Volume 10 Number 3 1982 Nucleic Acids Research

Modification profiles of bacterial genomes

Joan E.Brooks and Richard J.Roberts

P.O.Box 100, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

Received 27 October 1981; Revised 16 December 1981; Accepted 21 December 1981

ABSTRACT

DNAs were prepared from twenty-six bacterial species and digested with a variety of restriction endonucleases to determine what modifications the DNAs carry. Several general conclusions could be made: 1) First, in no instance was the DNA of a restriction enzyme strain cleaved by its own restriction enzyme. 2) The specificity of the DNA modification was the same as that of its restriction counterpart; there were no cases of the DNAs being modified against a less specific class of restriction enzymes. 3) In most (but not all) cases, the resistance of a bacterium's DNA to its own restriction enzyme could be generalized to include resistance to all other restriction enzymes with the same specificity (isoschizomers). 4) DNA modified within the central tetramer of a recognition sequence is usually protected against cleavage by all related hexameric enzymes possessing that central tetramer. Only three families of DNA presented in this study disobey this rule. 5) Finally, a significant number of cases emerge where bacterial DNA carries a modification but no corresponding restriction endonuclease activity.

INTRODUCTION

Much is still unknown about the mechanisms used by bacteria to protect their own DNAs from the action of endogenous restriction enzymes. In all systems that have been studied, cells produce a modification methylase in addition to the restriction endonuclease; both enzymes recognize the same specific DNA sequence (1-3). Restriction systems fall into three classes based on structural and biochemical characteristics (3, 4). Type I restriction enzymes require divalent cations, ATP and S-adenosyl methionine (AdoMet) and cut the DNA a random distance from the recognition site. Type II restriction enzymes require only divalent cations for activity and cut at fixed points relative to (usually within) the recognition sequence. Type III restriction enzymes require divalent cations and ATP for activity, are stimulated by (but don't require) AdoMet, and cut DNA at fixed points relative to the recognition sequence. The Type II enzymes have proven to be the most useful to molecular biologists. Over 250 of these enzymes have been characterized that, as

0305-1048/82/1003-0913$2.00/0

a group, possess more than 70 different recognition sequences (5). While these Type II enzymes have become extremely useful as research tools, relatively little is known of their role in the bacterium: few actually have been shown to "restrict" the entry of foreign DNA into the cell (6, 7). In some cases there is evidence against such a role (8).

The means by which bacteria protect their own DNA from their restriction enzymes have not been fully investigated. Certainly the best characterized means of protection is the specific methylation of an adenine or cytosine residue within the recognition sequence by a site-specific methylase. All known DNA methylases require AdoMet as the methyl donor; no other cofactors are required (3, 4, 9). Not many DNA methylases have been studied in detail, but all of those studied methylate either adenine to N^6methyladenine or cytosine to 5-methylcytosine (10, 11). Most methylases modify a single specific residue within the recognition sequence; recently, however, it has been reported that a methylase from Moraxella species can methylate both cytosine residues within its recognition sequence (12). There is a group of Type II enzymes having asymmetric recognition sites. These include HgaI (GAGGC), HphI (GGTGA), and MnlI (CCTC). It is not clear in these cases how methylation on one strand (which is different from the sequence on the other strand) can prevent the action of the restriction endonuclease. In these cases, one could easily envision the existence of an auxiliary protein to regulate activity of the restriction endonuclease. In fact, there are phage encoded proteins known which act to inhibit the activity of specific restriction enzymes on nonmodified phage DNAs (13, 14).

Aside from what biological role the restriction modification systems play within the bacteria, several other questions regarding the activity and interaction of the restriction and modification counterparts remain. For example, it is still not known whether all restriction enzymes are accompanied by a specific DNA-modifying enzyme. Nor is it known whether on the bacterial chromsome DNA modification always exists as methylation. In systems where DNA modification methylation does exist, is it always as specific as the restriction enzyme in recognition sequence? Or, for example, can a bacterium containing a restriction enzyme specific for a six-base sequence, contain a modification methylase that recognizes the central four bases of that sequence? To begin to answer some of these fundamental questions, chromosomal DNAs were prepared from several restriction enzyme producing strains of bacteria and examined for sensitivity to cleavage by a variety of restriction enzymes.

MATERIALS AND METHODS

Bacterial DNA preparations: All the bacterial DNAs used are found in Table I. Acinetobacter calcoaceticus, Anabaena variabilis, Providencia stuartii and Xanthomonas malvacearum DNAs were obtained from N.E. Biolabs; Moraxella species DNA was a gift from D. Levy. The remainder of the bacterial DNAs were purified by the method of Marmur (15).

Control DNA preparations: Adenovirus-2 DNA was purified by the method of Pettersson and Sambrook (16); λcI857Sam7 DNA was prepared by the method of Schrenk and Weisberg (17). pBR322 DNA was isolated from GM119 cells after chloramphenicol amplification of the plasmid (18).

Analysis of DNA: AccI, AccII, AvaI, AvaII, BamHI, BstNI, HincII, HindIII, HphI, MboI, MspI, PvuI, PvuII, SalI, SmaI, XbaI and XhoI were purchased from N.E. Biolabs. EcoRII, HpaI and DpnI enzymes were obtained from BRL. These enzymes were used with the buffers recommended by their respective manufacturers. The remainder of the enzymes listed in Table I were prepared in this laboratory. The buffer used in these digests contains 6 mM TRIS, pH 7.9, 6 mM $MgCl_2$ and 6 mM Mercaptoethanol. For all digests, 4 units of enzyme were used to digest 2 μg of DNA in a 50 μl reaction mixture. All reactions were incubated at 37°C except BclI, BstNI and TaqI, which were done at 68°C.

After 3 hr. incubations, the reactions were terminated by the addition of 0.1 volume 0.1 M EDTA and 0.1 volume loading dye containing 50% sucrose and 0.15% Bromphenol blue. The restriction digests were then loaded onto 1.0% agarose slab gels (20 cm x 20 cm x 0.3 cm) containing E buffer (40 mM TRIS-acetate, pH 7.8, 5 mM sodium acetate, 1 mM EDTA) with 0.5 μg/ml ethidium bromide added. Electrophoresis were carried out at 100V for 3 hrs; the gels were then photographed under short wave UV light.

RESULTS

The results of digesting a variety of bacterial DNAs with a range of Type II restriction endonucleases are given in Table I. In each assay, 2 μg of the DNA being tested was incubated with an excess of the restriction enzyme (see Materials and Methods). While the extent of digestion varied considerably with the DNA-enzyme pair being tested, a DNA was scored as being resistant to an enzyme only if no digestion of the substrate was detected after electrophoresis on a 1% agarose gel. In a few cases, contaminating nonspecific nuclease activity in the enzyme preparation led to a slight smearing of the DNA during electrophoresis, but not to banding. The smearing

was not detectable with the enzymes used under normal digest conditions, but only becomes apparent under conditions of excess enzyme and excess incubation times necessary to these experiments. Such smearing is clearly distinct from specific digestion (see Figure 1, slot #5).

DNAs are resistant to restriction enzymes for which they code.

In no case tested was the purified DNA of a bacterial strain that produces a restriction enzyme cleaved by that restriction enzyme (Table I). This is true even for enzyme systems having asymmetric recognition sites (i.e., HgaI GACGC; HphI GGTGA; and MboII GAAGA). In bacteria producing multiple restriction enzymes the DNA was resistant to cleavage by each of its endogenous enzymes. This observation also indicates that the DNA preparation being tested came, in fact, from the desired bacterium and not a contaminant. In these and other cases of resistance to cleavage, two control experiments were performed. First, each of the DNA preparations were shown to be cleaved by at least one restriction enzyme; second, a mixed

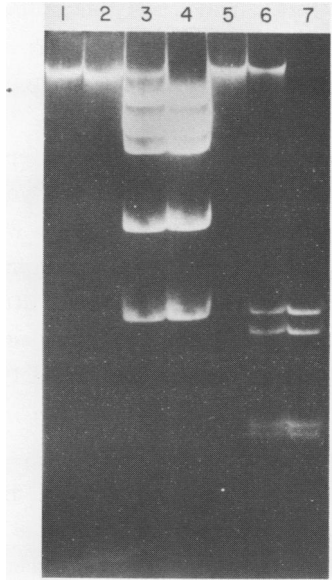

Figure 1: Control restriction digests of X. holcicola DNA.
X. holcicola DNA, either alone or in combination with Adenovirus-2 or pBR322 DNAs, was treated with XhoI or XhoII endonuclease and subjected to electrophoresis, as described in Materials and Methods. Lanes (1), (2) and (5) contain X. holcicola DNA; (3) contains Adenovirus-2 + X. holcicola DNAs; (4) Adenovirus-2 DNA; (6) pBR322 + X. holcicola DNAs and (7) pBR322 DNA. Lane (1) shows uncut X. holcicola DNA; (2-4) the DNAs are treated with XhoI; in lanes (5-7), XhoII enzyme.

digest was set up with the resistant DNA and a DNA known to be susceptible (either Adenovirus-2, λ or pBR322 DNA). Together, the two controls show that the observed resistance was due neither to inhibitors in the DNA preparation, nor to some general interfering modification of the DNA. An example of a set of control digests is given in Figure 1. One implication of these results is that although other means of controlling restriction enzyme activity may be present within the cell, they are not used in lieu of DNA level modification.

Specificities of the modifications.

The specificity of the DNA modification was found to parallel that of the associated restriction enzyme. In no case was the DNA modification less stringent: a strain producing an enzyme with a six-base specificity was never found to be totally resistant to an enzyme having a recognition specificity of just the central four bases. Similarly, bacterial DNA from strains producing restriction enzymes having a "relaxed core" were not resistant to enzymes recognizing a related but less specific four-base sequence (this is summarized in Table II).

DNAs resistant to cleavage by related restriction enzymes.

In most cases examined, a DNA that is resistant to one restriction enzyme is also resistant to isoschizomers of that enzyme. For example, DNA from A. calcoaceticus, which produces the AccII enzyme (recognizing CGCG), is also resistant to cleavage by BceR and FnuDII, two isoschizomers of AccII. Similarly, DNA from X. malvacearum, which codes for XmaI (CCCGGG) and XmaII (CTGCAG) cannot be cleaved by SmaI, an XmaI isoschizomer, nor by PstI and SflI, two isoschizomers of XmaII (see Table I). Furthermore, in most instances, modification of a DNA within the central tetramer of a recognition sequence will protect against all restriction enzymes having that central tetramer. The protection against cleavage is also found with DNAs containing restriction enzymes with "relaxed cores" when challenged by related nondegenerate hexameric enzymes. A summary of these results are given in Table III.

DNAs sensitive to cleavage by related restriction enzymes.

In Table IV there are three families of DNAs which show distinctive behavior -- the E. coli/E. cloacae pair, the M. species/H. parainfluenzae pair and the S. aureus/M. bovis/E. coli group. Each of these groups contain DNAs sensitive to digestion by some enzymes, but resistant to digestion by isoschizomers of those enzymes. Digests of E. cloacae and E. coli; M. species and H. parainfluenzae; M. bovis, S. aureus3A and E. coli DNAs are given in Figures 2, 3, and 4, respectively. Figure 2 shows that E. coli and E. cloacae

TABLE I

Modification Pofiles of Bacterial DNAs

Organism	Known Restriction System(s)[a]	Known Methylation Site(s)[a]	Cleaved by[b]	Resistant to[b]
Acinetobacter calcoaceticus	AccI GT($\frac{A}{C}$)($\frac{G}{T}$)AC AccII CGCG AccIII ?	none	TaqI TCGA	AccI GT($\frac{A}{C}$)($\frac{G}{T}$)AC SalI GTCGAC AccII CGCG BceR CGCG FnuDII CGCG SstII CCGCGG
Anabaena variabilis	AvaI CPyCGPuG AvaII GG($\frac{A}{T}$)CC AvaIII ATGCAT	none	HpaII CCGG MspI CCGG TaqI TCGA	AvaI CPyCGPuG SmaI CCCGGG XmaI CCCGGG XhoI CTCGAG AvaII GG($\frac{A}{T}$)CC

Organism	Known Restriction System(s)[a]	Known Methylation Site(s)[a]	Cleaved by[b]	Resistant to[b]
Arthrobacter luteus	AluI AGCT	none	MboI GATC Sau3A GATC	AluI AGCT HindIII AAGCTT PvuII CAGCTG SstI GAGCTC
Bacillus amyloliquefaciens H	BamHI GGATCC	M. BamI GGAT^{m}CC M. BamII G^{m}C($^{A}_{T}$)GC(19)	MboI GATC Sau3A GATC XhoII PuGATCPy BglII AGATCT PvuI CGATCG XorII CGATCG BclI TGATCA	BamHI GGATCC
Bacillus caldolyticus	BclI TGATCA	none	MboI GATC Sau3A GATC DpnI G^{m}ATC XhoII PuGATCPy BamHI GGATCC BglII AGATCT PvuI CGATCG XorII CGATCG	BclI TGATCA

Organism	Known Restriction System(s)[a]	Known Methylation Site(s)[a]	Cleaved by[b]	Resistant to[b]
Bacillus globigii	BglI $GCCN_5GGC$ BglII AGATCT	none	MboI GATC Sau3A GATC BamHI GGATCC PvuI CGATCG XorII CGATCG	BglI $GCCN_5GGC$ BglII AGATCT
Enterobacter cloacae	EclI ? EclII $CC(^{A}_{T})GG$	none	BstNI $CC(^{A}_{T})GG$ AacI $CC(^{A}_{T})GG$	EclII $CC(^{A}_{T})GG$ EcoRII $CC(^{A}_{T})GG$
Escherichia coli strain HB101	none	M. Eco dam $G^{m}ATC(20)$ M. Eco dcm $C^{m}C(^{A}_{T})GG(21)$	Sau3A GATC XhoII PuGATCPy BglII AGATCT BamHI GGATCC PvuI CGATCG XorII CGATCG EcoRII $CC(^{A}_{T})GG$ EclII $CC(^{A}_{T})GG$	MboI GATC BclI TGATCA BstNI $CC(^{A}_{T})GG$ AacI $CC(^{A}_{T})GG$

Organism	Known Restriction System(s)[a]	Known Methylation Site(s)[a]	Cleaved by[b]	Resistant to[b]
Haemophilus gallinarum	HgaI GACGC	none[c]	Sau3A GATC	HgaI GACGC MboI GATC[c]
Haemophilus influenzae Rc	HincII GTPyPuAC	M. HincII GTPyPu^{m}AC	TaqI TCGA	HincII GTPyPuAC HpaI GTTAAC SalI GTCGAC
Haemophilus parahaemolyticus	HphI GGTGA	none[c]	Sau3A GATC	HphI GGTGA MboI GATC[c]
Haemophilus parainfluenzae	HpaI GTTAAC HpaII CCGG	M. HpaI GTTA^{m}AC M. HpaII C^{m}CGG	MspI CCGG	HpaI GTTAAC HpaII CCGG NciI CC($^{C}_{G}$)GG SmaI CCCGGG XmaI CCCGGG
Klebsiella pneumoniae	KpnI GGTACC	none	RsaI GTAC	KpnI GGTACC

Organism	Known Restriction System(s)[a]	Known Methylation Site(s)[a]	Cleaved by[b]	Resistant to[b]
Moraxella bovis	MboI GATC	none	Sau3A GATC	MboI GATC
			DpnI G^{m}ATC	BclI TGATCA
	MboII GAAGA		XhoII PuGATCPy	BamHI GGATCC
				BglII AGATCT
				PvuI CGATCG
			HaeIII GGCC	XorII CGATCG
			FnuDII CGCG	
				MboII GAAGA
				SmaI CCCGGG
				SstII CCGCGG
Moraxella species	MspI CCGG	M. MspI $^{m}C^{m}CGG$(12)	NciI CC($^{C}_{G}$)GG	MspI CCGG
			SmaI CCCGGG	HpaII CCGG
				XmaI CCCGGG
Proteus vulgaris	PvuI CGATCG	none[c]	Sau3A GATC	MboI GATC[c]
			BamHI GGATCC	PvuI CGATCG
	PvuII CAGCTG		BglII AGATCT	XorII CGATCG

Organism	Known Restriction System(s)[a]	Known Methylation Site(s)[a]	Cleaved by[b]	Resistant to[b]
			AluI AGCT	
			HindIII AAGCTT	PvuII CAGCTG
			SstI GAGCTC	
Providencia stuartii	PstI CTGCAG	none[c]		PstI CTGCAG
				SflI CTGCAG
			Sau3A GATC	MboI GATC[c]
Pseudomonas aeruginosa	PaeR7 CTCGAG	none	TaqI TCGA	PaeR7 CTCGAG
				XhoI CTCGAG
Serratia marcesens	SmaI CCCGGG	none[c]	HpaII CCGG	SmaI CCCGGG
			MspI CCGG	XmaI CCCGGG
			NciI $CC(^{C}_{G})GG$	
Staphylococcus aureus 3A	Sau3A GATC	none	MboI GATC	Sau3A GATC
			BclI TGATCA	BglII AGATCT
				XhoII PuGATCPy

Organism	Known Restriction System(s)[a]	Known Methylation Site(s)[a]	Cleaved by[b]	Resistant to[b]
				PvuI CGATCG
				XorII CGATCG
				BamHI GGATCC
Thermus aquaticus	TaqI TCGA	M. TaqI TCG^{m}A	AccI GT($^{A}_{C}$)($^{G}_{T}$)AC	TaqI TCGA
				SalI GTCGAC
	TaqII ?			XhoI CTCGAG
				ClaI ATCGAT
Xanthomonas badrii	XbaI TCTAGA	none	MboI GATC	XbaI TCTAGA
			Sau3A GATC	
Xanthomonas holcicola	XhoI CTCGAG	none	TaqI TCGA	XhoI CTCGAG
				PaeR7 CTCGAG
	XhoII PuGATCPy		MboI GATC	
			Sau3A GATC	XhoII PuGATCPy
			PvuI CGATCG	BamHI GGATCC
			XorII CGATCG	BglII AGATCT

Organism	Known Restriction System(s)[a]	Known Methylation Site(s)[a]	Cleaved by[b]	Resistant to[b]
Xanthomonas malvacearum	XmaI CCCGGG	none	HpaII CCGG	XmaI CCCGGG
			MspI CCGG	SmaI CCCGGG
	XmaII CTGCAG		NciI CC(C/G)GG	
			HaeIII GGCC	PstI CTGCAG
	XmaIII CGGCCG			SflI CTGCAG
Xanthomonas nigromaculans	XniI CGATCG	none	MboI GATC	PvuI CGATCG
			Sau3A GATC	XorII CGATCG
			BamHI GGATCC	
			BglII AGATCT	
Xanthomonas oryzae	XorI CTGCAG	none	MboI GATC	PstI CTGCAG
				SflI CTGCAG
			Sau3A GATC	
	XorII CGATCG		BamHI GGATCC	XorII CGATCG
			BglII AGATCT	PvuI CGATCG

[a] Data taken from (5), except where otherwise indicated.

[b] Each DNA was digested with twice the amount of enzyme for twice the length of time necessary to digest either Adenovirus-2 or λ DNA to completion. The DNAs were scored for either their total resistance to or ability to be cleaved by each restriction enzyme. No attempt was made to assess partial protection of some of the sites against cleavage.

[c] All Enterobacteriaceae and Haemophilus species tested contain dam methylase homologous to M. Ecodam that modifies all the GATC sequences and prevents cleavage by MboI (Brooks, et al., in preparation).

TABLE II

Sensitivity of bacterial DNAs to cleavage by related restriction enzymes

DNA	Restriction Enzyme Produced	Core Cutter	Sensitivity to Cleavage
A. calcoaceticus	AccI GT($\frac{A}{C}$)($\frac{G}{T}$)AC	TaqI TCGA	+
A. variabilis	AvaI CPyCGPuG	HpaII CCGG	+
		MspI CCGG	+
		TaqI TCGA	+
B. amyloliquefaciens	BamHI GGATCC	MboI GATC	+
		Sau3A GATC	+
B. caldolyticus	BclI TGATCA	MboI GATC	+
		Sau3A GATC	+
B. globigii	BglII AGATCT	MboI GATC	+
		Sau3A GATC	+
H. influenzae Rc	HincII GTPyPuAC	TaqI TCGA	+
K. pneumoniae	KpnI GGTACC	RsaI GTAC	+
P. vulgaris	PvuI CGATCG	MboI GATC	-[a]
		Sau3A GATC	+
	PvuII CAGCTG	AluI AGCT	+
P. aeruginosa	PaeR7 CTCGAG	TaqI TCGA	+
S. marcesens	SmaI CCCGGG	HpaII CCGG	+
		MspI CCGG	+
X. holcicola	XhoI CTCGAG	TaqI TCGA	+
	XhoII PuGATCPy	MboI GATC	+
		Sau3A GATC	+
X. malvacearum	XmaI CCCGGG	HpaII CCGG	+
		MspI CCGG	+
	XmaIII CGGCCG	HaeIII GGCC	+
X. nigromaculans	XniI CGATCG	MboI GATC	+
		Sau3A GATC	+
X. oryzae	XorII CGATCG	MboI GATC	+
		Sau3A GATC	+

[a] P. vulgaris, like other Enterobacteriaceae, contains a dam methylase homologous to M·Ecodam that modifies all the GATC sequences and prevents cleavage by MboI (Brooks, et al., in preparation).

DNAs act the same with the enzymes used. Both DNAs are resistant to cleavage by EclII and EcoRII enzymes, but are extensively cleaved by AacI and BstNI. All four enzymes are isoschizomers recognizing the sequence CC($\frac{A}{T}$)GG.

Figure 3 shows digests of M. species and H. parainfluenzae DNAs. Both DNAs are resistant to cleavage by HpaII (CCGG) and XmaI (CCCGGG) enzymes. However, the two DNAs differ when treated with MspI, NciI or SmaI enzymes.

TABLE III
Bacterial DNAs that are resistant to cleavage by related restriction enzymes

DNA	Coding for	Resistant to
A. calcoaceticus	AccI GT($\frac{A}{C}$)($\frac{G}{T}$)AC	SalI GTCGAC
	AccII CGCG	BceR CGCG
		FnuDII CGCG
		SstII CCGCGG
A. variabilis	AvaI CPyCGPuG	SmaI CCCGGG
		XmaI CCCGGG
		XhoI CTCGAG
A. luteus	AluI AGCT	HindIII AAGCTT
		PvuII CAGCTG
		SstI GAGCTC
H. influenzae Rc	HincII GTPyPuAC	HpaI GTTAAC
		SalI GTCGAC
T. aquaticus	TaqI TCGA	SalI GTCGAC
		XhoI CTCGAG
		ClaI ATCGAT
X. holcicola	XhoII PuGATCPy	BamHI GGATCC
		BglII AGATCT

M. species DNA is resistant to MspI (CCGG) digestion, whereas H. parainfluenzae DNA is extensively cleaved. In contrast, M. species DNA is susceptible to cleavage by NciI (CC($\frac{C}{G}$)GG) and SmaI (CCCGGG); neither enzyme can cut H. parainfluenzae DNA.

Finally, the largest group represented is the E. coli/M. bovis/S. aureus3A series which is shown in Figure 4. In these digests E. coli DNA can be cut by Sau3A (GATC), XhoII (PuGATCPy), BglII (AGATCT), BamHI (GGATCC), PvuI (CGATCG) and XorII (CGATCG), but is resistant to BclI (TGATCA) and MboI (GATC). S. aureus3A DNA behaves exactly the opposite: it is resistant to Sau3A, XhoII, BglII, BamHI, PvuI and XorII, but is cleaved by both MboI and BclI. M. bovis DNA is somewhat intermediate. It is resistant to MboI and BclI, but also BamHI, BglII, PvuI and XorII; it is cleaved by Sau3A and XhoII. What is known about the modifications of these three

TABLE IV
Bacterial DNAs with differential sensitivity to cleavage by related restriction enzymes

Group A		$CC(^{A}_{T})GG$ DNAs
Enzymes	E. coli	E. cloacae
EcoRII $CC(^{A}_{T})GG$	-	-
EclII $CC(^{A}_{T})GG$	-	-
AacI $CC(^{A}_{T})GG$	+	+
BstNI $CC(^{A}_{T})GG$	+	+

Group B		CCGG DNAs
Enzymes	M. species	H. parainfluenzae
HpaII CCGG	-	-
MspI CCGG	-	+
NciI $CC(^{C}_{G})GG$	+	-
SmaI CCCGGG	+	-
XmaI CCCGGG	-	-

Group C		GATC DNAs	
Enzymes	E. coli	S. aureus 3A	M. bovis
MboI GATC	-	+	-
BclI TGATCA	-	+	-
Sau3A GATC	+	-	+
XhoII PuGATCPy	+	-	+
BamHI GGATCC	+	-	-
BglII AGATCT	+	-	-
PvuI CGATCG	+	-	-
XorII CGATCG	+	-	-

families will be discussed below.

DISCUSSION

From the results there are four major observations that can be made: First, in all cases tested, the DNA itself is protected against all restriction enzymes produced by the bacterium. This protection, in all bacterial systems thus far characterized, is in the form of adenine or cytosine methylation. From these experiments, however, it is not possible to determine what form the modification takes. These results also do not preclude the existence

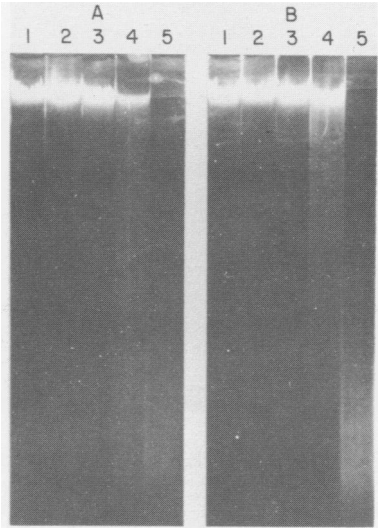

Figure 2: Restriction digests of E. cloacae and E. coli DNAs. E. cloacae and E. coli DNAs were digested by various restriction enzymes and analyzed by gel electrophoresis as described in Materials and Methods. Set A digests contain E. coli DNA; Set B contain E. cloacae DNA. The various lanes include undigested DNA (1) and DNAs cut with EcoRI (2); EclII (3); AacI (4); and BstNI (5).

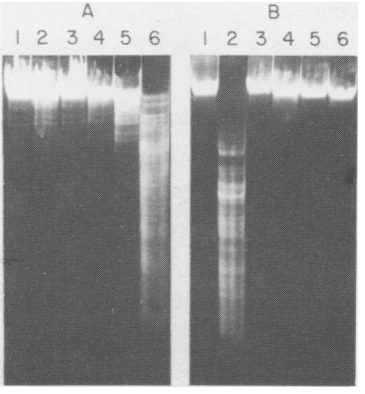

Figure 3: Restriction digests of M. species and H. parainfluenzae DNAs. M. species and H. parainfluenzae DNAs were digested by a series of restriction enzymes and subjected to electrophoresis as described in Materials and Methods. Set A digests contain M. species DNA; Set B digests contain H. parainfluenzae DNA. The various lanes include undigested DNA (1) and DNAs treated with MspI (2); HpaII (3); XmaI (4); SmaI (5) and NciI (6).

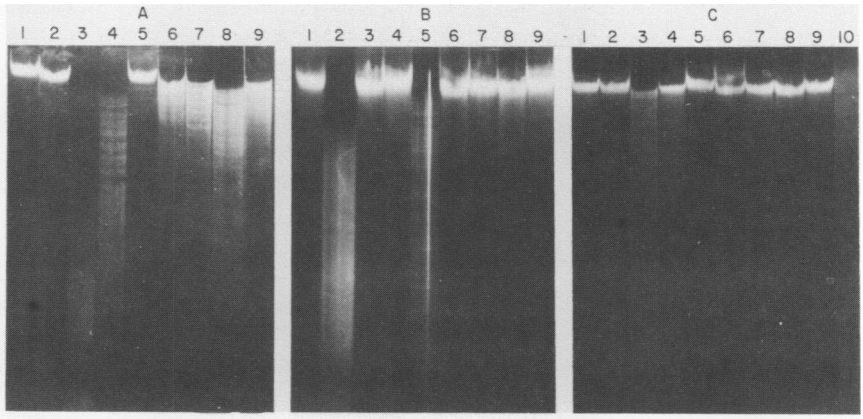

Figure 4: Restriction digests of E. coli, S. aureus3A and M. bovis DNAs.
These three DNAs were digested with the restriction endonucleases enumerated below and the digests analyzed by gel electrophoresis, as described in Materials and Methods. Set A digests contain E. coli DNA; Set B contains S. aureus 3A DNA and Set C, M. bovis DNA. The various lanes contain uncut DNA (1) and DNAs cut by MboI (2); Sau3A (3); XhoII (4); BclI (5); BamHI (6); BglII (7); PvuI (8); XorII (9). In Set C lane (10) contains DNA cut by DpnI.

of specific regulatory proteins for the restriction enzymes as have been found to occur in phage systems (13, 14); however, such proteins would work in addition to DNA modifications.

Second, in no case tested was the DNA modification less specific than its restriction counterpart. For example, there were no cases found where the DNA from an organism making a restriction enzyme specific for a hexameric sequence was protected against a restriction enzyme specific for the central tetrameric sequence (see Table II). In the Type II systems that have been characterized the restriction endonuclease and modification methylase are coded for by separate genes (22, 23, 24); therefore, there must be strong selection pressure to keep the specificities identical. It is possible that excess methylation may be detrimental to the bacterium.

Third, in most cases, a DNA protected against a certain restriction enzyme was protected against all isoschizomers of that restriction enzyme.

Finally, in most instances, DNA modified within the central tetramer of a recognition sequence is protected against cleavage by all related hexameric enzymes having that central tetrameric core (Table III). From these data it is not possible to say whether the protection is due to the modification methylases in each system working on the same residue within the recognition

sequence, or whether most restriction enzymes are sensitive to modifications anywhere within the recognition sequence. The information that is presently available about modification methylases supports both possibilities. For example, HpaI enzyme cannot cut H. influenzae Rc DNA because both are modified at the same position within the sequence (see Table I). However, there are other examples where methylation at a site other than the cognate methylation site can prevent cleavage. M·HindIII prevents HindIII cleavage by methylation of the external adenine in the sequence AAGCTT; however, HindIII enzyme cannot cut A. luteus DNA, whose modification methylation must be within the tetramer AGCT. There are two other examples where two different sites of modification within a recognition sequence can block cleavage. M·TaqI methylates the A residue in the sequence TCGA to protect against TaqI cleavage (25). T. aquaticus DNA is resistant to cleavage by SalI (GTCGAC) and XhoI (CTCGAG); it is known from work with eukaryotic DNAs that SalI and XhoI also cannot cleave DNAs modified at the internal cytosine residue of that core sequence (26). Therefore, for these enzymes at least, methylation at two different positions and on two different bases within the recognition sequence can prevent cleavage (it is not yet known for either SalI or XhoI what the cognate methylation site is).

Apparently methylation anywhere within the central tetramer tends to block restriction endonuclease activity, but there are exceptions. Of course, the most exceptional case is that of DpnI; it can only cut DNA containing methyladenine within its recognition sequence (27). It is the only known restriction enzyme that requires a methylated base for activity. In the case of the EcoRII isoschizomers tested, EclII and EcoRII are unable to cleave DNA modified to $C^{m}C(^{A}_{T})GG$; however, this DNA is extensively cleaved by AacI and BstNI, two other isoschizomers. It has previously been reported that BstNI, unlike its known isoschizomers, can cleave DNA substrates having both cytosine residues on one strand methylated (28). Therefore, there may be three qualitatively different types of restriction systems that recognize the sequence $CC(^{A}_{T})GG$.

One of the most perplexing aspects of our results concerns the cleavage of M. species and H. parainfluenzae DNAs, and is summarized in Table V. It has previously been shown that H. parainfluenzae DNA is methylated at the internal cytosine of the HpaII sequence $C^{m}CGG$ (29). This modification prevents cleavage by HpaII, and also by NciI, SmaI and XmaI. However, this DNA is susceptible to cleavage by MspI. It has also been shown by others that MspI is unable to cleave DNA in which the outer cytosine residue ($C^{m}CGG$) is methylated (30). However, the Msp methylase does not merely modify this outer cytosine residue, but in fact is able to methylate both cytosine residues within

TABLE V

Cleavage of M. species and H. parainfluenzae DNAs

Enzyme		M. species DNA	H. parainfluenzae DNA
MspI	CCGG	resistant	cleaved
HpaII	CCGG	resistant	resistant
NciI	CC($^{C}_{G}$)GG	cleaved	resistant
SmaI	CCCGGG	cleaved	resistant
XmaI	CCCGGG	resistant	resistant

this sequence. A natural inference from this observation would be that M. species DNA carries modifications of both cytosines within this sequence. This would account for the fact that HpaII and XmaI are unable to cleave M. species DNA since both enzymes are known to be inhibited by methylation at the internal cytosine residue. However, it does not account for the observation that both NciI and SmaI are able to cleave M. species DNA since, from our results with H. parainfluenzae DNA, we find that both of these enzymes are unable to cleave when the internal cytosine residue is methylated. Since XmaI is also unable to cleave H. parainfluenzae DNA, and recognizes the same sequence as SmaI, it is hard to rationalize the differential cleavage potential of these two enzymes on M. species DNA. One possible interpretation is that in the case of NciI and SmaI the double modification somehow negates the effect of a single modification at the internal cytosine residue. Clearly, further studies are required to firmly establish the cause of this apparent paradox. In particular, it will be important to establish whether all of the CCGG sequences in M. species DNA are fully modified at both cytosines, and perhaps also to establish whether modification other than methylation plays a role in the protection of M. species DNA.

Another interesting aspect of this series of digests is the differential cleavage activity of SmaI and XmaI enzymes, two isoschizomers recognizing the sequence CCCGGG: SmaI can cut M. species DNA while XmaI enzyme cannot. This difference could prove useful when it is better characterized how the enzymes interact with methylated substrates.

Among the family of enzymes recognizing sequences containing the central tetranucleotide GATC, it is apparent that two different kinds of modification exist. As summarized in Table VI, the enzymes MboI and BclI are inhibited by A-methylation, while the remaining members of this family are inhibited by C-methylation. It has been shown directly that MboI is blocked by dam methylation ($G^{m}ATC$) (5), while for BamHI, the analysis of methylated sequences in

TABLE VI

Modifications preventing cleavage by the GATC family of enzymes

G^mATC	GAT^mC
MboI (GATC)	Sau3A (GATC)
BclI (TGATCA)	BamHI (GGATCC)
	BglII (AGATCT)
	PvuI (CGATCG)
	XorII (CGATCG)
	XhoII (PuGATCPy)

B. amyloliquefaciens DNA reveals that 5-methylcytosine is present in the sequence $GGATC^mC$ (19). By inference from our results, the other members of this family can be ordered according to their ability to cleave one another's DNAs. The only unexpected results concern the cleavage of M. bovis DNA. Although this DNA is resistant to both MboI and BclI, and can be cut by Sau3A and XhoII, it is resistant to cleavage by BamHI, BglII, PvuI and XorII. Clearly, adenine modification alone cannot be the reason for this, since these enzymes will cleave DNA, carrying only dam-like methylation. We suspect that our results are a consequence of additional modifications which may be quite extensive in M. bovis DNA. For instance, this DNA is resistant to cleavage by several unrelated restriction enzymes like SmaI (CCCGGG) and SacII (CCGCGG), and is only poorly cleaved by HaeIII (GGCC), HpaII (CCGG) and FnuDII (CGCG). A similar situation occurs with DNA from the closely-related species Neisseria gonorrhoea, which is resistant to cleavage by BamHI and SacII, and is poorly cleaved by many other enzymes (31). It should be mentioned here that in most of the bacteria that have been rigorously studied, more methylases and/or modified bases than can be accounted for by known restriction modification systems have been found (e.g. E. coli (32), H. influenzae (33), B. amyloliquefaciens (19), B. brevis (19) and B. subtilis (34)).

In summary, this detailed analysis of bacterial DNA modifications has proven useful in two ways. First, the analysis has shown many general rules as to how DNAs are modified against the action of restriction enzymes. Second, the analysis has revealed several interesting examples of restriction enzymes or modified DNAs that do not follow the rules and are therefore worthy of further investigation.

ACKNOWLEDGMENTS

We thank Phyllis Myers for preparation of the restriction enzymes and

help in growing the many bacteria used in these experiments. We are grateful to Dan Levy for his gift of M. species DNA and Drs. Robert Blumenthal and Thomas Gingeras for helpful comments on the manuscript. We thank D. O'Loane for her technical assistance, N. D'Anna for preparation of the manuscript and M. Ockler for photography. The work was supported by National Science Foundation Grant # PCM 7919-882.

REFERENCES

1. Arber, W. (1974) in progress in Nucleic Acids Research and Molecular Biology, Cohn, W.E. Ed., Vol XIV, pp. 1-37, Academic Press, New York.
2. Smith, H.O. (1979) Science 205, 455-466.
3. Modrich, P. (1979) Q. Rev. Biophys. 12, 315-369.
4. Yuan, R. (1981) Ann. Rev. Biochem. 50, 285-315.
5. Roberts, R.J. (1981) Nucl. Acids Res. 9, r75-96.
6. Nathans, D. and Smith, H.O. (1975) Ann. Rev. Biochem. 44, 273-293.
7. Roberts, R.J. (1976) CRC Crit. Rev. Biochem. 4, 123-164.
8. Stuy, J.H. (1976) J. Bacteriol. 128, 212-220.
9. Hattman, S. (1981) in The Enzymes, Vol 15, in press.
10. Gold, M. and Hurwitz, J. (1963) Cold Spring Harbor Symp. Quant. Biol. 28, 275-298.
11. Borek, E. and Srinivasan, P.R. (1966) Ann. Rev. Biochem. 35, 275-298.
12. Jentsch, S., Günthert, U. and Trautner, T. (1981) Nucl. Acids Res. 9, 2753-2759.
13. Studier, F.W. (1975) J. Mol. Biol. 94, 283-295.
14. Makino, O. and Saito, H. (1980) Mol. Gen. Genet. 179, 463-468.
15. Marmur, J. (1961) J. Mol. Biol. 3, 208-218.
16. Pettersson, U. and Sambrook, J. (1973) J. Mol. Biol. 73, 125-130.
17. Schrenk, W.J. and Weisberg, R.A. (1975) Mol. Gen. Genet. 137, 101-107.
18. Clewell, D.B. and Helinski, D.R. (1972) J. Bacteriol. 110, 1135-1146.
19. Hattman, S., Keister, T. and Gottehrer, A. (1978) J. Mol. Biol. 124, 701-711.
20. Hattman, S., Brooks, J.E. and Masurekar, M. (1978) J. Mol. Biol. 126, 367-380.
21. May, M.S. and Hattman, S. (1975) J. Bacteriol. 123, 768-770.
22. Yoshimori, R., Roulland-Dussoix, D. and Boyer, H.W. (1972) J. Bacteriol. 112, 1275-1279.
23. Mann, M.B., Rao, R.N. and Smith, H.O. (1978) Gene 3, 97-112.
24. Walder, R.Y., Hartley, J.L., Donelson, J.E. and Walder, J.A. (1981) Proc. Nat. Acad. Sci., USA 78, 1503-1507.
25. Sato, S., Nakazawa, K. and Shinomiya, T. (1980) J. Biochem. 88, 737-747.
26. van der Ploeg, L.H.T. and Flavell, R.A. (1980) Cell 19, 947-958.
27. Lacks, S. and Greenberg, B. (1975) J. Biol. Chem. 250, 4060-4072.
28. Gruenbaum, Y., Cedar, H. and Razin, A. (1981) Nucl. Acids Res. 9, 2509-2515.
29. Mann, M.B. and Smith, H.O. (1977) Nucl. Acids Res. 4, 4211-4221.
30. Sneider, T.W. (1980) Nucl. Acids Res. 8, 3829-3840.
31. Norlander, L., Davies, J.K., Hagblom, P. and Normark, S. (1981) J. Bacteriol. 145, 788-795.
32. Marinus, G.M. and Morris, N.R. (1973) J. Bacteriol. 114, 1143-1150.
33. Roy, P.H. and Smith, H.O. (1973) J. Mol. Biol. 81, 427-444.
34. Trautner, T.A., Pawlek, B., Günthert, U., Canosi, U., Jentsch, S. and Freund, M. (1980) Molec. Gen. Genet. 180, 361-367.