q   R   T   P   Q   D   N   Q   L   T   Y   V   V   I   A   I   L   T   V   V  120
GCC GCA ACC ATG GCA AAC GAG ATG GGT TTC CTG GAA AAA ACG AAG AAA GAT CTC GGA TTG GGA AGC ATT ACA ACC CAG CAA CCC GAG AGC  450
 A   A   T   M   A   N   E   M   G   F   L   E   K   T   K   K   D   L   G   L   G   S   I   T   T   Q   Q   P   E   S  150
AAC ATC CTG GAC ATA GAT CTA CGT CCC GCA TCA GCA TGG ACG CTG TAT GCT GTG GCC ACA ACA TTT GTC ACA CCA ATG TTG AGA CAC AGC  540
 N   I   L   D   I   D   L   R   P   A   S   A   W   T   L   Y   A   V   A   T   T   F   V   T   P   M   L   R   H   S  180
ATT GAA AAT TCC TCA GTG AAC GTG TCC CTA ACA GCT ATT GCC AAC CAA GCC ACA GTG TTA ATG GGT CTT GGG AAA GGA TGG CCA TTG TCA  630
 I   E   N   S   S   V   N   V   S   L   T   A   I   A   N   Q   A   T   V   L   M   G   L   G   K   G   W   P   L   S  210
AAG ATG GAC ATC GGA GTT CCC CTT CTC GCC ATT GGA TGC TAC TCA CAA GTC AAC CCC ATA ACT CTC ACA GCA GCT CTT TTC TTA CTG GTA  720
 K   M   D   I   G   V   P   L   L   A   I   G   C   Y   S   Q   V   N   P   I   T   L   T   A   A   L   F   L   L   V  240
GCA CAT TAT GCC ATC ATA GGG CCA GGA CTC CAA GCA AAA GCA ACC AGA GAA GCT CAG AAA AGA GCA GCA GCG GGC ATC ATG AAA AAC CCA  810
 A   H   Y   A   I   I   G   P   G   L   Q   A   K   A   T   R   E   A   Q   K   R   A   A   A   G   I   M   K   N   P  270
ACT GTC GAT GGA ATA ACA GTG ATT GAC CTA GAT CCA ATA CCC TAT GAT CCA AAG TTT GAA AAG CAG TTG GGA CAA GTA ATG CTC CTA GTC  900
 T   V   D   G   I   T   V   I   D   L   D   P   I   P   Y   D   P   K   F   E   K   Q   L   G   Q   V   M   L   L   V  300
CTC TGC GTG ACT CAA GTG TTG ATG ATG AGG ACT ACA TGG GCT CTG TGT GAG GCT TTA ACC TTA GCG ACC GGG CCT ATC TCC ACA TTG TGG  990
 L   C   V   T   Q   V   L   M   M   R   T   T   W   A   L   C   E   A   L   T   L   A   T   G   P   I   S   T   L   W  330
GAA GGA AAT CCA GGG AGG TTT TGG AAC ACT ACC ATT GCA GTG TCA ATG GCT AAC ATT TTT AGA GGG AGT TAC TTG GCC GGA GCT GGA CTT 1080
 E   G   N   P   G   R   F   W   N   T   T   I   A   V   S   M   A   N   I   F   R   G   S   Y   L   A   G   A   G   L  360
CTC TTT TCC ATC ATG AAG AAC ACA ACC AAC ACG AGA AGG GGA ACT GGC AAC ATA GGA GAG ACG CTT GGA GAG AAA TGG AAA AGC CGA TTG 1170
 L   F   S   I   M   K   N   T   T   N   T   R   R   G   T   G   N   I   G   E   T   L   G   E   K   W   K   S   R   L  390
AAC GCA TTG GGG AAA AGT GAA TTC CAG ATC TAC AAG AAA AGT GGA ATC CAG GAA GTG GAT AGA ACC TTA GCA AAA GAA GGC ATT AAA AGA 1260
 N   A   L   G   K   S   E   F   Q   I   Y   K   K   S   G   I   Q   E   V   D   R   T   L   A   K   E   G   I   K   R  420
GGA GAA ACG GAC CAT CAC GCT GTG TCG CGA GGC TCA GCA AAA CTG AGA TGG TTC GTC GAG AGA AAT ATG GTC ACA CCA GAA GGG AAA GTA 1350
 G   E   T   D   H   H   A   V   S   R   G   S   A   K   L   R   W   F   V   E   R   N   M   V   T   P   E   G   K   V  450
GTG GAC CTC GGT TGC GGC AGA GGA GGC TGG TCA TAC TAT TGT GGG GGA CTA AAG AAT GTA AGA GAA GTC AAA GGC CTA ACA AAA GGA GGA 1440
 V   D   L   G   C   G   R   G   G   W   S   Y   Y   C   G   G   L   K   N   V   R   E   V   K   G   L   T   K   G   G  480
CCA GGA CAT GAA GAA CCC ATC CCC ATG TCA ACA TAT GGG TGG AAT CTA GTA CGT CTT CAA AGT GGA GTT GAC GTT TTC TTC ACT CCG CCA 1530
 P   G   H   E   E   P   I   P   M   S   T   Y   G   W   N   L   V   R   L   Q   S   G   V   D   V   F   F   T   P   P  510
GAA AAG TGT GAC ACA TTG TTG TGT GAC ATA GGG GAG TCG TCA CCA AAT CCC ACG GTA GAA GCA GGA CGA ACA CTC AGA GTC CTT AAC TTA 1620
 E   K   C   D   T   L   L   C   D   I   G   E   S   S   P   N   P   T   V   E   A   G   R   T   L   R   V   L   N   L  540
GTG GAA AAT TGG TTG AAC AAC AAC ACC CAA TTT TGC ATA AAG GTT CTC AAC CCA TAC ATG CCC TCA GTC ATA GAA AAA ATG GAA GCA CTA 1710
 V   E   N   W   L   N   N   N   T   Q   F   C   I   K   V   L   N   P   Y   M   P   S   V   I   E   K   M   E   A   L  570
CAA AGG AAA TAT GGA GGA GCC TTA GTG AGG AAT CCA CTT TCA CGA AAC TCC ACA CAT GAG ATG TAC TGG GTA TCC AAT GCC TCC GGG AAC 1800
 Q   R   K   Y   G   G   A   L   V   R   N   P   L   S   R   N   S   T   H   E   M   Y   W   V   S   N   A   S   G   N  600
ATA GTG TCA TCA GTG AAC ATG ATT TCA AGG ATG TTG ATC AAC AGA TTC ACA ATG AGA CAC AAG AAA GCC ACT TAC GAG CCA GAT GTA GAC 1890
 I   V   S   S   V   N   M   I   S   R   M   L   I   N   R   F   T   M   R   H   K   K   A   T   Y   E   P   D   V   D  630
CTC GGA AGC GGA ACC CGC AAC ATC GGA ATT GAA AGT GAG                                                                      1929
 L   G   S   G   T   R   N   I   G   I   E   S   E                                                                        643
```

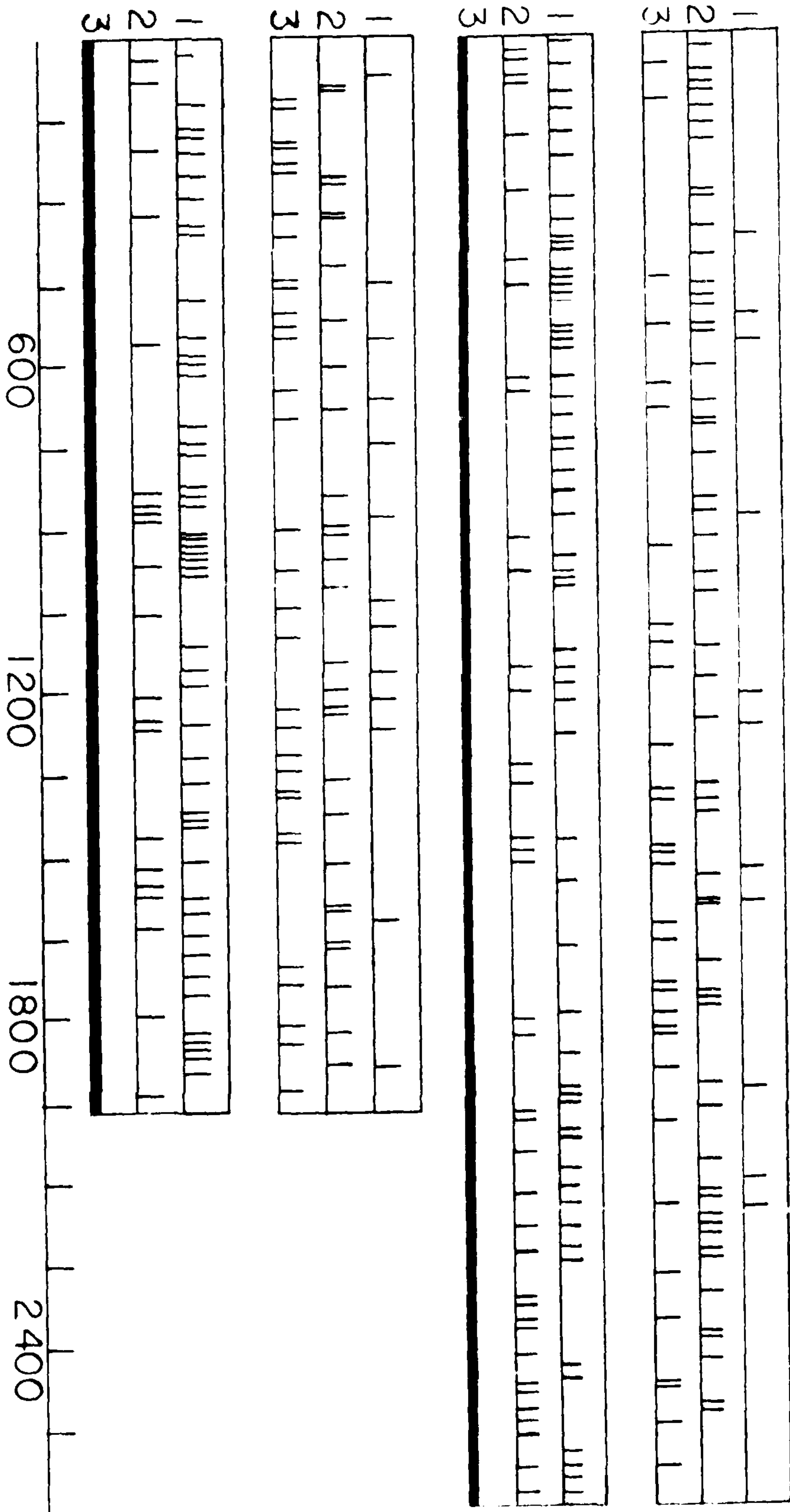


Fig.7

LITERATURE CITED

1. Barnes, W.M. and Bevan, M. (1983) Nuc. Acids Res. 11, 349-368.

2. Guo, L.-H., Yang, R.C.A. and Wu, R. (1983) Nuc. Acids Res. 5521-5540.

3. Maniatis, T., Fritsch. E.F. and Sambrook, J. (1982) Molecular cloning-a laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

4. Maxam, A.M. and Gilbert, W. (1980) Methods in enzymology (Grossman, L. and Moldave, K., eds) Academic press, New York, pp. 488-560.

5. Rice, C.M., Lenches, E.M., Eddy, S.R., Shin, S.J., Sheets, R.L. and Strauss, J.H. (1985) Science, 229, 726-733.

Personnel Supported from September 1, 1984 thru August 31, 1985

Personnel	Title	Period Supported by Contract
Dr. Vikram Vakharia	Research Associate	Sept. 1., 1984 - Feb. 1, 1985
Dr. Tazuko Yaegashi	Research Associate	Sept. 1, 1984 - Aug. 31, 1985
Mrs. Karin Page	Research Assistant	Oct. 10, 1984 - June 1, 1985
Dr. Suhas Kohlekar	Research Associate	Sept. 1, 1984 - Feb. 1, 1985
Baljit Kaur	Research Assistant	Sept. 1, 1984 - June 1, 1985
Dr. Radha K. Padmanabhan	Principal Investigator	Sept. 1, 1984 - present