Biochem. J. (1980) **185**, 65–75
Printed in Great Britain

The Nucleotide Sequences Recognized by Endonucleases *Ava*I and *Ava*II from *Anabaena variabilis*

Stephen G. HUGHES* and Kenneth MURRAY
Department of Molecular Biology, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, Scotland, U.K.

(*Received* 4 *June* 1979)

Determination of the 5′-terminal sequences flanking all the individual cleavage sites for endonuclease *Ava*I in bacteriophage-λ DNA has shown that this enzyme recognizes the hexanucleotide sequence:

5′-C↓-Y-C-G-R-G-3′
G-R-G-C-Y-C↑

This sequence is cut as shown by the arrows to give single-stranded 5′-tetranucleotide protrusions (cohesive ends). Endonucleases *Sma*I, *Xho*I and *Xma*I recognize different symmetrical subsets of this sequence and provide independent evidence for the occurrence of these subsets at particular endonuclease-*Ava*I cleavage sites in the bacteriophage-λ genome. Further evidence for this structure came from the demonstration that DNA fragments generated by endonuclease *Ava*I can be ligated to form a discrete set of larger molecules and from nearest-neighbour analysis which showed that cytosine residues occurred at the 3′-side of cleavage points. The observation that endonuclease *Ava*II recognized a subset of the sites recognized by *Asu*I [Hughes, Bruce & Murray (1979) *Biochem. J.* **185**, 59–63] led to the deduction that *Ava*II recognizes the pentanucleotide sequence:

5′-G↓-G-A-C-C-3′
C-C-T-G-G↑

and breaks internucleotide bonds at the positions indicated by the arrows.

The endonucleases *Ava*I and *Ava*II from the blue-green alga (cyanobacterium) *Anabaena variabilis* cut double-stranded DNA at specific sites. The 5′-terminal sequences of fragments generated by endonuclease *Ava*I are pC-C-G or pT-C-G and those of fragments generated by endonuclease *Ava*II are pG-A-C and pG-T-C. From these and other results it was inferred that both enzymes recognized degenerate hexanucleotide sequences, and it was stated, incorrectly as will be shown here, that neither enzyme produced DNA fragments with cohesive ends (Murray *et al.*, 1976). This analysis was followed by the location of the eight cleavage sites for endonuclease *Ava*I in the chromosome of bacteriophage λ (Hughes, 1977*a*; Rosenvold & Honigman, 1977).

From the structure proposed for the recognition target for endonuclease *Ava*I (Murray *et al.*, 1976) it was expected that some of the cleavage sites in bacteriophage-λ DNA should also be cut by endonucleases *Bsu* or *Hae*III, which break the central bond of the sequence -G-G-C-C- (Bron & Murray, 1975). However, this proved not to be the case, and so the target sequence for endonuclease *Ava*I was investigated in more detail, beginning with the determination of both 5′-terminal oligonucleotide sequences generated by the action of the enzyme at each of the eight cleavage sites in the chromosome of bacteriophage λ. These results led to the conclusion that the recognition target for endonuclease *Ava*I is a rotationally symmetrical but degenerate hexanucleotide sequence in which the enzyme makes staggered breaks, 4 base-pairs apart, to leave 5′ single-stranded projections terminated in a pyrimidine. This structure was reinforced by the results of an analysis of the products of digestion of bacteriophage-λ DNA with mixtures of endonucleases *Ava*I and *Xma*I, and *Ava*I and *Xho*I. It was further confirmed by demonstration that the fragments of bacteriophage-λ DNA generated by endonuclease *Ava*I could be joined together again by the action of DNA ligase under conditions requiring the joining of cohesive ends, and by nearest-

* Present address: Unilever Research, Colworth Laboratory, Colworth House, Sharnbrook, Beds. MK44 1LQ, U.K.

0306-3275/80/010065-11 $1.50/1

neighbour analysis on the products of such reactions to determine the 3'-nucleotide of the DNA fragments released by the action of endonuclease *Ava*I.

Electrophoretic analysis showed that the digestion products of several DNA substrates with endonucleases *Ava*II (Murray *et al.*, 1976) and *Asu*I (Hughes *et al.*, 1979) contained many components in common. It was deduced from experiments in which DNA samples were digested with the two endonucleases that endonuclease *Ava*II cuts one subset of the sites cut by endonuclease *Asu*I, and that degeneracy recognized by *Ava*II is limited to adenine or thymine. This result has been reported previously (Hughes, 1977*b*), but the supporting data were not presented. Identical conclusions on the recognition targets for endonucleases *Ava*I and *Ava*II have been reached independently and by different means by Sutcliffe & Church (1978).

Materials and Methods

Conventions and nomenclature

In the nomenclature of restriction endonucleases the acronymic system of Smith & Nathans (1973) has been followed.

The map of the wild-type bacteriophage-λ chromosome is presented in the linear form as described by Davidson & Szybalski (1971). Cleavage sites for individual restriction endonucleases are numbered consecutively from the left and carry the appropriate suffix according to the convention proposed by Arber & Linn (1969), e.g. *savaI* and *shindIII* refer to sites cut by endonucleases *Ava*I and *Hin*dIII respectively.

Nucleotide sequences are written in the direction 5' to 3', and where both strands of the DNA duplex are shown this orientation is retained in the upper line.

DNA fragments in restriction-endonuclease digests are labelled alphabetically in the upper case according to size (i.e. inversely related to electrophoretic mobility in gels) and are given a lower-case prefix to denote the restriction endonuclease (in addition to *Ava*I) which was used to generate them, *r* referring to endonuclease *Eco*RI, *b* to endonuclease *Bam*I, *h* to endonuclease *Hin*dIII, and *rh* to *Hin*dIII and *Eco*RI. Fragments obtained with endonuclease *Ava*I alone have no prefix.

Enzymes and digestion conditions

All the restriction endonucleases used were from the laboratory collection and had been prepared by methods essentially as described in the references indicated for the enzymes as follows: endonuclease *Eco*RI (Yoshimori, 1971); endonuclease *Hin*dIII (Old *et al.*, 1975); endonucleases *Ava*I and *Ava*II (Murray *et al.*, 1976); *Xma*I (Endow & Roberts, 1977); *Xho*I (Gingeras *et al.*, 1978); *Bsu* (Bron & Murray, 1975); *Bam*HI (Wilson & Young, 1975). The amount of each enzyme required to digest 1 μg of bacteriophage-λ DNA was determined in a series of trial digestions with increasing quantities of the enzyme. All digestions were carried out in the presence of 100mM-NaCl/20mM-Tris/HCl (pH7.5)/10mM-$MgCl_2$/20mM-2-mercaptoethanol, except for digestions with endonuclease *Hin*dIII where the concentration of NaCl was 50mM.

DNA ligase was isolated from cells of *Escherichia coli* infected with a derivative of bacteriophage λ carrying a segment of phage T4 which encodes the T4 DNA ligase (Murray *et al.*, 1979). Polynucleotide kinase was isolated and used in labelling 5'-termini of DNA molecules as described by Richardson (1965). Venom phosphodiesterase and pancreatic deoxyribonuclease were purchased from Worthington Corp., Freehold, NJ, U.S.A. Bacterial alkaline phosphatase was from Whatman Biochemicals, Maidstone, Kent, U.K. Calf intestinal phosphatase used in some experiments was purchased from Boehringer/Mannheim A.G.

DNA isolation

DNA from bacteriophage λ carrying the *cI*857 and *Sam*7 mutations was isolated as described previously (Murray *et al.*, 1976).

Electrophoretic methods

Methods for separation of DNA fragments by zone electrophoresis on agarose gels and for the separation and identification of radioactively labelled nucleotides on ion-exchange papers have been described in detail (Sharp *et al.*, 1973; Thompson *et al.*, 1974; Murray, 1973; Murray *et al.*, 1976). For radioautography, agarose gels were dried down slowly on glass plates in a hot-air oven before being exposed to X-ray film by direct contact. Zones of the gel containing radioactivity were then circumcised with a razor blade, and the piece of dried gel was peeled from the glass plate for further analysis.

5'-Terminal labelling of DNA fragments, and nucleotide sequence determination

The methods used for dephosphorylation of DNA fragments, labelling with ^{32}P in reactions with polynucleotide kinase, and digestion and analysis of the products were as described by Murray (1973) and Murray *et al.* (1976), and relevant details are given in the legend to Fig. 2.

Results and Discussion

Determination of 5'-nucleotide sequences at each of the eight cleavage sites for endonuclease AvaI in bacteriophage-λ DNA

From the maps of the positions of cleavage sites for endonucleases *Ava*I (Hughes, 1977*a*), *Bam*I (Haggerty & Schleif, 1976; Perricaudet & Tiollais,

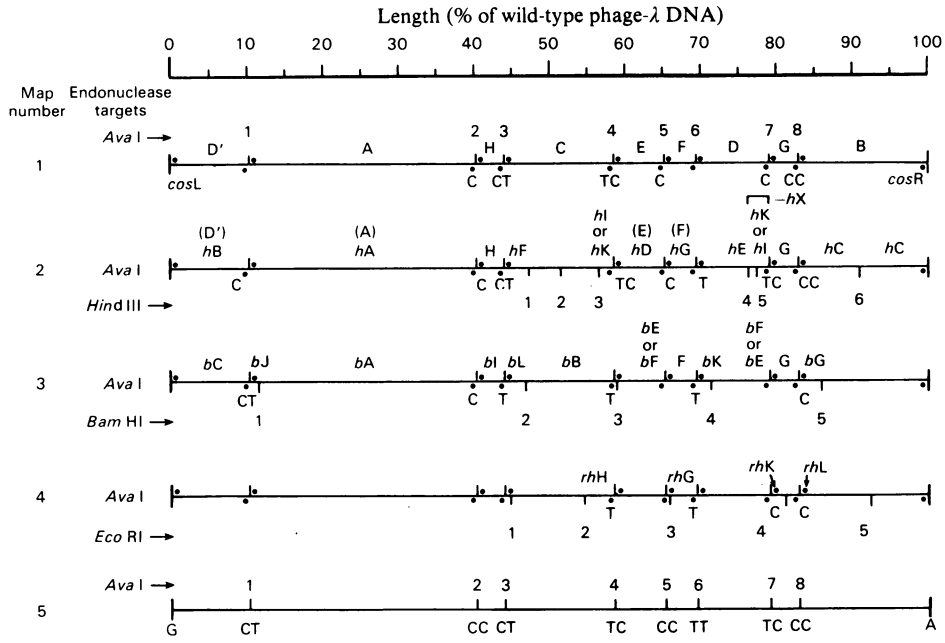


Fig. 1. *Positions in the bacteriophage-λ chromosome of targets for endonucleases AvaI, HindIII, BamHI and EcoRI and of DNA fragments produced in various endonuclease digests*

Targets are located by short vertical lines and the corresponding fragments are identified above the lines of the appropriate maps. The origin and assignment of the DNA fragments are shown in Fig. 2 and their nomenclature is described in the Methods and Materials section. *cos*L and *cos*R refer to the natural 5′ cohesive ends (left and right respectively) of bacteriophage-λ DNA. The labelled 5′-termini of the DNA fragments are denoted by dots, and terminal nucleotides that have been identified unequivocally (Table 1) are shown beneath the lines on the respective maps. Cumulative results from DNA fragments in the endonuclease-*Ava*I digest (map 1) and the various double digests (maps 2, 3 and 4 and Table 1) are summarized in map 5. Terminal sequences for each endonuclease-*Ava*I target are given in Fig. 6.

1975), *Hin*dIII (Murray & Murray, 1975) and *Eco*RI (Thomas & Davis, 1975) most of the fragments obtained by digestion of bacteriophage-λ DNA with pairs of these enzymes can be identified unambiguously. Thus digestion of an endonuclease *Ava*I digest of phage-λ DNA with endonuclease *Bam*I or *Hin*dIII or *Eco*RI generated fragments from known positions on the map, each of which carried only one terminus from a defined cleavage site for endonuclease *Ava*I (Fig. 1), For determination of the nucleotide sequences at each of these termini, the endonuclease-*Ava*I fragments were dephosphorylated and labelled with ^{32}P at their 5′-ends by treatment with polynucleotide kinase and [γ-^{32}P]ATP before digestion with the second endonuclease. Separation of the labelled fragments by electrophoresis in an agarose gel is illustrated in Fig. 2, and the analyses of their 5′-terminal (*Ava*I) sequences are summarized in Table 1.

Assignment of the labelled DNA fragments in the double digests to positions on the map of the phage-λ chromosome was made principally on the basis of their size, which was deduced from their electrophoretic mobilities. Ambiguities arose in placing some of the smaller fragments because their size was such that they could be allotted to more than one position (e.g. either fragment *b*K or *b*L could be placed at *saval*3 or *saval*6 on this basis). In these instances, however, the information derived from another double digest permitted the determination of the 5′-terminal nucleotides flanking the intervals within which such fragments lay, as explained in the legend to Fig. 2. In this way a composite map of the distribution of 5′-thymine and 5′-cytosine residues at the individual targets for endonuclease *Ava*I was built up as shown in Fig. 1 (map 5).

More extensive 5′-terminal sequences were determined from pancreatic deoxyribonuclease digests of the radioactively labelled fragments. The resultant array of oligonucleotides was separated and analysed by electrophoresis on Whatman AE81 paper at pH 3.5 (Fig. 3 and Table 2). Four series of related

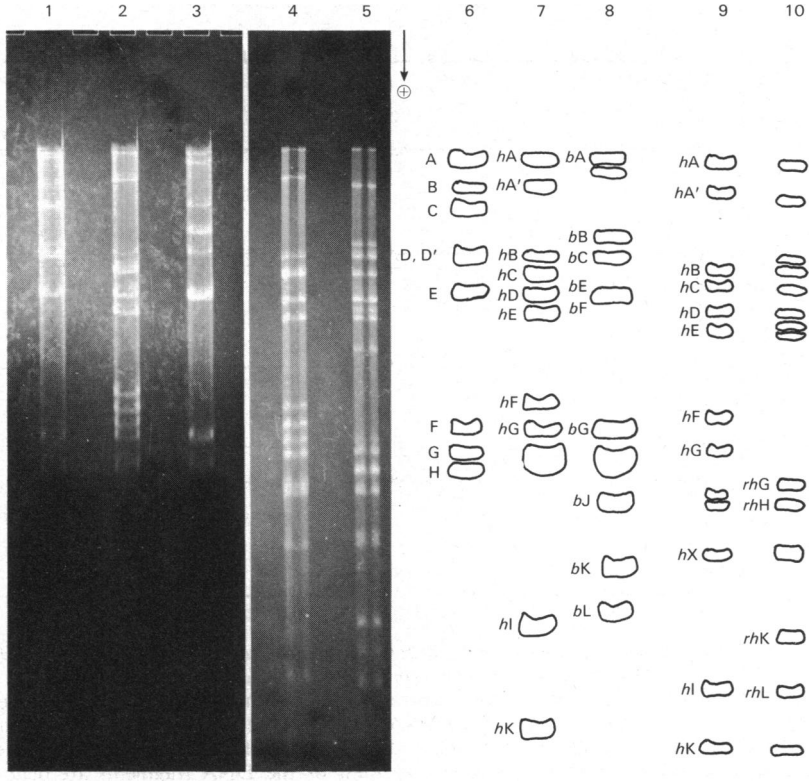


Fig. 2. *Electrophoresis in agarose gels of digests of bacteriophage-λ DNA with endonuclease AvaI and digests with mixtures of endonucleases*

Bacteriophage-λ DNA (100μg) was restricted with endonuclease *Ava*I and the reaction stopped by heating at 70°C for 5min. 5′-Terminal phosphate groups were removed with bacterial alkaline phosphatase and replaced with radioactive phosphates from [γ-^{32}P]ATP in reactions with polynucleotide kinase as described previously (Murray *et al.*, 1976). The labelled DNA fragments were dialysed against 0.3M-NaCl for 20h to remove the bulk of unchanged [^{32}P]ATP and free orthophosphate. A part of the material (about one-tenth) was then purified further by passage through a Sephadex G-100 column (50cm×2cm diameter) in 0.3mM-NaCl. (This procedure was not followed with most of the material, since residual radioactive ATP was adequately separated from the labelled DNA fragments during subsequent electrophoresis in agarose gels.) The purified labelled material contained about 2000c.p.m./labelled 5′-terminus per μg of DNA. Samples of the labelled DNA (10000–20000c.p.m.) were then digested further with endonucleases *Bam*HI, *Hin*dIII or *Hin*dIII and *Eco*RI together, and the digests were analysed by electrophoresis on 1% agarose gels as described in the Materials and Methods section. Gels were stained with ethidium bromide and photographed (tracks 1–5), then dried and radioautographed. The radioautograph was traced and the fragments were numbered with the nomenclature described in the Materials and Methods section (tracks 6–10). The digests in the various tracks are as follows: tracks 1 and 6, *Ava*I; tracks 2 and 7, *Ava*I and *Hin*dIII; tracks 3 and 8, *Ava*I and *Bam*HI; tracks 4 and 9, *Ava*I and *Hin*dIII; tracks 5 and 10, *Ava*I, *Hin*dIII and *Eco*RI. Sections of the dried gel containing radioactive fragments were circumcised and incubated at 37°C in 100μl of a reaction mixture containing venom phosphodiesterase (0.1mg/ml) and pancreatic deoxyribonuclease (0.5mg/ml) in 10mM-$MgCl_2$/20mM-Tris/HCl, pH7.5. Nucleotides recovered from the gel were applied to pennants of Whatman DE81 paper and eluted with 30% triethylamine carbonate, evaporated to dryness and redissolved in water (three times) before being analysed by electrophoresis on Whatman AE81 paper at pH3.5. Labelled 5′-mononucleotides were identified by their mobilities relative to nucleotide standards and are listed in Table 1. The sizes of individual DNA fragments on the gel were determined from their mobilities relative to fragments of known size (*Ava*I fragments and *Ava*I/*Eco*RI fragments of phage-λ DNA; Hughes, 1977*a*). Fragments were assigned to positions on the bacteriophage-λ DNA map (Fig. 1) on the basis of their size. However, some ambiguities arose from doublet bands on the agarose gels and in some instances two or more small fragments could be assigned equally well to more than one position on the map. It was not necessary to resolve all the ambiguities completely because, as Fig. 1 shows, the overall distribution of terminal nucleotides was built up with information from double digests with endonucleases *Bam*I or *Hin*dIII, so that an ambiguity in one could usually be resolved with information from the other. For example, consider the two ends of endonuclease-*Ava*I fragment D. Either fragment *b*K or fragment *b*L (which have the same 5′-terminus but different extended terminal

oligonucleotide sequences were observed typified by those from fragments *b*B, *h*B, *h*C and *h*G, each of which contained a single labelled terminus and thus gave rise to a unique overlapping series of endonuclease-digestion products. It was then possible to identify the individual nucleotides within each series from their mobility differences (Murray, 1973). Thus fragment *b*B gave a family of nucleotides consistent with the terminal sequence pT-C-G-G-G, *h*B gave pC-C-G-(A/T)-G, *h*C gave pC-C-G-G-G- and *h*G gave pT-C-G-(A/T)-G. The ambiguity with regard to the internal adenine and thymine residues arises because these nucleotides give similar mobility shifts on Whatman AE81 paper. The terminal sequence of each individual fragment was determined in this way from the families of oligonucleotides (Tables 1, 2 and 3). The sequences C-C-G-G-G and T-C-G-A-G would result from cleavage of the self-complementary hexanucleotide sequences C-C-C-G-G-G and C-T-C-G-A-G-. When the various 5′-terminal sequences

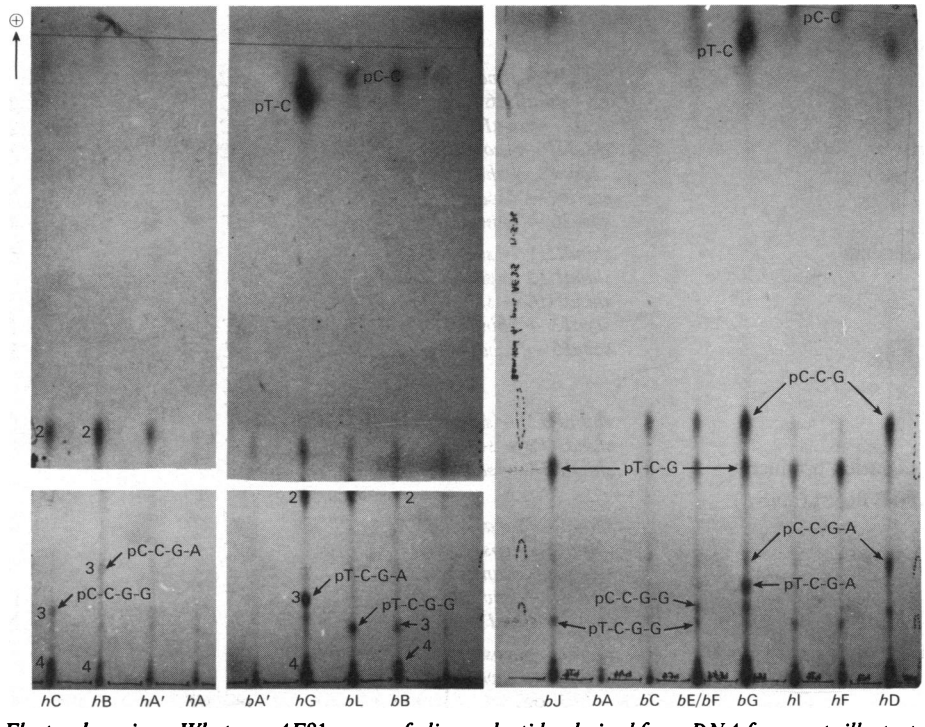


Fig. 3. *Electrophoresis on Whatman AE*81 *paper of oligonucleotides derived from DNA fragments illustrated in Fig.* 2
Radioactively labelled fragments were removed from excised zones of the dried agarose gel (Fig. 2) and incubated overnight at 37°C with pancreatic deoxyribonuclease [0.1 mg/ml in 0.02M-Tris/HCl (pH 7.5)/ 5mM-$MgCl_2$; 100μl]. Nucleotides leached from the gel were then applied directly to pennants of Whatman DE81 paper from which they were eluted with 30% triethylamine carbonate. The eluted material was evaporated to dryness and redissolved in distilled water (three times) before application to AE81 paper for electrophoresis at pH 3.5 until the blue marker dye, Xylene Cyanol FF, had moved 45cm from the origin. The paper was dried and radioautographed and nucleotides were identified by their mobilities and *M* values (Murray, 1973) which are given in Table 2. The numbers refer to the overlapping nucleotides in the sequence family to which the DNA fragment belongs.

sequences) comes from the left-hand end of fragment D (map 3), but this ambiguity is resolved by *h*E (map 2) which can be assigned to this position and belongs to the same terminal sequence family as fragment *b*K. This places fragment *b*K (terminal sequence family 4; the same as fragment *h*E) at *saval*6 (the left-hand end of fragment D), and fragment *b*L (terminal sequence family 1) must therefore constitute the right-hand side of *saval*3 (the left-hand end of endonuclease-*Ava*I fragment C). Fragment *h*I or fragment *h*K comes from the right-hand end of fragment D, but since both give the same 5′-terminus (T) and the same extended terminal sequence family (1), this ambiguity is irrelevant. Similarly, the ambiguity in placing fragment *b*E or *b*F at *saval*5 or *saval*7 is unimportant, because the sequences at these sites are defined from fragments in other digests.

Table 1. *Identification of DNA fragments separated by gel electrophoresis (Fig. 2) after digestion of bacteriophage-λ DNA with endonuclease AvaI and with combinations of endonucleases AvaI with HindIII, BamHI and EcoRI*
Location of the fragments on the phage-λ chromosome is shown on the maps in Fig. 1.

Fragment	Map interval (Fig. 1)	5′-Terminus	5′-Terminal sequence family (Table 2)
R.*Ava*I digest (map 1)			
A		pC, pT	
B		pC	
C		pT	
DD′		pC, pT	
E		pC	
F		pC, pT	
G		pC	
H		pC	
R.*Ava*I+R.*Hin*dIII digest (map 2)			
*h*A (A)	*sava*I1 → *sava*I2	pC, pT	
*h*A′	*h*B+*shindIII*6 → *cos*R	pC	2
*h*B (D′)	*cos*L → *sava*I1	pC	2
*h*C (doublet)	*shinIII*6 → *cos*R and *sava*I8 → *shindIII*6	pC	3
*h*D (E)	*sava*I4 → *sava*I5	pC	2, 3
*h*E	*sava*I6 → *shindIII*4	pT	4
Unlabelled fragments	*shindIII*1 → *shindIII*2 *shindIII*2 → *shindIII*3 *shindIII*4 → *shindIII*5		
*h*F	*sava*I3 → *shindIII*1	pT	1
*h*G (F)	*sava*I5 → *sava*I6	pC, pT	3, 4
(G)		pC	
(H)		pC	
*h*I	*shindIII*3 → *sava*I4	pT	1
*h*K	*shindIII*5 → *sava*I7	pT	1
*h*X (partial-digestion product)	*shindIII*4 → *sava*I7	pT	
R.*Ava*I+R.*Bam*HI digest (map 3)			
*b*A	*sbamI* → *sava*I2	pC	
*b*A′	*sbamI*5 → *cos*R + *cos*L → *sbam*I1	pC	
*b*B	*sbamI*2 → *sava*I4	pT	1
*b*C	*cos*L → *sava*I1	pC	2
Unlabelled fragment	*sbamI*5 → *cos*R		
*b*E, *b*F	*sbamI*3 → *sava*I5 or *sbamI*4 → *sava*I7	pC, pT	1, 3
*b*G (doublet) including F	*sava*I5 → *sava*I6 *sava*I8 → *sbamI*5	pC, pT	3, 4 4
*b*H (G)		pC	
*b*I (H)		pC	
*b*J	*sava*I1 → *sbamI*1	pT	1
*b*K	*sava*I3 → *sbam*2	pT	4
*b*L	*sava*I6 → *sbamI*4	pT	1
R.*Ava*I+R.*Eco*RI+R.*Hin*dIII *digest* (map 4)			
*rh*G, *rh*H	*sr*I3 → *sava*I6 *sr*I2 → *sava*I4	pT	
*rh*K	*sava*I7 → *sr*I4	pC	
*rh*L	*sr*I4 → *sava*I8	pC	

were assigned to their positions on the map (Figs. 1 and 6), these two sequences occurred only at those sites where either a cytosine or a thymine residue occurred at both 5′-termini resulting from cleavage with endonuclease *Ava*I. Furthermore, the sequences T-C-G-G-G- and C-C-G-(A/T)-G occurred at those sites that gave fragments with a 5′-thymine residue from one side of the break and fragments with a

Table 2. *5'-Terminal oligonucleotides derived by digestion of some of the 5'-terminally labelled DNA fragments (Table 1) with pancreatic deoxyribonuclease*

The electrophoretograms are illustrated in Fig. 3. Mobilities are with reference to the blue dye Xylene Cyanol FF.

DNA fragment	Sequence family	Nucleotide no.	Mobility	*M* value	Identity of oligonucleotide
*b*B	1	1	0.980	—	pT-C
		2	0.309	2.2	pT-C-G
		3	0.083	2.7	pT-C-G-G
		4	0.025	2.3	pT-C-G-G-G
*h*B	2	1	0.930	—	pC-C
		2	0.389	1.4	pC-C-G
		3	0.180	1.2	pC-C-G-(A/T)
		4	0.045	3.0	pC-C-G-(A/T)-G
*h*C	3	1	0.930	—	pC-C
		2	0.380	1.4	pC-C-G
		3	0.109	2.5	pC-C-G-G
		4	0.030	2.6	pC-C-G-G-G
*h*G	4	1	0.980	—	pT-C
		2	0.300	1.9	pT-C-G
		3	0.130	1.3	pT-C-G-(A/T)
		4	0.040	2.3	pT-C-G-(A/T)-G

Table 3. *Nearest-neighbour analyses to show the 3'-terminal nucleotides resulting from breakage of bacteriophage-λ DNA with endonuclease AvaI*

Bacteriophage-λ DNA (10 μg) was digested to completion with endonuclease *AvaI* in a volume of 120 μl, and the reaction stopped by heating at 70°C for 6 min. Polynucleotide ligase (1 μl) (isolated from *E. coli* infected with a derivative of phage λ carrying the T4 ligase gene; Murray *et al.*, 1979) and 0.2 μl of 0.1 M-ATP were then added and the digest was incubated at 37°C for 1 h. The intestinal phosphatase (Efstratiadis *et al.*, 1977) (1 μl) was added and incubation continued at 37°C for 15 min. The phosphatase was inactivated by heating at 70°C for 5 min. [γ-^{32}P]ATP (Amersham; 0.03 nmol; specific radioactivity 3000 Ci/mmol), unlabelled ATP (0.3 μmol) and 5 μl of polynucleotide kinase were then added to the reaction, and incubation was continued at 37°C for 2 h. Then 5 μl of additional polynucleotide ligase and 0.1 μl of 0.1 M-ATP were added and incubation was continued at 10°C for 40 h. Bacterial alkaline phosphatase (Worthington; 3 μl) was then added and the reaction mixture heated at 70°C for 10 min to remove 5'-terminal label from unligated fragments. This reaction was terminated by extraction with phenol, and unincorporated and released radioactive isotope were removed from the DNA by gel filtration on a column (500 mm×8 mm diameter) of Sephadex G-50 in 0.3 mM-NaCl. The excluded fraction (DNA) was concentrated to a volume of 50 μl and divided into two fractions, A and B. Fraction A was redigested with endonuclease *AvaI*. Both fractions were then incubated overnight with spleen endo- and exo-nucleases (a gift from Dr. G. Bernardi, I.B.M., Université de Paris, Paris, France), and the resulting mixture of 3'-mononucleotides was separated by electrophoresis on Whatman AE81 paper at pH 3.5. The positions of each of the 3'-mononucleotides were determined by radioautography and their identities determined from their mobilities relative to the marker dye Xylene Cyanol FF (Gp and Ap are not resolved in this system). The corresponding radioactive spots were cut out from the electrophoretogram and counted for radioactivity in a liquid-scintillation spectrometer. The results in fraction A show an average non-specific incorporation of 330 c.p.m. of each mononucleotide. Deduction of this from the values in fraction B shows clearly that Cp is the predominant 3'-terminal nucleotide.

		Net radioactivity (c.p.m.)	
Nucleotide	Identity	Fraction A	Fraction B
1	Tp	363	411
2	Cp	288	1291
3	Gp+Ap	567	644

5'-cytosine residue from the other. Again, these two sequences could be base-paired to form a partially symmetrical hexanucleotide sequence in which the asymmetrical nucleotides lay adjacent to the cleavage points; such arrangements were found at endonuclease *AvaI* sites, 1, 3, 4 and 7. These results thus

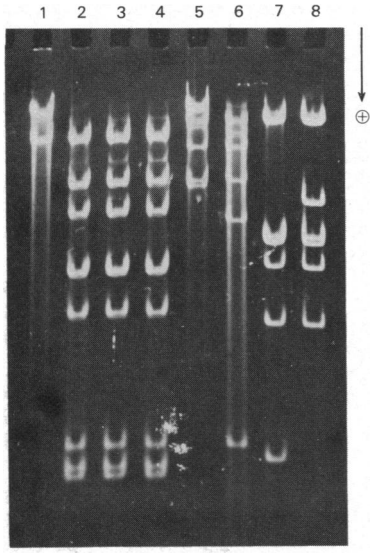


Fig. 4. *Co-digestion of bacteriophage-λ DNA with endonucleases XhoI and XmaI and endonuclease AvaI*
Digests were carried out as described in the Materials and Methods section and analysed by electrophoresis on a 1% agarose gel as follows. Tracks 1 and 2 contain endonuclease-*Xho*I digests, of which the sample in track 2 was also digested with endonuclease *Ava*I. Track 3 contains an endonuclease-*Ava*I digest. Tracks 4, 5 and 6 contain endonuclease-*Xma*I digests, of which that in track 5 was also digested with endonuclease *Ava*I, and that in track 6 with endonuclease *Xho*I. Tracks 7 and 8 contain an endonuclease-*Eco*RI digest, that in track 7 having been digested also with endonuclease *Xho*I. The smallest fragment in the latter digest is identical in size with that generated by cutting between sites *sava*I6 and *sr*I3 (fragment *rh*G in Figs. 1 and 2) and is clearly cut from the fragment lying between *sr*I3 and *sr*I4.

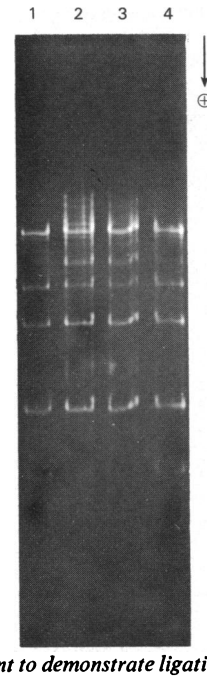


Fig. 5. *Experiment to demonstrate ligation of endonuclease-AvaI fragments of phage-λ DNA*
Bacteriophage λ C1857 *Sam*7 DNA (2.5 μg) was digested to completion with 10 μl of endonuclease *Ava*I in 50 mM-NaCl/10 mM-$MgCl_2$/20 mM-Tris/HCl, pH 7.5, in a volume of 28 μl. The reaction was stopped by heating at 70°C for 7 min and then 2.5 μl of 10× ligation buffer (Murray *et al.* 1979) was added. Four 6 μl fractions of the digest were taken: fraction 1 (track 1) received no further treatment; fraction 2 (track 2) was incubated with 0.5 μl of polynucleotide ligase+0.1 μl of 0.1 M-ATP for 16 h before being heated at 70°C for 5 min; fraction 3 (track 3) was treated identically with fraction 2 except for the addition of 2 μl of polynucleotide kinase; fraction 4 (track 4) was treated identically with fraction 3, except that it was redigested with 2 μ. of endonuclease *Ava*I after ligation.

indicate the generalized recognition sequences:

$$\begin{array}{l} \overset{\downarrow}{\text{C-Y-C-G-R-G-}} \\ \text{G-R-G-C-Y-C} \\ \qquad\qquad\quad \uparrow \end{array}$$

Products of co-digestion of bacteriophage-λ DNA with endonucleases XmaI or XhoI and AvaI

Endow & Roberts (1977) reported that the recognition sequence for endonuclease *Xma*I is C-C-C-G-G-G-, which is one subset of the generalized sequence proposed above for endonuclease *Ava*I. Fig. 4 shows that this sequence lies at *sava*I2, *sava*I5 and *sava*I8, and it was duly noted that the three cleavage sites for endonuclease *Xma*I in the bacteriophage-λ chromosome (MacParland *et al.*, 1976) are located at or very close to these sites. Co-digestion of bacteriophage-λ DNA with both these enzymes gave the same pattern of fragments as that obtained with endonuclease *Ava*I alone, which shows that the cleavage sites for this enzyme are in fact coincident with those for endonuclease *Xma*I. The DNA sequence recognized by endonuclease *Xho*I is C-T-C-G-A-G- (Gingeras *et al.*, 1978), which represents a second subset of the recognition sites proposed for endonuclease *Ava*I. Co-digestion of bacteriophage-λ DNA with endonucleases *Xho*I and *Ava*I (Fig. 4) confirmed that the single site for endonuclease *Xho*I in the bacteriophage-λ chromosome is in fact coincident with an endonuclease *Ava*I site. The products of co-digestion of bacteriophage-λ DNA with endonucleases *Eco*RI or *Hin*dIII with endonuclease *Xho*I are consistent with the single site

for *Xho*I in phage-λ DNA being coincident with *saval*5.

Together, these results support the proposed structure for the site recognized by endonuclease *Ava*I.

*Ligation of DNA fragments generated by endonuclease Ava*I

The 5′-terminal sequences of fragments of DNA generated by endonuclease *Ava*I are such that no single fragment should carry a pair of fully complementary termini and thus the individual fragments should not be able to circularize. Failure of single fragments to form circles was observed previously, and was used as an argument against the formation of cohesive ends on the digestion products (Murray *et al.*, 1976). However, it should be possible to observe new fragments generated by annealing and ligation of an endonuclease-*Ava*I digest of phage-λ DNA. Fig. 5, which gives the results of an experiment in which an endonuclease-*Ava*I digest of phage-λ DNA was incubated with DNA ligase at 10 °C, shows that this is indeed the case, and that such fragments give the standard digestion pattern for phage-λ DNA with endonuclease *Ava*I when the digest is reincubated with *Ava*I. [It had been shown previously (S. G. Hughes & K. Murray, unpublished work) that the DNA ligase preparation used did not effect the ligation of 'blunt-ended' DNA fragments, even at ligase concentrations 10 times higher than those used in these experiments.]

This result confirms that endonuclease *Ava*I cuts DNA to leave single-stranded projections ('cohesive ends') and offers an approach to determination of the nucleotide at the 3′-side of the points of cleavage via a nearest-neighbour experiment in which a radioactive 5′-phosphate is transferred to an adjacent 3′-hydroxy terminus by the action of polynucleotide ligase. Such an experiment was carried out with endonuclease-*Ava*I fragments of bacteriophage-λ DNA, the 5′-termini of which were labelled by using polynucleotide kinase (Table 3). In a preliminary experiment, a high degree of random incorporation of radioactivity was observed. It was concluded that this resulted from a high proportion of single-strand breaks in the substrate DNA, and for this reason the experiment reported in Table 3 included a ligation step at 37 °C after restriction-endonuclease digestion but before labelling with ^{32}P. It was reasoned that this should also seal up the natural cohesive ends of the phage-λ DNA molecule and minimize the contribution from this source to the background. The results show that Cp is the predominant labelled nucleotide and that the radioactivity in this nucleotide is reduced to the same background level as the other nucleotides by preincubation of the labelled religated DNA with endonuclease *Ava*I (Table 3, fraction A), showing that the additional radioactivity was in fact incorporated at the endonuclease-*Ava*I cleavage sites. This experiment does not rule out the possibility that other nucleotides could occur at the 3′-side of the cuts made by endonuclease *Ava*I. However, taken together, the large excess of radioactivity in Cp, the similar but lower amounts of radioactivity found in the other three nucleotides, and the insensitivity of this radioactivity to digestion with endonuclease *Ava*I (Table 3, fraction B) support the conclusion that Cp is the predominant 3′-nucleotide. This further substantiates the conclusion that endonuclease *Ava*I recognizes the sequence -C-Y-C-G-R-G-.

A map has thus been constructed showing the structure and location of each of the sites in phage-λ DNA recognized and cut by endonuclease *Ava*I (Fig. 6). Extensive sequence analysis by others has independently confirmed some of these allocations. The attachment region of bacteriophage-λ DNA contains the sequence C-C-C-G-A-G- at position 272–277 (Davies *et al.*, 1977), which is in good agreement with Fig. 6 where this sequence is shown in the same orientation at *saval*4 adjacent to the *att* region. In addition, the sequence C-C-C-G-A-G- lies at position 174–179 within the *Cro* gene of bacteriophage λ (Roberts *et al.*, 1977) in agreement with the position, sequence and orientation of *saval*7 shown in Fig. 6.

Recognition site of endonuclease AvaII

The preliminary comparison of endonuclease-*Asu*I and endonuclease-*Ava*II digests of a number of DNA substrates showed a large proportion of fragments of the same size, especially among the smaller frag-

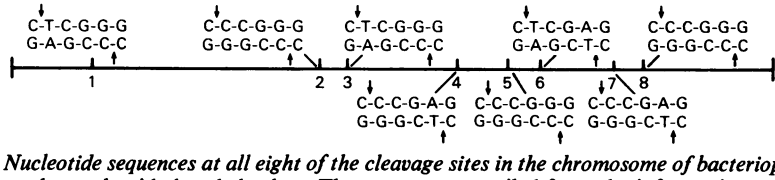


Fig. 6. *Nucleotide sequences at all eight of the cleavage sites in the chromosome of bacteriophage λ*
The arrows indicate the nucleotide bonds broken. The map was compiled from the information presented in Tables 1 and 2 and Figs. 2 and 3.

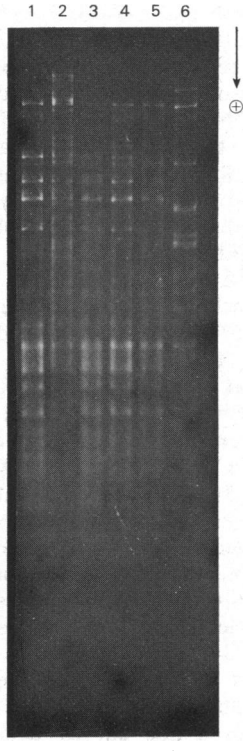
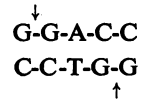


Fig. 7. *Digests and co-digests of bacteriophage-λ DNA with endonucleases AvaII and AsuI*

All reactions were carried out at 37°C for 2h in the presence of 10mM-$MgCl_2$/20mM-Tris/HCl (pH7.5)/100mM-NaCl and contained 1 μg of bacteriophage-λ DNA. Track 1, reaction carried out with the amount of *Asu*I (20 μl) required for complete digestion; track 2, with one-quarter of that amount of *Asu*I (5 μl); track 3, with 5 μl of endonuclease *Ava*II+20 μl of *Asu*I; track 4, with 5 μl of endonuclease *Ava*II+10 μl of *Asu*I; track 5, with 5 μl of endonuclease *Ava*II +5 μl of *Asu*I; track 6, with 5 μl of endonuclease *Ava*II.

ments. Further analysis with bacteriophage-λ DNA (Fig. 7) showed that a co-digest with endonucleases *Asu*I and *Ava*II appeared the same as an endonuclease-*Asu*I digest and that a partial *Asu*I digest could be driven almost to completion by addition of endonuclease *Ava*II. Since endonuclease *Asu*I has been shown to recognize the sequence G-G-N-C-C- leaving 5′-termini pG-N-C-C- (Hughes *et al.*, 1979), it is clear that endonuclease *Ava*II recognizes a subset of these sequences. The previous analysis showed that only the terminal dinucleotides G-A and G-T were obtained with endonuclease *Ava*II (Murray *et al.*, 1976), and thus endonuclease *Ava*II recognizes the sequence shown above for *Asu*I in which N is limited to A or T, giving the structure:



In a previous analysis of the sequence recognized and cleaved by endonuclease *Ava*II (Murray *et al.* 1976) sequences were found in addition to the families of nucleotides related to G-A-C-C- and G-T-C-C- in which the second cytosine was replaced by any of the four nucleotides. This led to the conclusion that the sequence became degenerate beyond the first cytosine. However, re-examination of the electrophoretograms shown in Murray *et al.* (1976) (Plate 4) reveals that the G-A-C-C- and G-T-C-C- families gave by far the strongest spots, and that the degenerate series formed a minor component of the nucleotides observed, which reconciles the previous observations with the sequence deduced here. It is possible that the minor set of degenerate nucleotides observed previously resulted from a relaxation of the specificity of endonuclease *Ava*II analogous to that observed for endonuclease *Eco*RI (*Eco*RI*) (Polisky *et al.*, 1975) and *Bsu* (Heininger *et al.*, 1977). Alternatively, the batch of enzyme used in the initial experiments may have been contaminated with a small amount of an additional endonuclease.

The relationship between the recognition properties of endonucleases *Asu*I and *Ava*II may be of interest to those concerned with the evolution and mechanisms of restriction endonucleases. Further study of the endonucleases of other *Anabaena* species may reveal a series of further endonucleases of related specificity such as have been observed among species of *Xanthomonas* (Roberts, 1976). This would provide additional evidence for a proposed role of restriction endonucleases as prokaryotic genetic isolating mechanisms which act to stabilize divergence during speciation.

We thank Miss Karen Brown, Mrs. Sandra Bruce and Mrs. Jo Rennie for their expert assistance in this work, which was supported by a grant from the Science Research Council.

References

Arber, W. & Linn, S. (1969) *Annu. Rev. Biochem.* **38**, 467–500

Bron, S. & Murray, K. (1975) *Mol. Gen. Genet.* **143**, 25–33

Davidson, N. & Szybalski, W. (1971) in *The Bacteriophage Lambda* (Hershey, A. D., ed.), p. 74, Cold Spring Harbor Laboratory, New York

Davies, R. W., Schreier, P. H. & Büchel, D. E. (1977) *Nature (London)* **270**, 757–760

Efstratiadis, A., Vournakis, J. N., Donis-Keller, H., Chaconas, G., Dougall, D. K. & Kafatos, F. C. (1977) *Nucleic Acids Res.* **4**, 4165–4174

Endow, S. & Roberts, R. J. (1977) *J. Mol. Biol.* **112**, 521–529

Gingeras, T. R., Myers, P. A., Olson, J. A., Hanberg, F. A. & Roberts, R. J. (1978) *J. Mol. Biol.* **118**, 113–122
Haggerty, D. M. & Schleif, R. E. (1976) *J. Virol.* **18**, 659–663
Heininger, J., Hörz, W. & Zachau, H. G. (1977) *Gene* **1**, 291–303
Hughes, S. G. (1977*a*) *Biochem. J.* **163**, 503–509
Hughes, S. G. (1977*b*) Ph.D. Thesis, University of Edinburgh
Hughes, S. G., Bruce, T. & Murray, K. (1979) *Biochem. J.* **185**, 59–63
MacParland, R. H., Brown, L. R. & Pearson, G. D. (1976) *J. Virol.* **19**, 1006–1011
Murray, K. (1973) *Biochem. J.* **131**, 569–583
Murray, K. & Murray, N. E. (1975) *J. Mol. Biol.* **98**, 551–564
Murray, K., Hughes, S. G., Brown, J. S. & Bruce, S. A. (1976) *Biochem. J.* **159**, 317–322
Murray, N. E., Bruce, S. A. & Murray, K. (1979) *J. Mol. Biol.* in the press
Old, R. W., Murray, K. & Roizes, G. (1975) *J. Mol. Biol.* **92**, 331–339
Perricaudet, M. & Tiollais, P. (1975) *FEBS Lett.* **56**, 7–11
Polisky, B., Greene, P., Garfin, D. E., McCarthy, B. J., Goodman, H. M. & Boyer, H. W. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3310–3314
Richardson, C. C. (1965) *Proc. Natl. Acad. Sci. U.S.A.* **54**, 158–165
Roberts, R. J. (1976) *Crit. Rev. Biochem.* **4**, 122–164
Roberts, T. M., Shimatake, H., Brady, C. & Rosenberg, M. (1977) *Nature (London)* **270**, 274–275
Rosenvold, E. C. & Honigman, A. (1977) *Gene* **2**, 273–278
Sharp, P. A., Sugden, W. & Sambrook, J. (1973) *Biochemistry* **12**, 3055–3063
Smith, H. O. & Nathans, D. (1973) *J. Mol. Biol.* **81**, 419–423
Sutcliffe, J. G. & Church, G. M. (1978) *Nucleic Acids Res.* **5**, 2313–2319
Thomas, M. & Davis, R. W. (1975) *J. Mol. Biol.* **91**, 315–328
Thompson, R., Hughes, S. G. & Broda, P. M. A. (1974) *Mol. Gen. Genet.* **133**, 141–149
Wilson, G. A. & Young, F. E. (1975) *J. Mol. Biol.* **97**, 123–126
Yoshimori, R. (1971) Ph.D. Thesis, University of California