EcoR124I: from Plasmid-Encoded Restriction-Modification System to Nanodevice

James Youell and Keith Firman*

School of Biological Sciences, University of Portsmouth, King Henry Building, King Henry I Street, Portsmouth, Hants PO1 2DY, United Kingdom

INTRODUCTION

Plasmid R124 and the restriction-modification (R-M) system that it encodes have an unusual history, one that is confused by mistakes and misidentification. These mistakes range from confusion over the origin of the plasmid (Escherichia coli versus Salmonella enterica serovar Typhimurium) through the misidentification of the R-M enzyme encoded by R124 as the EcoRI system. This situation is compounded by a great deal of confusion about the incompatibility grouping of the plasmid. This review is written in part to clarify this confusion but also to tell an interesting story about the identification, purification, and eventual bionanotechnological exploitation of the R-M enzyme EcoR124I.

Background of R-M Systems

Although this is a historical review about plasmid R124, one of the main features of this plasmid is the presence of a unique R-M system (see below). Therefore, it is useful to briefly review R-M systems and briefly detail type I R-M systems, of which the R124 R-M system is a member.

R-M systems can be viewed as being equivalent to a bacterial immune system, providing the bacteria with protection from invading DNA, and, as such, is an endonuclease and directly affords the bacterial host with protection from invading DNA, and a modification enzyme, which is responsible for DNA cleavage and, as such, is an endonuclease and directly affords the bacterial host with protection from invading DNA, and a modification enzyme, which is a DNA methyltransferase (MTase) and methylates DNA at the target recognition sequence of the restriction enzyme, thus preventing the host chromosome from being cleaved and destroyed by the host restriction enzyme. However, variations to this general theme exist (e.g., the type IV MrcBC system [53], which has no separate MTase and cleaves only modified DNA substrates, or the type IIM sub-group, which recognizes methylated DNA). The restriction enzymes fall into a wide variety of families and subtypes (72), and these all produce different DNA products depending upon the mechanism of DNA cleavage (94).

Most R-M systems consist of at least two components: a restriction enzyme, which is responsible for DNA cleavage and, as such, is an endonuclease and directly affords the bacterial host with protection from invading DNA, and a modification enzyme, which is a DNA methyltransferase (MTase) and methylates DNA at the target recognition sequence of the restriction enzyme, thus preventing the host chromosome from being cleaved and destroyed by the host restriction enzyme. However, variations to this general theme exist (e.g., the type IV MrcBC system [53], which has no separate MTase and cleaves only modified DNA substrates, or the type IIM sub-group, which recognizes methylated DNA). The restriction enzymes fall into a wide variety of families and subtypes (72), and these all produce different DNA products depending upon the mechanism of DNA cleavage (94).

Type I R-M enzymes are multifunction, multisubunit enzymes that bind at a specific DNA sequence but are able to translocate the DNA substrate and cleave only when translocation is stalled; often, this stalling, on circular DNA, is due to DNA topology (42, 85, 86; P. Janscák and K. Firman, presented at the 4th New England Biolabs Workshop on Biological DNA Modification, Innsbruck, Austria, 1997), and this leads to the production of random-sized fragments of DNA. The R-M enzyme is capable of both endonuclease activity and methylase activity dependent upon the initial methylation status of the DNA, and it is through this mechanism that the R-M enzyme is able to discriminate host and invading DNA. DNA translocation involves a motor activity, which has led to these enzymes being included in the SFII superfamily of closely related DNA helicases (22). However, this motion is not due to a simple helicase activity (83).

Type II R-M enzymes were first identified during the 1970s.
(49, 75) and are much more simple, usually with independent endonuclease (often functioning as a dimer) and methylase (usually a monomer) enzymes, which have evolved to recognize the same DNA sequence. These enzymes usually cut at relatively short (4- to 8-bp) recognition sequences to which they also bind specifically. Their simplicity and ability to cleave DNA into discrete fragments and the ease of purification of these enzymes led to their use in genetic engineering, which in turn led to an explosion in the isolation, purification, and use of these enzymes (71, 73).

Type III R-M enzymes are almost as complex as type I R-M enzymes and are also members of the SFII superfamily, with combined endonuclease and methylase activities associated with a single enzyme (40, 52). They are unique in that their DNA recognition sequence consists of two inverted repeats of the same DNA sequence (55), which produce a loop of DNA, bringing the bound enzymes together for DNA cleavage (46).

More recently, the clear classification of restriction enzymes into different types has been blurred by the discovery of subclasses and new types. These include type IIIs, which usually cut DNA at an offset site reminiscent of type III cleavage, and further subgroups of type I from type IIA through type IIM (72). A new type IV group of enzymes was proposed by Janulaitis et al. in 1992, typified by Eco57I (47), but this designation was changed in 2003 by Roberts et al. (72), who designated the EcoKMcrBC system as type IV and redesignated Eco57I as type IIIG.

**ISOLATION AND CHARACTERIZATION OF PLASMID R124**

Plasmid R124 was first documented by Meynell and Datta in 1966 (60) and is described as “coming from Salmonella [enterica] serovar Typhimurium”... as a strain sent to the Enteric Reference Laboratory, Central Public Health Laboratory, London for phage-typing.” This original source for R124 has led to subsequent confusion over the naming of the R-M system that it encodes.

The plasmid itself was described as an $f^+$ plasmid (that is, it inhibits the transfer of the sex factor $F^+$) and is therefore closely related to $F^+$, with some 65% DNA identity (2, 39, 60; H. Smith, personal communication [2]), and carries the conjugation/transfer system of $F^+$. The bacterial strain carrying the plasmid was likely to have been first isolated because R124 also carries a tetracycline resistance determinant possibly producing a Salmonella infection that was resistant to this antibiotic.

During the early 1960s, the detection and isolation of novel R-M systems (11, 65, 89) were of major interest to many research groups. As the 1970s arrived, these interests intensified as the biochemistry of restriction enzymes was gradually uncovered (3, 59, 74) and as their potential to cleave DNA into discrete fragments and the ease of purification of these enzymes led to their use in genetic engineering, which in turn led to an explosion in the isolation, purification, and use of these enzymes (71, 73).

Bannister and Glover (6) analyzed a collection of 153 antibiotic resistance plasmids for the presence of restriction activity (the ability to reduce the plating efficiency of bacteriophage lambda) and identified 61 plasmids as being restriction proficient. Interestingly, R124 was the only plasmid in this collection that was both $f^+$ and restriction proficient, suggesting that a novel R-M system might be present, and this system was named hspI. Unfortunately, this is where the first major confusion involving R124 was to occur. The other group of restriction-proficient plasmids was $f^-$ and belonged to the incompatibility group N. The R-M system of these plasmids was designated hspII (91) and included the well-described EcoRI enzyme (95). In 1971, Yoshimori (96) also isolated plasmids carrying R-M systems, which he designated R1 and RII, which subsequently became known as EcoRI and EcoRII. The plasmid carrying the RI R-M system was (like R124) $f^+$, and it seems that the reviewers of the time mistakenly identified the RI and the hspI restriction systems as being the same. In fact, Bannister and colleague (5, 6) had already shown subtle differences in the genetic phenotypes of the RI system and the hspI system (Bannister was able to identify both Res$^-$Mod$^-$ mutants, while Yoshimori identified only Res$^-$ mutants, suggesting a possible genetic difference between these systems, but this was not readily noted in the literature at the time).

However, the confusion did not involve only R124, and in fact, even the story involving EcoRI is confusing. It was later shown that the $f^+$ plasmid, which was described as carrying the EcoRI R-M system (plasmid RY5), was in fact a bacterium containing a complex mix of plasmids and that EcoRI actually resided on a nonconjugative, but transmissible, ColE1 plasmid, NTP13 (8, 82). As a consequence, during the early 1970s, the EcoRI and EcoRII R-M systems were often confused with one another, a situation that in Russia almost survived the cold war (V. Zinkevich, personal communication).

Hedges and Datta (34) showed that R124 was, interestingly, the first member of a unique, new plasmid incompatibility group. The assay for incompatibility was based on the failure of two incompatible plasmids to coexist in the same bacterial cell, and therefore, by definition, R124 was found to coexist with all plasmids available at that time. This incompatibility group was a subgroup of the F family (based on the $f^+$ property of the plasmid) and was designated IncFIV.

**GENETIC CHARACTERIZATION OF THE R-M GENES OF R124**

As mentioned above, the first genetic characterization of the R-M system carried by R124 was carried out by Bannister (5), who showed that this system could restrict the growth of bacteriophages T1, φ80, P1, and λ, and she named this system host specificity system I (hspI).

In order to clarify the situation with regard to the two R-M systems hspI (R124) and RI (RY5), Hughes carried out a more comprehensive analysis of both of these systems (39). Unfortunately (because it resulted in more confusion involving R124), he named the enzyme that he isolated from plasmid R124 EcoR124, but as mentioned above, plasmid R124 was actually isolated from S. enterica serovar Typhimurium, and the enzyme should have been named StyR124. However, his detailed analysis of the plating efficiencies of bacteriophage grown on the RI and hspI R-M systems showed that they were distinctly different systems (39). Therefore, the isolation and biochemical characterization of the restriction enzyme were now important in order to confirm this evidence.

**EcoR124 AS A TYPE I R-M ENZYME**

Purification of the now-named EcoR124 enzyme was problematic, and early results were contaminated by an un-
known, additional, nonspecific nuclease activity (39). This led to a mistaken conclusion that EcoR124 DNA cleavage activity, like EcoRI DNA cleavage activity, was ATP and $S$-adenosyl methionine independent (39). Phosphocellulose chromatography allowed Hughes to purify the EcoR124 enzyme away from the nonspecific nuclease activity, but this also led to an unexpected concomitant loss of EcoR124 enzyme activity. This confusing situation led to one of the most unexpected results from experimentation in this field. Hughes thought that the contaminating nuclease might be exonuclease III and decided to transfer plasmid R124 into a new genetic background that lacked this enzyme (an $xth$ mutant strain of $E. coli$). He isolated five exconjugants following the conjugation of the plasmid into bacterial strain $E. coli$ BW9091 and showed that four of these exconjugants carried the expected EcoR124 R-M activity. Unexpectedly, one exconjugant carried a plasmid with a novel R-M activity, and this plasmid was named R124/3 (39). At the time, this was the first example of the isolation of a natural variant of R-M specificity, and that author showed that the specificity of this system was different from those of all other known systems available at that time. To have isolated the plasmid from a simple conjugation transfer was particularly unusual, a story that was to become even more confusing.

FIG. 1. Restriction fragment map (EcoRI) of plasmid R124. The restriction fragment map of R124 is based on the map produced by Campbell and Mee (13). The sizes of the EcoRI fragments (in kbp) are as follows: E1, 15.08; E2, 13.83; E3, 13.11; E4, 11.32; E5, 10.56; E6, 8.27; E7, 8.14; E8, 7.06; E9, 6.84; E10, 6.19; E11, 6.04; E12, 5.85; E13, 3.23; E14, 3.06; E15, 2.25; E16, 1.43; E17, 0.97; E18, 0.93; E19, 0.59; E20, 0.42; E21, 0.33; E22, 0.20. Fragment E3 includes an internal EcoRI recognition site, which is methylated by the EcoR124 R-M enzyme (dotted line) (see the text for further details). The location of the hsd genes is based on the observations described by Firman et al. (19, 21), but the nomenclature used for numbering of the EcoRI fragments differs by 1 after E6 (which was thought to be one band and not two). Therefore, the hsd genes were described by Firman et al., following Tn7-based mutagenesis, as being located on EcoRI fragment 8 (equivalent to EcoRI fragment 9 in the map described above).

FIG. 2. Recognition sequence overlap for EcoR124I, EcoRI, and SalI. The recognition sequence of the EcoR124I enzyme was determined by Price et al. (68), and alignment with the recognition sequences of EcoRI and SalI restriction enzymes shows how methylation by the EcoR124I enzyme would also methylate the overlapping EcoRI and SalI sites.

INITIAL BIOCHEMICAL CHARACTERIZATION OF EcoR124

As mentioned above, the initial purification of EcoR124 described by Hughes (39) was a difficult process. The problem was related to an ATP-independent, contaminating nuclease activity, which coeluted from all ion-exchange columns used. In fact, Hughes was able to obtain reasonably pure protein only following the “accidental introduction of a step in the elution gradient” (39), which indicated the use of a stepwise gradient of NaCl (0 to 0.5 M) for the elution of bound protein from a phosphocellulose column. This purified EcoR124 no longer included the contaminating nuclease activity, and Hughes was able to show that the enzyme required ATP as a cofactor (indicative of a type I R-M enzyme [type III enzymes had not yet been described]) and that the DNA cleavage activity was at least stimulated by $S$-adenosyl methionine. Along with the genetic evidence described by Bannister and Glover (6), this suggested a type I R-M system.

Hughes was also able to purify the EcoR124/3 enzyme (39) and demonstrate identical properties, although the enzyme must have a different DNA specificity to explain the novel restriction and modification activities observed in vivo.

R124 AND R124/3, TWO CLOSELY RELATED PLASMIDS

The isolation of plasmid R124/3 renewed interest in the R-M system encoded by plasmid R124. Our own interest was to determine the nature of the changes between the two plasmids, which led to a novel R-M phenotype for R124/3. Therefore, we set about physically mapping both plasmids, locating the hsd genes on each plasmid, and mutating the hsd genes in order to better understand the R-M system. Hughes had already produced agarose gels showing differences between the EcoRI and SalI restriction restriction pattern fragments of R124 and R124/3 (Fig. 1) (39). He attributed these changes to the changed DNA specificity of EcoR124/3, suggesting that methylation by the new R-M enzyme was allowing cleavage by these restriction enzymes (Fig. 2), but the evidence was unfortunately weak (however, later determination of the recognition sequence of EcoR124I demonstrated that this methylation was possible). At about the same time, we had also isolated plasmid DNA from both R124 and R124/3 and shown that there were major physical differences between the plasmids (Fig. 3) (30). It appears that the changes between these two plasmids that we observed were greater than those observed by Hughes, suggesting that further DNA rearrangements may have occurred. In particular, the large EcoRI fragment E1 was deleted by some 1.5 kbp, in addition to the cleavage of the “phantom” EcoRI fragment E3 into two new fragments, as described by Hughes.

At that time, we were interested in isolating restriction-deficient mutants of both R-M systems in order to better un-
understand the genetics of the system. In order to isolate such restriction-deficient mutants, an enrichment procedure is required. Therefore, we decided to investigate the use of an incoming F lac plasmid as a first-stage enrichment step. The theory behind such a selection step is that restriction-proficient incoming F lac bacteria will more readily accept the F lac plasmid than restriction-deficient bacteria will restrict the incoming F lac, while restriction-deficient bacteria will more readily accept the F lac plasmid. The first stages of these genetic experiments were an ideal third-year research project for an undergraduate, and this was how the work began. The very first experiment was to determine the level of restriction that EcoR124I produces against an incoming F lac plasmid. The transfer frequency of F lac was restricted by 2 orders of magnitude compared to its normal transfer frequency (77). The first experiments were to determine the level of restriction that EcoR124I produces against an incoming F lac plasmid. The transfer frequency of F lac was restricted by 2 orders of magnitude compared to its normal transfer frequency. This occurred when R124 and F + were present together in a bacterium. The restriction of EcoR124I against an incoming F lac plasmid was found to be restriction deficient (17). This was an unexpected and unusual situation: without mutagenesis, we had produced restriction-deficient mutants of R124 in the presence of the F plasmid (with which it is supposedly compatible) (see above for details on incompatibility). We applied the same restriction fragment analysis to plasmid R124, isolated as restriction deficient, from F lac16[R124] bacteria. Figure 4 shows that there are major DNA arrangements associated with plasmid R124 isolated from F lac16. Therefore, the definition of two plasmids coexisting as compatible plasmids was incorrect: the plasmids coexisted only if they had undergone the complex DNA rearrangements observed, and all plasmids isolated showed rearrangements, although many showed variation in the restriction fragment sizes, which is indicative of more than one type of rearrangement.

Two papers published in 1983 described these complex DNA rearrangements observed between plasmid R124 or R124/3 and F + in detail (17, 30). However, the outcome of this work is too detailed and complex to repeat here. One interesting observation was the transfer of the R-M genes between R124 and F +, replacing the phi region of F + with a region of R124 that carries the hsd genes of EcoR124I. The interactions between F + and R124 appear to involve regions of both plasmids that are involved in replication and/or incompatibility. In the F + plasmid, these rearrangements involve the phi region (which is also responsible for sensitivity of F + to phage MS2) (Fig. 5), and in turn, this region is transferred to R124 (21). It seems likely that such rearrangements are “enabling” the coexistence of these two plasmids and that the categorization of R124 as IncFIV may have been incorrect.

What is even more interesting with regard to these DNA rearrangements is the frequent switch in DNA specificity between the R124 and the R124/3 phenotypes (30), which led us to propose a mechanism based on the transposition or inversion of insertion sequences. It seems now that these changes involved an induced high-frequency recombination between the two plasmids that was recA independent and in some as-yet-unexplained way was linked to interactions between the origins of replication of these plasmids. Such a recombination event may also have occurred in E. coli BW9091 and led to the initial isolation of R124/3 by Hughes.

A restriction fragment map of R124 was produced in 1985 (13), but this failed to resolve the complex nature of the recombination events and DNA rearrangements that occurred when R124 and F + were present together in a bacterium. The location of the transposon Tn10, adjacent to one replication/incompatibility region of R124, could account for rearrangements involving the nearby R-M genes, but it is difficult to account for the changes involving EcoRI fragment E1 observed in both R124/3 and the Res− mutants, which are associated with an incompatibility region (Fig. 1). Our data from Tn7 transposition mutagenesis showed that HindIII fragment 4 was likely to carry the EcoR124I R-M genes (67), opening the way to isolate the genes. However, the lesson learned from these mapping studies was that plasmid DNA is subject to many more rearrangements than originally understood, and this may still not be fully recognized even today. It seems sensible that other incompatibility subgroups should be investigated in terms of DNA rearrangements.

**ISOLATING GENES FOR EcoR124I**

The first step toward isolating the genes for the EcoR124I R-M system was to isolate all of the HindIII fragments of
R124 in the vector pUR51 (76). The resulting recombinant plasmids that were identified as carrying the R124 R-M genes were named pCP1005 and pCP1009, respectively (67). The isolation of the genes took our research away from studies with plasmid R124 and focused us on the EcoR124I R-M enzyme.

R124 in the vector pUR51 (76). The resulting recombinant plasmids that were identified as carrying the R124 R-M genes were named pCP1005 and pCP1009, respectively (67). The R-M genes of the closely related plasmid R124/3 were also isolated but in a different vector (pUC18), producing the recombinant plasmid pUNG31. The isolation of the genes took our research away from studies with plasmid R124 and focused us on the EcoR124I R-M enzyme.

FIG. 4. DNA rearrangements of R124 and F+ associated with coexistence in the same bacterial cell. Plasmids R124 and F+ were isolated from E. coli J55[R124][F+] using acridine orange or ethidium bromide curing, and the plasmid DNA was isolated, cleaved with EcoRI, and separated on a 1% agarose gel. Significant DNA rearrangements are shown by arrows. EcoRI fragments E1, E3, E5, and E9 of R124 appear to be altered, although the small EcoRI fragment, visible in the R124/3 EcoRI digests and generated by a loss of methylation by the EcoR124I R-M enzyme, is absent in the Res- mutants. This suggests that fragment E3 undergoes some deletion that removes the “phantom” EcoRI site.

FIG. 5. The phi region of the F+ plasmid and acquisition of the hsd genes of R124. The phi region of F+ is responsible for sensitivity to bacteriophage MS2. This sensitivity is often lost following the coexistence of R124 and F+ in a single bacterium, and the associated DNA rearrangements (Fig. 3) suggest that the R-M genes of R124 are acquired by F+ replacing the Phi region. It may be significant to the mechanism of this process that the Phi region lies between two origins of replication of the F+ plasmid.
DNA SEQUENCE OF THE EcoR124 I-R-M GENES

The nature of the difference between the R-M system present on R124 and that present on R124/3 was one of changed DNA specificity, and it was Bickle’s group who showed conclusively what this change in DNA specificity involved. The two recognition sequences were determined by identifying restriction fragments of known DNA sequence that were methylated by the two enzymes (70). The difference between the two specificities was unexpectedly simple: EcoR124I was shown to recognize GAAn7RTCG, while EcoR124/3 was shown to recognize the same specific bases but with the inclusion of an extra nonspecific base in the spacer region of the recognition sequence (GAAn-RTCG). This was an exciting observation, as it supported the two-domain model for HsdS (originally proposed by Argos [4] and later extended in studies by Murray and others [24, 25]) in which a spacer region separates two target recognition domains (TRDs) that bind the specific nucleotides in the recognition sequence. It seemed likely that this spacer was extended in EcoR124/3. Bickle’s group also confirmed that there were no obvious structural differences between the two enzymes, which was also confirmed by their antigenic relationship with antibodies to EcoR124I recognizing all subunits of EcoR124II: the three subunits were of identical sizes (116 kDa for HsdR, 55 kDa for HsdM, and 43 kDa for HsdS) and were therefore similar to other type I R-M systems, but they were confirmed as being antigenically different from other type I R-M enzymes.

Therefore, on the basis of the antigenic differences and the biochemical properties, EcoR124I and EcoR124II (then still called EcoR124 and EcoR124/3, respectively) were classified by Price et al. as a new family of type I enzymes, type IC (69), yet another first for the R124 system.

Determination of the DNA sequence of the R-M genes isolated from R124 and R124/3 provided a clear insight into how the new DNA specificity of EcoR124/3 was generated. It was found that within the central conserved region of the hsdS genes of both systems, there is a repeat sequence of 12 bp (coding for TAEL); in EcoR124I, there are two 12-bp repeats, while in EcoR124/3, there are three 12-bp repeats (68). Therefore, the hypothesis that type I R-M enzymes might vary their DNA specificities through an increase in the length of the nonspecific spacer that separates the two TRDs of HsdS was confirmed. Further work by Gubler and Bickle not only confirmed this hypothesis but also showed the upper and lower limits of such expansion (32). That work supported the growing evidence that type I R-M systems were highly adaptable with respect to their abilities to generate novel DNA specificities, especially through recombination events including domain swapping (4, 24, 25).

FUNCTIONAL STUDIES WITH EcoR124I AND EcoR124II

As described above, the name attached to the variant of EcoR124I, which has a novel DNA specificity, was EcoR124/3. However, a survey of those working on R-M systems led to the suggestion that a more systematic nomenclature of all restriction enzymes was required (leading to a change for EcoR124/3 to EcoR124II), and this was eventually incorporated into the literature in 2003 (72). This systematic nomenclature also prevented the future use of StyRI24I, which had caused some confusion in the literature.

The observation that EcoR124I and EcoR124II differed in their DNA specificities due to a change in the length of the spacer region between the two TRDs of the enzyme renewed interest in the structure of the DNA-binding subunit of type I R-M enzymes.

The Central Conserved Region of HsdS

One of the first people to analyze the secondary structure of the DNA specificity subunits of type I R-M enzymes was Argos, who identified repeating, conserved domains within the HsdS subunit (4). Argos suggested that the conserved regions of HsdS “would provide the structural configuration necessary for interaction at the two specificity sites in the DNA recognition span.” However, it is not clear if he really identified the TRDs of HsdS as being the regions of maximum variation between the various HsdS sequences known at the time. In hindsight, perhaps one interesting comment from Argos was the idea of a “pseudodimer,” based on the idea of repeating domains. Described later in this review is the existence of such a pseudodimer consisting of a repeating half-HsdS subunit (1, 54). At the same time as Argos published his paper on repeating domains within the HsdS subunit of type I R-M enzymes, Fuller-Pace et al. (23) described a rearrangement of DNA specificity based on recombination between conserved regions within hsdS that “swapped” DNA specificities by swapping domains within the HsdS subunit. In fact, both Bickle’s and Murray’s groups were able to show that type I R-M enzymes can change DNA specificity through domain swapping (in which the TRDs were exchanged between familial members) (23, 25, 33). This eventually led to the idea of a circular structure of repeated domains within HsdS where the DNA recognition domains were nonconserved regions (TRD), and these were separated by structurally important conserved regions (51).

Although, to date, there is no actual protein structure for either HsdS or the MTase of the EcoR124I system, the recent crystallization of related HsdS subunits (12, 50) has confirmed the circular structure proposed for HsdS and allowed us to model in silico not only the HsdS subunit of EcoR124 but also the MTase of this enzyme (62). One of the interesting observations from this structural model was the ability of the model to allow a prediction for the mechanism of DNA binding. It was previously shown that upon specifically binding substrate DNA, the EcoR124I MTase underwent a major conformational shift, dramatically reducing the diameter of the protein (87). The structural model shows regions adjacent to the HsdM subunits, where a conformational shift involving one domain of the HsdM subunit would account for the observed changes in diameter (Fig. 6).

However, the EcoR124I system was to provide yet another exciting and unexpected observation. In our laboratory, we isolated a point mutation within the hsdS gene that showed a new DNA specificity (1), while Meister et al. reported a new DNA specificity that was discovered following insertional mutagenesis using the Tn5 transposon (56). Both new specificities were due to the production of a truncated HsdS subunit, which was able to dimerize, producing the recognition sequence GA...
An\textsubscript{TTC} in the case of the point mutation of EcoR124I, which was shown to produce a stop codon and hence the truncated subunit. The inverted symmetry of the sequence and the dimerization of the “one-half” HsdS subunit confirmed the circular model for the structure of HsdS proposed by Kneale (51) and expanded the view that type I R-M enzymes possessed the most adaptable system for evolving novel DNA specificity currently known (2).

Further studies of deletion mutants within the \textit{hsdS} gene of EcoR124I identified the limits of such changes to DNA specificity and showed that a minimum one-half HsdS subunit existed (2), which defines a key structural feature within the protein.

**Functional Studies of EcoR124I**

Work on the restriction activity of EcoR124I was always hampered by low enzyme concentrations. As with all type I R-M enzymes, the cleavage reaction does not show any turnover, and each cleavage event requires a single enzyme molecule. While the enzyme was available only in microgram quantities produced from plasmid R124, or recombinant plasmid pCP1005, grown in large volumes, little biochemistry was carried out on the enzyme.

In 1996, Dreier et al. (16) conducted the first systematic study of DNA cleavage by EcoR124II by using a genetically engineered vector carrying a single recognition sequence for the enzyme. They showed that cutting was a two-stage event, with nicking occurring first followed by a double-strand cleavage event, suggesting that each HsdR subunit was responsible for nicking one strand of the DNA and that following cleavage, no further cleavage of the now-linear DNA occurred. They also demonstrated that the cleavage of linear DNA required two sites in \textit{cis} on the DNA substrate. The orientation of these two sites was unimportant to the DNA cleavage event, and EcoR124I recognition sites available in \textit{trans} on short oligoduplex DNA made no difference to the cleavage reaction (although they did stimulate the ATPase activity) (15).

Interestingly, when a two-site plasmid was used for the DNA cleavage studies, if the separation of the two sites was greater...
than 40 bp, multiple cleavage events occurred, as expected, eventually leading to a smear of DNA on a gel. However, if the separation of the two sites was less than 40 bp, only a single cleavage site occurred (16). On linear DNA with two sites, only a single cleavage event occurred when the separation of the two recognition sequences was less than 23 bp. The conclusion was that the binding of the enzyme to one site inhibited the binding to the second site, and this was confirmed a few years later by DNA footprinting studies (57, 58).

Increased availability of the EcoR124I R-M enzyme (see below for further details) was provided by the availability of recombinant plasmids that produced both the independent MTase (63) (Fig. 7) and the HsdR subunit (from a compatible vector using a two-plasmid system (98), but an early observation from that work was that EcoR124I readily, but unexpectedly, dissociated into a mixture of a R1M2S1 complex (R1-complex) and free HsdR (41, 43).

At this stage, although the mechanism of DNA cleavage by type I R-M enzymes was reasonably well understood, there were still clear gaps in our knowledge of the precise mechanism. The above-described data showed that a nick in the DNA did not inhibit restriction; therefore, the question of what inhibits the translocation/cleavage process was asked. The question was addressed in a series of experiments using the enzyme EcoR124II. Dreier et al. (16) were able to show that a non-covalently-bound protein, such as the binding of LacR to a DNA sequence on the DNA, did not induce DNA cleavage but that changes to the DNA topology, in the form of a Holliday junction, were able to induce DNA cleavage by EcoR124II, indicating that it halts translocation (45). Janscak et al. went on to suggest that any topological barrier to translocation, including the translocation of a whole circle of DNA on a single-site plasmid, would lead to the blockage of translocation and, thus, the observed highly efficient cleavage of the DNA. Furthermore, that paper was able to confirm bidirectional translocation by the EcoR124II holoenzyme. However, one problem remained unresolved by that paper, which was that DNA cleavage of a single-site plasmid did not appear to require a cooperative interaction with a second enzyme, as suggested by the model described by Studier and Bandyopadhyay (85) for two-site substrates.

A more complete analysis of the DNA cleavage reaction was provided by Jindrova et al. (48), who analyzed the nature of the DNA ends produced by a variety of type I restriction endonucleases and suggested a model for cleavage for the EcoR124I enzyme, which takes into account the dissociation of the holoenzyme. In this model, the HsdR subunit can produce only a single cleavage of the phosphodiester backbone of the DNA but can cooperate with another HsdR subunit to produce full DNA cleavage, producing DNA with overhanging ends of single-stranded DNA. On a single-site plasmid, cleavage requires the association of HsdR, present in solution, to the DNA-bound R1-complex/hoenzyme, allowing the HsdR-HsdR interactions required to complete the cleavage of both DNA strands. However, on a two-site plasmid, cleavage is enabled through HsdR-HsdR interactions resulting from the “collision” of two translocating enzymes. (Reprinted from reference 48 by permission of Oxford University Press.)
duce a Res− Mod− phenotype, because the mutation should destroy the DNA-binding capability of the HsdS subunit (27–29, 37). However, for a number of years, there was speculation among those people studying these systems that it might be possible to alter HsdS-HsdR protein-protein interactions through mutagenesis and, as a consequence, produce a “nonclassical” phenotype: Res−/HsdS Mod−/HsdR. In fact, such mutants were first isolated for the EcoKI system as temperature-sensitive mutations (38), but subsequent to this, nonclassical phenotypes of hsdS were also observed for EcoR124I (93), and one such mutation was indeed shown to alter HsdS-HsdR protein-protein interactions (92) although in an unexpected manner. It was shown that the Trp212Arg mutation prevented DNA binding for the MTase, but the addition of HsdR overcame this problem and allowed DNA binding at a reduced level. We believe that this may reflect the nature of the DNA-binding event itself, where it has been shown that a large conformational change occurs in the MTase upon binding of DNA (87). If the mutant MTase is altered in the ability to undergo this conformational change and “wrap” around the DNA, the presence of HsdR may “push” the MTase back into the required conformation.

CELLULAR LOCALIZATION OF EcoR124I

For the bacterial cell harboring, or acquiring, an R-M system, there is always an important requirement that the restriction activity is controlled and cannot readily function without prior modification activity: the host chromosome must be fully methylated at the recognition sites of the R-M endonuclease to avoid cleavage and subsequent cell death.

Type I R-M enzymes have an inbuilt control based on a switch between methylation activity and endonuclease activity, which occurs when the enzyme “reads” the methylation status of the recognition site. Only unmethylated DNA allows an ATP-dependent switch to endonuclease activity (97). It was proposed that the dissociation of the EcoR124I holoenzyme to a restriction-deficient R1-complex may also be an additional means of controlling these two opposing functions (18), which may be required because of the transmissibility of plasmid R124. Other mechanisms for the control of restriction and modification have been described, including the phenomenon of restriction alleviation (RA) (9), which involves the proteolysis of the HsdR subunit of the EcoKI system (14). The mechanism by which RA occurs appears to be family dependent, and the proteolysis of HsdR observed for EcoKI was not observed with EcoR124I. The explanation for this was unclear for a number of years until a series of elegant experiments by Holubova et al. (35, 36) showed first that the EcoKI holoenzyme was localized within the cell membrane of the host and later that the precise localization within the cell was also dependent on the type I family to which the enzyme belongs, providing a possible explanation for the different modes of RA activity. The EcoR124I holoenzyme was found to be particularly exposed on the periplasmic side of the cytoplasmic membrane, with EcoKI partially exposed in the periplasm, but EcoAI is not exposed to the periplasm at all. This may provide another interesting mechanism for the separation of the R-M functions. While it is far from clear what role the periplasmic exposure of the R-M enzyme would achieve, localization to the cell periphery may hasten endonuclease activity, guarding against incoming DNA. This is to some extent supported by the observation that EcoKI can act as a molecular motor, binding to the first 850 bp of injected DNA and translocating bacteriophage T7 DNA into the cell (26), suggesting localization of the enzyme at the periphery of the cell.
SINGLE-MOLECULE STUDIES WITH THE EcoR124I MOLECULAR MACHINE

In recent years, there have been a number of important papers describing single-molecule studies of DNA-binding enzymes, and for this review, we point to our own work involving single-molecule studies with this enzyme carried out in collaboration with Dekker’s group at Technology University of Delft. This exciting work points toward potential uses for this type I R-M enzyme as a nanoactuator within a biosensor (64) in the field of bionanotechnology, something we could hardly have believed possible in the early 1970s when this work started. In fact, in the early 1970s, it was impossible to imagine that it would be possible to measure DNA translocation for a single molecule let alone show that the motor could pull a micrometer-sized magnetic bead over a distance of several micrometers, an amazing capability for a motor measuring only a few nanometers in diameter (90).

The first direct measurements of the speed of this DNA translocation were carried out using bulk biochemistry based on strand displacement of a triple-helix-forming oligonucleotide (20) and showed that the motor is able to move DNA, through the DNA-bound complex, at approximately 400 bp s$^{-1}$. This bulk biochemistry was later confirmed by single-molecule studies (80), which also allowed the speed to be refined, and was shown to be 550 bp s$^{-1}$ for each active molecule. This involved the use of a magnetic tweezer setup (Fig. 9) in which an inverted microscope was used to determine the height of a magnetic bead attached to the DNA and located above the objective. By stretching the DNA using an external magnetic field, the motor activity, pulling the bead toward the surface, could subsequently be determined from the height of the magnetic bead above the objective (31, 99; J. Zlatanova and S. H. Leuba, presented at the Annual West Coast Chro-

FIG. 10. Proposed nanoactuator device. This image, captured from an animation of the proposed device (http://www.bionano-switch.info/archive.htm), shows a microfluidic channel of a Lab-on-a-Chip device in which magnetic sensors have been located on the bottom of the flow cell. Above these sensors, a DNA molecule is attached, which is stretched vertically through the effect of an external magnetic field on the DNA-attached magnetic bead. The incoming EcoR124I motor is shown in dark green, with the MTase component bound at the recognition site on the DNA (vertical tube), while the motor subunits (horizontal tubes) bind at the MTase and translocate the DNA. Translocation “pulls” the magnetic bead toward the sensor, which induces an electronic output (illustrated by a red sensor) acting as a “molecular dynamo” that links biological input to silicon-measured output. The inset shows an example of the data obtained from the magnetic tweezer setup (y axis, bead height; x axis, time), showing vertical movement of the bead due to translocation and thermal motion as a random movement of the bead throughout.
matin and Chromosomes Conference, Pacific Grove, CA, 2002). This work also confirmed that each motor (HsdR subunit) was able to work independently, with the holoenzyme acting as a bidirectional motor, but perhaps, unexpectedly, that the motor subunit of EcoR124I could also readily dissociate from the translocating complex. This provides a mechanism for the control of the direction of translocation, in vitro (78), by the appropriate location of the recognition sequence for EcoR124I on surface-attached DNA, locating the recognition sequence near the surface and making use of steric hindrance to prevent the assembly of the R_2-complex. Such a situation is likely to find uses in the proposed nanoactuator device (Fig. 10 and see below). It was this dissociation, which leads to the resetting of the motor, that accounts for the difference in measured speed between bulk measurements and single-molecule measurements. The R_1-complex was also shown to be a molecular motor but was unidirectional and slightly less processive (80).

Single-molecule work was also able to elucidate the nature of the events leading to the initiation of translocation. There is a major topological problem that a DNA translocase has to overcome prior to DNA translocation, which is due to the persistence length of DNA (~50 nm) (88). Footprinting studies have shown that the endonuclease does not “reach” over such a distance (57, 66), yet the enzyme can somehow “bend” adjacent DNA to initiate translocation despite the large energy input required. Studies using scanning probe microscopy by van Noort et al. (90) demonstrated that the DNA in the initial complex (captured using a nonhydrolyzable analogue of ATP) was present as a single-stranded DNA bulge, one strand of which was visible in the scanning probe microscopy and accessible to single-strand-specific nucleases and the other of which was buried in the enzyme. Therefore, the motor is able to “melt” DNA adjacent to its binding site and hold these strands apart prior to translocation.

Dynamics, kinetics, and equilibrium studies have now thoroughly characterized EcoR124I as a biological molecular motor capable of pulling DNA and an attached micrometer-sized bead over a distance of several micrometers (78–80, 83, 84). The ability to guide the directionality of the translocation using gaps in the DNA and the ability of the enzyme to translocate past nicks suggest that the motor will have potential uses in the bionanotechnology sector (81, 84).

THE FUTURE: EcoR124I AS A NANOACTUATOR

The possibility of single-molecule analysis of biological enzymes would never have been predicted in the 1960s or 1970, but the possibility of a commercial use for a type I R-M enzyme has been a long-term dream, which may yet be realized.

Perhaps the greatest surprise from the work with EcoR124I has been the success of these single-molecule studies; similar work was attempted with the closely related EcoKI system (D. Dryden and D. Bensimon, personal communication) but with no results. This may reflect a fundamental difference between EcoKI and EcoR124I in their modes of operation: the EcoKI enzyme has been shown to dimerize during translocation (7), but this is not the case for EcoR124I. While it has not been shown that dimerization is required for translocation by EcoKI, and the role of the dimerization event is not clear, although it certainly happens, it may be that single-molecule studies of translocation by EcoKI were affected by the inability of such assemblies to occur in the single-molecule magnetic tweezer setup.

Following two European-funded collaborations, we have clearly demonstrated that the molecular motor can move a magnetic bead attached to a DNA substrate, and we have demonstrated that the motor can “work” for as long as 1,800 min and that the micrometer-sized magnetic bead can be moved over a distance of several micrometers. We are now on the verge of incorporating this molecular motor into a microchip device using the motor to pull a magnetic bead toward a sensor, which in turn is capable of outputting an electronic signal, a device that links the biological world of ATP, as a fuel, with the silicon world of computers through this electronic output, a single-molecule molecular dynamo (Fig. 10). This molecular dynamo can measure events from single molecules of DNA, providing a highly sensitive biosensor for detecting events that interrupt translocation events. Such a transducer is likely find uses in biosensing, drug discovery, and many, as-yet-unknown applications.

ACKNOWLEDGMENTS

We thank Clive Price and Pavel Janscak for their input into the drafting of this article.

REFERENCES

García, L. R., and I. J. Molineux.
Hedges, R. W., and N. Datta.
15: 27. YOEULL AND FIRMAN MICROBIOL. MOLL. BIOI. REV.