Enzyme Information

Enzyme	Recognition Sequence	Sites cut	Sites not cut
EcoR I	GAATTC	GAATTh^{m5}C	G^{m6}AATTC^{b} GA^{m6}ATTC# GAATT^{m5}C^{b}
EcoR II	CCWGG	^{m5}CCWGGb	^{m4}CCWGG C^{m4}CWGG C^{m5}CWGG# CC^{m6}AGG h^{m5}Ch^{m5}CWGG
EcoR V	GATATC	GATAT^{m5}C^{b}	G^{m6}ATATC# GAT^{m6}ATC
EcoR 124	GAAN6RTCG^{b}	?	GA^{m6}AN6RTCG GAAN6RmTCG
EcoR 124/3	GAAN7RTCG^{b}	?	^{m6}A
Ehe I	GGCGCC	?	GG^{m5}CGCC
Esp I	GCTNAGC	GCTNAG^{m5}C	G^{m5}CTNAGC
Fnu4H I	GCNGC	?	G^{m5}CNGC GCNG^{m5}C
FnuD II	CGCG	?	^{m5}CGCG CG^{m5}CG
Fok I	CATCC	CAT^{m5}CC CATC^{m5}C^{b}	GG^{m6}ATG C^{m6}ATCC CATC^{m4}C
Fse I	GGCCGGCC	?	GG^{m6}CCGG^{m6}CC GGC^{m5}CGGCC GG^{m5}CCGGCC
Fsp I	TGCGCA	?	TG^{m5}CGCA
Hae II	RGCGCY^{b}	?	RG^{m5}CGCY RGh^{m5}CGh^{m5}CY
Hae III	GGCC	GGC^{m5}C	GG^{m5}CC#^{b} GGh^{m5}Ch^{m5}C
Hap II	CCGG	?	C^{m5}CGG#
Hga I	GACGC	?	GA^{m5}CGC GACG^{m5}C
HgiA I	GRGCYC	GRGCY^{m5}C	GRG^{m5}CYC
HgiC I	GGYRCC	?	GGYRC^{m5}C
HgiC II	GGWCC	?	GGWC^{m5}C
HgiE I	GGWCC	?	GGWC^{m5}C
HgiJ II	GGYRCC	?	GGYRC^{m5}C
Hha I	GCGC	?	G^{m5}CGC# GCG^{m5}C Gh^{m5}CGh^{m5}C
Hha II	GANTC	?	G^{m5}ANTC#
Hinc II	GTYRAC	GTYRA^{m5}C	GTYR^{m6}AC GTYRAh^{m5}C
Hind II	GTYRAC	?	GTYR^{m6}AC#
Hind III	AAGCTT	A^{m6}AGCTT#	^{m6}AAGCTT# AAG^{m5}CTT AAGh^{m5}CTT
Hinf I	GANTC	GANT^{m5}C^{b}	G^{m5}ANTC GANTh^{m5}C
HinP I	GCGC	?	G^{m5}CGC
Hpa I	GTTAAC	GTTAA^{m5}C	GTTA^{m6}AC# GTTAAh^{m5}C
Hpa II	CCGG	?	^{m4}CCGG ^{m5}CCGG^{b} C^{m4}CGG^{b} C^{m5}CGG# h^{m5}Ch^{m5}CGG
Hph I	TCACC	TCAC^{m5}C	T^{m5}CACC# GGTG^{m6}A
Kpn I	GGTACC^{b}	GGTA^{m5}CC GGTAC^{m5}C GGTA^{m5}C^{m5}C^{b} GGT^{m6}ACC	GGT^{m6}A^{m6}CC GGTAC^{m4}C
Kpn2 I	TCCGGA	TCCGG^{m6}A	T^{m5}CCGGA TC^{m5}CGGA
Ksp I	CCGCGG	?	^{m5}CCGCGG C^{m5}CGCGG
Mae II	ACGT	?	A^{m5}CGTb
Mam I	GATN4ATC	?	G^{m6}ATN4^{m6}ATC
Mbo I	GATC^{b}	GAT^{m4}C GAT^{m5}C^{b}	G^{m6}ATC# GATh^{m6}C
Mbo II	GAAGA	T^{m5}CTT^{m5}C^{b} G^{m6}AAGA	GAAG^{m6}A#
Mfl I	REGATCY^{b}	?	RG^{m6}ATCY RGAT^{m4}CY RGAT^{m5}CY
Mlu I	ACGCGT	^{m6}ACGCGT	A^{m5}CGCGT
Mme II	GATC	?	G^{m6}ATC
Mnl I	CCTC^{b}	?	^{m5}CCTC ^{m5}C^{m5}CT^{m5}C
Mro I	TCCGGA	TCCGG^{m6}A	T^{m5}CCGGA TC^{m5}CGGA
Mse I	TTAA	TTm6AA	?
Msp I	CCGG^{b}	^{m4}CCGG C^{m4}CGG C^{m5}CGG	^{m5}CCGG# h^{m5}Ch^{m5}CGG
Mst II	CCTNAGG	^{m5}CCTNAGG	?
Mun I	CAATTG	?	CA^{m6}ATTG
Mva I	CCWGG	C^{m5}CWGG^{b} ^{m5}CCWGG CC^{m6}AGG^{b}	C^{m4}CWGG# ^{m4}CCWGG ^{m5}C^{m5}CWGG^{b}
Mvn I	CGCG	?	m5CGCG
Nae I	GCCGGC^{b}	?	G^{m5}CCGGC GC^{m5}CGGC GCCGG^{m5}C
Nar I	GGCGCC	GGCGC^{m5}C	GG^{m5}CGCC GGCGC^{m4}C
Nci I	CCSGG	^{m5}CCSGG	C^{m4}CSGG C^{m5}CSGG^{b}
Nco I	CCATGG	CC^{m6}ATGG	^{m4}CCATGG^{b} ^{m5}CCATGG
Nde I	CATATG	^{m5}CATATG^{b}	^{m6}A
Nde II	GATC	GAT^{m5}C^{b}	G^{m6}ATC
Nhe I	GCTAGC	?	GCTAG^{m5}C
Nla III	CATG	?	C^{m6}ATG#
Not I	GCGGCCGC	GCGGCCG^{m5}C	GCGG^{m5}CCGC GCGGC^{m5}CGC
Nru I	TCGCGA	TCG^{m5}CGA	T^{m5}CGCGA TCGC^{m6}A
Nsi I	ATGCAT	?	ATGC^{m6}AT ATG^{m5}CAT
Nsp I	RCATGY	?	RC^{m6}ATGY
NSpB II	CMGCKG	C^{m5}CGCKG	?
PflM I	CCAN5TGG	?	C^{m4}CAN5TGG C^{m5}CAN5TGG

LEGEND

a. # denotes canonical modification mTase specificity.

b. See notes section of reference 2.

*Dpn I and its isoschizomers require the presence of 6-methyladenosine within the recognition sequence GATC

Recognition sequences given use the standard abbreviations [*Eur. J. Biochem* (1985) [50:15] to represent ambiguity:

R=G or A
M=A or C
S=G or C
H=A or C or T
V=G or C or A
N=A or C or G or T
Y=C or T
K=G or T
W=A or T
B=G or T or C
D=G or A or T

^{m4}C = N4-methylcytosine
^{m5}C = C5-methylcytosine
^{hm5}C = hydroxymethylcytosine
^{m}C = methylcytosine
N_4 or C_5-methylcytosine unspecified
^{m6}A = N_6-methyladenine

Sequences are in 5′ - 3′ order.

Enzyme Information, continued

Enzyme	Recognition Sequence	Sites cut	Sites not cut
Pfu I	CGTACG	?	CGTA^{m5}CG
PaeR7 I	CTCGAG	?	CTCG^{m6}AG# CT^{m5}CGAG
Pml I	CACGTG	?	CA^{m5}CGTG
PpuAI	CGTACG	?	CGTA^{m5}CG
PspA I	CCCGGG	C^{m5}CGGG CC^{m5}CGGG	^{m5}CCCGGG
Pst I	CTGCAG	?	^{m5}CTGCAG CTGC^{m6}AG#
Pvu I	CGATCGb	CG^{m6}ATCG CGAT^{m5}CG	CGAT^{m4}CG
Pvu II	CAGCTG	?	CAG^{m4}CTG# CAG^{m5}CTG
Rsa I	GTACb	GTA^{m5}C^{b}	GT^{m6}AC
Rsr I	GAATTC	?	G^{m6}AATTC GA^{m5}ATTC#^{b}
Rsr II	CGGWCCG	?	^{m5}CGGWCCG CGGW^{m5}CCG CGGWC^{m5}CG
Sac I	GAGCTC	G^{m6}AGCTC	GAG^{m5}CTC
Sac II	CCGCGG	?	^{m5}CCGCGG
Sal I	GTCGAC	GTCGA^{m5}C	GT^{m5}CGAC GTCG^{m6}AC#
Sau3A I	GATC^{b}	G^{m6}ATC	GAT^{m5}C#^{b} GAT^{m4}C GATh^{m5}C
Sau96 I	GGNCC	?	GGN^{m5}CC# GGNC^{m5}C GGNh^{m5}Ch^{m5}C
Sca I	AGTACT	AGTA^{m5}CT	?
ScrF I	CCNGG	^{m5}CCNGG	C^{m5}CNGG C^{m4}CNGG
SfaN I	GATGC	GATG^{m5}C	G^{m6}ATGC
Sfi I	GGCCN5GGCC	GG^{m5}CCN5GG^{m6}CC^{b} GGCCN5GGC^{m6}C	GGC^{m6}CN5GGCC
SfI I	CTGCAG	?	CTGC^{m6}AG
SgrA I	CRCCGGYG	?	CRC^{m6}CGGYG
Sin I	GGWCC	?	GGW^{m5}CC#
Sma I	CCCGGG	C^{m5}CCGGG	^{m4}CCCGGG ^{m5}CCCGGG^{b} C^{m4}CCGGG^{b}
SnaB I	TACGTA	?	TA^{m6}CGTA T^{m6}ACGT^{m6}A
Sno I	GTGCAC	?	GTG^{m5}CA^{m5}C
Spe I	ACTAGT	?	^{m6}ACTAGT A^{m5}CTAGT
Sph I	GCATGC	GCATG^{m5}C Gh^{m5}CATGh^{m5}C	GC^{m6}ATGC
Spl I	CGTAGC	CGTm6ACG	?
Spo I	TCGCGA	TCGCG^{m6}A	T^{m5}CGCGA TCG^{m5}CGA
Srf I	GCCC/GGGC	GCCC/GGG^{m5}C	G^{m5}CCC/GGGC GC^{m5}CC/GGGC GCC^{m6}C/GGGC
Ssp I	AATATT	^{m6}AATATT	?
Sst I	GAGCTC	?	GAG^{m5}CTC GAGh^{m5}CTh^{m5}C
Stu I	AGGCCT	?	AGGm5CCT AGGCm5CT AGGCm4CT
StySP I	AACN6GTRC^{b}	?	A^{m6}ACN6GmTRC#^{b}
Taq I	TCGA	T^{m5}CGA^{b} Th^{m5}CGA^{b}	TCG^{m6}A#
Taq II	GACCGA CACCCA	?	G^{m6}ACCGA
Tfi I	GAWTC	GAWT^{m5}C	?
Tha I	CGCG	^{m5}CGCG	^{m5}CGCG h^{m5}CGh^{m5}CG
Xba I	TCTAGA	?	TCTAG^{m6}A# T^{m5}CTAGA Th^{m5}CTAGA
Xho I	CTCGAG^{b}	?	CT^{m5}CGAG CTCG^{m6}AG ^{m5}CTCGAG
Xho II	RGATCY	RG^{m6}ATCY	RGAT^{m5}CY^{b}
Xma I	CCCGGG	CC^{m5}CGGG^{b}	^{m4}CCCGGG ^{m5}CCCGGG C^{m4}CCGGG CC^{m4}CGGG
Xma III	CGGCCG	?	CGGm5CCG
Xmn I	GAAN4TTC	GA^{m5}AN4TTC	G^{m6}AAN4TTC GAAN4TT^{m6}C^{b}
Xor II	CGATCG	CG^{m6}ATCG	CGAT^{m5}CG h^{m6}CGATh^{m5}CG

LEGEND

a. # denotes canonical modification mTase specificity.

b. See notes section of reference 2.

Recognition sequences given use the standard abbreviations [*Eur. J. Biochem* (1985) [50:15] to represent ambiguity:

R=G or A
M=A or C
S=G or C
H=A or C or T
V=G or C or A
N=A or C or G or T

Y=C or T
K=G or T
W=A or T
B=G or T or C
D=G or A or T

^{m4}C = N4-methylcytosine
^{m5}C = C5-methylcytosine
^{hm5}C = hydroxymethylcytosine
^{m}C = methylcytosine
N_4 or C_5-methylcytosine unspecified
^{m6}A = N_6-methyladenine
Sequences are in 5′ - 3′ order.

*Dpn I and its isoschizomers require the presence of 6-methyladenosine within the recognition sequence GATC

REFERENCE

1. Roberts, R.J. and Macelis, D. (1991) *Nucleic Acids Res.* 19: 2077-2109.
2. Nelson, M. and McClelland, M. (1991) *Nucleic Acids Res.* 19:2045-2071