The Type IIB restriction endonucleases

Jacqueline J. T. Marshall1 and Stephen E. Halford

The DNA-Protein Interactions Unit, Department of Biochemistry, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, U.K.

Abstract

The endonucleases from the Type IIB restriction-modification systems differ from all other restriction enzymes. The Type IIB enzymes cleave both DNA strands at specified locations distant from their recognition sequences, like Type IIS nucleases, but they are unique in that they do so on both sides of the site, to liberate the site from the remainder of the DNA on a short duplex. The fact that these enzymes cut DNA at specific locations mark them as Type II systems, as opposed to the Type I enzymes that cut DNA randomly, but in terms of gene organization and protein assembly, most Type IIB restriction-modification systems have more in common with Type I than with other Type II systems. Our current knowledge of the Type IIB systems is reviewed in the present paper.

Types and subtypes of R-M (restriction-modification) systems

Bacteria defend themselves against bacteriophage and other external DNA by using R-M systems to destroy foreign DNA as it enters the cell while protecting the host DNA [1,2]. These systems display two activities: a modification MTase (methyltransferase) that recognizes a specific DNA sequence and transfers methyl groups from AdoMet (S-adenosylmethionine) on to particular bases in the sequence, one in each strand, to eventually generate DNA methylated in both strands; and an REase (restriction endonuclease) that recognizes the same sequence as the MTase and then cleaves the DNA, but only if the target is not methylated in either strand [3]. The MTase thus protects the host DNA from the REase, even after its semi-conservative replication when only one strand is methylated, while the REase cleaves any DNA entering the cell that lacks the host’s methylation pattern [2].

The vast majority of the bacterial genomes sequenced to date contain candidate genes for multiple R-M systems [4], over 20 in the case of Helicobacter pylori, many of which code for functional R-M systems [5]. Despite their multiplicity, R-M systems can be categorized into four main types [4]. Types I and II are the most common: only these are discussed in the present paper.

Most Type I systems contain three contiguous genes in the order hsdR-hsdM-hsdS, that each code for one polypeptide: HsdS, which recognizes the target DNA sequence; HsdM, which methylates the target sequence; and HsdR, which carries out the DNA cleavage reaction, but at non-specific sequences often kilobases away from the recognition site [6]. The HsdR subunit also possesses an ATP-dependent DNA-translocation motif symptomatic of a DNA helicase [7]. The three subunits assemble to form a R2M2S unit with both MTase and REase activities: the former functions optimally on hemimethylated DNA, whereas the latter acts only on completely unmethylated DNA [6]. The single HsdS subunit fulfils all of the DNA specificity functions. The recognition sites for Type I systems consist of two specified elements 3 or 4 bp long separated by a short segment of 6–8 bp of unspecified DNA, and HsdS contains two TRDs (target-recognition domains) that each contact one of the specified elements [6].

Nearly all Type I systems follow this pattern, although they can be placed into separate complementation groups known as Types IA, IB, IC and so forth [6]. However, some differ from the common pattern by carrying the functions for DNA recognition, methylation and cleavage in a single polypeptide [8,9].

The hallmark of the Type II R-M systems is that the REase cleaves both DNA strands at fixed positions, either inside or at fixed locations outside (but generally within 20 bp) of the recognition site [4]. This contrasts with the random, often distant, locations by Type I REases. Consequently, only the Type II REases are widely used as tools for the dissection, analysis and reconstruction of DNA [10]. Most Type II systems feature separate REase and MTase proteins, often a homodimeric REase and a monomeric MTase that both recognize a palindromic recognition sequence 4–8 bp long [3,4,11,12]. Such systems encompass many well-known REases such as EcoRI, EcoRV and BamHI [11,12]. In these cases, the dimeric REase binds symmetrically to its palindromic recognition sequence so that the active site from one subunit is placed to cut the target phosphodiester in one strand whereas that from the second subunit acts on the complementary strand. The monomeric MTase has only one active site, but its main function is to transfer a single methyl group on to hemimethylated DNA.

The Type II REases are, however, a remarkably diverse group of proteins, much more diverse than the Type I systems, and can be categorized into a series of markedly different

Key words: DNA communication, endonuclease, methyltransferase, phosphodiester hydrolase, recognition sequence, restriction-modification system.

Abbreviations used: 3D, three-dimensional; AdoMet, S-adenosylmethionine; MTase, methyltransferase; REase, restriction endonuclease; R-M, restriction-modification; TRD, target-recognition domain.

1To whom correspondence should be addressed (email jacqui.marshall@bristol.ac.uk).
Machines on Genes: Enzymes That Make, Break and Move DNA and RNA

Table 1 | Subtypes of Type II REases
Table modified from [13], with permission of Oxford University Press.

<table>
<thead>
<tr>
<th>Subtype*</th>
<th>Defining feature</th>
<th>Examples</th>
<th>Recognition sequence†</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Asymmetric recognition sequence</td>
<td>FokI</td>
<td>GGATG (9/13)</td>
</tr>
<tr>
<td>B</td>
<td>Cleaves both sides of target on both strands</td>
<td>BcgI</td>
<td>(10/12) CGANNNNNNTGC (12/10)</td>
</tr>
<tr>
<td>C</td>
<td>R and M functions in one polypeptide (symmetric or asymmetric target)</td>
<td>GsuI</td>
<td>(7/13) GAYNNNNNRTC (14/9)</td>
</tr>
<tr>
<td>E</td>
<td>Two targets: one cleaved, one an effector</td>
<td>Nael</td>
<td>GCCG6GC</td>
</tr>
<tr>
<td>F</td>
<td>Two targets, both cleaved co-ordinately</td>
<td>FokI</td>
<td>GGATG (9/13)</td>
</tr>
<tr>
<td>G</td>
<td>Affected by AdoMet (symmetric or asymmetric target)</td>
<td>Eco57I</td>
<td>CTGAAG (16/14)</td>
</tr>
<tr>
<td>H</td>
<td>Similar to Type I gene structure (symmetric or asymmetric target)</td>
<td>Ahdl</td>
<td>GACNNN\nNCGT</td>
</tr>
<tr>
<td>M</td>
<td>Subtype IIP or IIA. Require methylated target</td>
<td>DpnI</td>
<td>GmαATC</td>
</tr>
<tr>
<td>P</td>
<td>Symmetric target and cleavage sites</td>
<td>EcoRV</td>
<td>GAT\nLAC</td>
</tr>
<tr>
<td>S</td>
<td>Asymmetric target and cleavage sites</td>
<td>BglII</td>
<td>GCCNNNN\nNGGC</td>
</tr>
<tr>
<td>T</td>
<td>R genes are heterodimers (symmetric or asymmetric target)</td>
<td>Bpu10I</td>
<td>CTCAAGC (−5/−2)</td>
</tr>
</tbody>
</table>

*Note that not all subtypes are mutually exclusive, e.g. BcgI (in bold italics) is the archetype of the Type IIB class, but is also in subtypes IIA, IIC, IIF, IIG and IIH.
†In cases where cleavage occurs outside the recognition sequence, the sites of cleavage are noted by two numbers: the first indicates the number of bp between recognition and cleavage sites in the strand shown and the second in the complementary strand. Cleavage sites within recognition sequences are marked by ↓. N, any nucleotide; Y, a pyrimidine; R, a purine.

subtypes (Table 1) [13]. The criteria for categorization include, among others, the following.

(i) The nature of their recognition sequences, whether symmetric or asymmetric (e.g. Type IIA, cf. Type IIP).

(ii) The positions of DNA cleavage, either at or away from the recognition site (Type IIP cf. Type IIS or IIB) and in the latter cases, whether they cut on one or on both sides of the site (Types IIS and IIB respectively).

(iii) The separation or otherwise of their REase and MTase polypeptides (the Type IIC systems carry both activities in one polypeptide).

(iv) Their co-factor requirements: all bar a few [14,15] need Mg2+ for nuclease activity, but the Type IIG group also require AdoMet. None needs the ATP that the Type I enzymes use.

(v) Their mode of action, whether they need to interact with two recognition sites before cutting the DNA at either one or both sites (Types IIE and IIF respectively).

These criteria are not mutually exclusive, so one Type II REase can belong to several subtypes. For example, BcgI falls into subtypes IIA, IIB, IIC, IIF, IIG and IIH (Table 1).

**Type IIB restriction endonucleases**

Whereas the Type IIS endonucleases cut both strands of the DNA on one side of their asymmetric sites, the defining characteristic of the Type IIB enzymes is that they cut both strands on both sides of their recognition sites [13]. Type IIB endonucleases thus excise from the remainder of the DNA a small fragment, 20–33 bp long depending on the enzyme, that contains the recognition sequence (Table 2). BcgI is the archetypal, and the best studied, Type IIB enzyme [17–21].

With one exception, NmeDI [22], their recognition sites are bipartite sequences, containing two specified elements 2–5 bp long separated by 4–7 unspecified bp (Table 2). The recognition sites for Type IIB enzymes (Table 2) thus resemble the bipartite sites for Type I systems [6]. In some cases, the two specified sets of bp are related by symmetry (e.g. AlfI and BpiII), but others are asymmetric (e.g. AloI and BcgI) and even of different lengths (e.g. BaeI and BsaXI). NmeDI differs significantly from all other Type IIB systems and will be discussed separately. Unless noted otherwise, our generalizations about Type IIB enzymes exclude NmeDI.
Examples of Type IIB restriction endonucleases

A total of 11 of the 23 Type IIB R-M systems identified to date (http://rebase.neb.com/cgi-bin/doublist on 1 August 2009) are listed.

### Table 2 | Examples of Type IIB restriction endonucleases

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Source organism</th>
<th>Recognition sequence</th>
<th>Excised DNA (bp)</th>
<th>Polypeptide chains</th>
<th>AdoMet</th>
<th>Number of sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>AlfI</td>
<td>Acinetobacter lwoffi BH32</td>
<td>(10/12) GCAN₆TGC (12/10)</td>
<td>32</td>
<td>ND</td>
<td>✓</td>
<td>2</td>
</tr>
<tr>
<td>AloI</td>
<td>Acinetobacter lwoffi KS-4B</td>
<td>(7/12) GAACN₆TCC (12/7)</td>
<td>27</td>
<td>1</td>
<td>×</td>
<td>2</td>
</tr>
<tr>
<td>Bael</td>
<td>Bacillus sphaericus</td>
<td>(10/15) ACN₆GAYT (12/7)</td>
<td>28</td>
<td>2 (ND)</td>
<td>✓</td>
<td>2</td>
</tr>
<tr>
<td>Bcgl</td>
<td>Bacillus coagulans</td>
<td>(10/12) CGAN₆TGC (12/10)</td>
<td>32</td>
<td>2 (A₂B₁)</td>
<td>✓</td>
<td>2</td>
</tr>
<tr>
<td>BplⅡ</td>
<td>Bacillus pumilus</td>
<td>(8/13) GAGN₆CTC (13/8)</td>
<td>27</td>
<td>2 (AB)</td>
<td>✓</td>
<td>1</td>
</tr>
<tr>
<td>BscXI</td>
<td>Bacillus stearothermophilus</td>
<td>(9/12) ACN₆TCC (10/7)</td>
<td>27</td>
<td>ND</td>
<td>✓</td>
<td>2</td>
</tr>
<tr>
<td>CjeI</td>
<td>Campylobacter jejuni</td>
<td>(8/14) CCAN₆GT (15/9)</td>
<td>28</td>
<td>1</td>
<td>✓</td>
<td>ND</td>
</tr>
<tr>
<td>CspCI</td>
<td>Citrobacter species 2144</td>
<td>(11/13) CAAN₆GTGG (12/10)</td>
<td>33</td>
<td>ND</td>
<td>✓</td>
<td>2</td>
</tr>
<tr>
<td>HaeVI</td>
<td>Haemophilus aegyptius</td>
<td>(7/13) GAYN₆RTC (14/9)</td>
<td>27</td>
<td>1</td>
<td>×</td>
<td>ND</td>
</tr>
<tr>
<td>NmeDI</td>
<td>Neisseria meningitidis</td>
<td>(12/7) RCCGGY (7/12)</td>
<td>20</td>
<td>2, separate R and M</td>
<td>×</td>
<td>2</td>
</tr>
<tr>
<td>PpI</td>
<td>Pseudomonas putida</td>
<td>(7/12) GAACN₆TCC (13/8)</td>
<td>28</td>
<td>1</td>
<td>×</td>
<td>2</td>
</tr>
</tbody>
</table>

*From [4]. For some enzymes, under some conditions, the point of cleavage varies by 1 or 2 bp from that indicated.
†The length is that of the duplex segment of the excised fragment. Except for NmeDI, the excised fragment carries single-strand extensions of 2–6 bases (depending on the enzyme) at the 3’ termini of both top and bottom strands. NmeDI leaves 5-base extensions at both 5’ termini.
‡The number of different types of polypeptide chain is indicated and, in cases containing two different chains, the ratio of the two chains. ND, not determined for either the number or ratio of chains.
§Enzymes that require AdoMet for nuclease activity are indicated with ticks and those that do not with crosses.
∥The number of recognition sites on a DNA substrate that the enzyme in question needs to display its full activity is indicated (data taken from [16] and [22]). ND, enzymes whose activities on DNA substrates with one or two sites have not been determined to date.

### Two polypeptide systems

Some Type IIB systems contain two polypeptide chains, others a single chain (Table 2). BcgI [18], BaeI [23] and BplⅡ [24] exemplify the former. BcgI is encoded by two overlapping genes, bcgIA followed by bcgIB, that each express a single protein: BcgIA carries both REase and MTase motifs; BcgIB confers DNA sequence specificity [18].

The various activities of BcgI were assigned to specific regions of each polypeptide by sequence alignments and mutational studies [18,20]. Mutations that affected only nuclease activity were all within the N-terminal portion of BcgIA and the nuclease active site, P⁵⁸EX₁₂EDK, was identified by site-directed mutagenesis [20]. The C-terminus of BcgIA is homologous with certain adenine MTases: the HsdM subunits of Type I systems and the MTases from some Type II systems [18]. The BcgIB polypeptide resembles the HsdS subunits of Type I systems in that it contains two putative TRDs, each of which was postulated to recognize one of the specified 3 bp elements in the recognition sequence [20]. This Type IIB enzyme thus has the same genetic and polypeptide organization as a Type I system, except that it carries the HsdR and HsdM subunits fused into a single polypeptide and that the fusion lacks the DNA motor functions of HsdR.

Neither BcgI subunit is catalytically active by itself, nor capable of binding DNA, but a combination of A and B subunits at a 2:1 A/B ratio is active as a MTase or as a REase, depending on the methylation status of the substrate [18,19]. In terms of Type I functionalities, the A₂B₁ unit of BcgI is equivalent to (R-M)₂S₁, which is effectively the same as the R₂M₂S₁ organization of the multisubunit Type I systems.

The two BcgIA subunits are perhaps located either side of the B protein so that once BcgIB has bound to its discontinuous sequence, one of the A subunits is placed to methylate the target adenine in one of the bipartite elements while the other modifies the specified adenine in the opposite strand of the second element. However, each A subunit seems to have only one catalytic centre for nuclease activity, yet this enzyme cuts four phosphodiester bonds at each copy of its recognition site, as it makes double-strand breaks on both sides of its site [17]. Hence, if BcgI is active as an A₂B₁ unit, then it may have to use each A subunit twice, to cut one particular phosphodiester bond before repositioning its catalytic centre on a second bond, either on the opposite side of the site or in the opposite strand. Although most Type II REases use one catalytic centre on each strand, the Type IIS endonuclease BfiI provides a precedent for switching a single active site between strands [25]. Alternatively, the A₂B₁ unit bound to the recognition site may have to recruit one or more additional A₂B₁ proteins to give an assembly with a sufficient number of nuclease sites to cut four phosphodiester bonds. This second possibility resembles the mechanism of
the monomeric Type IIS endonuclease FokI [26,27]. The monomer of FokI bound to its recognition site positions its single active site on the antisense strand, but has no activity until it associates with a second monomer; the latter attacks the sense strand [28]. Gel filtration of BcgI in free solution yielded an apparent Mr halfway between that for a single A2B2 unit and that for a dimeric unit [21]. This might reflect protein associations, but it could alternatively be due to the intrinsic inability of gel filtration to give true Mr values for proteins with large axial ratios.

BplI is another Type IIB system that contains two polypeptides: its A and B peptides are similar in size to BcgI A and BcgIB respectively, but they appear to be present in a 1:1 rather than a 2:1 ratio [24]. BplI displays both REase and MTase activities at a specific DNA sequence, but it is not yet known which functions are governed by which chain.

NmeDI provides a further example of a two-polypeptide Type IIB R-M system, in this case two entirely separate proteins: the NmeDI REase [22] and the NmeDI MTase [29]. The NmeDI system is thus organized like a conventional two-protein Type II system [3], unlike the multisubunit creatures of the Type I and other Type IIB systems. The REase appears from gel filtration to be a tetramer, but many Type II REases are tetrhomers, particularly those in the Type IIF group that cleave two recognition sites concertedly (Table 1 [30,31]). Even though the NmeDI REase makes double-strand breaks on both sides of its site and so has to be classified as a IIB system, it is distinctive in numerous respects: it has a continuous recognition sequence, as opposed to the discontinuous bipartite sites for all other Type IIB systems (Table 2); the NmeDI MTase modifies cytosine residues [29], whereas all other Type IIB enzymes act at adenine residues; it alone leaves 5' rather than 3' single-strand extensions.

One-polypeptide systems

Several Type IIB systems, such as AloI [32], CjeI [4,33] and HaeIV [34], carry their DNA methylation, cleavage and sequence-recognition functions all within a single polypeptide. In these cases, sequence alignments and/or deletion mutations indicate: (i) the N-terminus of the protein carries the nuclease activity; (ii) the central region contains several motifs characteristic of an adenine MTase; and (iii) the C-terminus is, at least in the case of AloI, homologous with the HsdS subunits of Type I systems and with BcgIB, so confers DNA sequence specificity. The two TRDs in this segment of AloI have been shown by ‘domain-swap’ experiments with similar one-polypeptide systems to each recognize one of the specified elements in the bipartite recognition sequence [35]. Therefore AloI, CjeI and HaeIV have genetically the same overall order of motifs as BcgI and the classical Type I systems, HsdR followed by HsdM and then HsdS. The one-polypeptide Type IIB systems are thus analogous, apart from the lack of a motor domain, to the one-polypeptide Type I schemes [8,9].

A one-polypeptide system must have its DNA-binding and both REase and MTase functions present at 1:1 ratios, provided the peptide contains one active site for each catalytic function. Hence, if one C-terminal DNA-binding domain covers the complete target sequence, it will be accompanied by only one active site for the MTase and one for the nuclease. Yet at their recognition sites, these enzymes methylate two adenine residues and cut four phosphodiester bonds. The single polypeptides thus need to associate to oligomers. Gel filtration of AloI and HaeIV suggest that the former may exist as a tetramer in solution and the latter a dimer [32,34]. But this arrangement then leaves a surplus of DNA-binding domains, when only one is needed to bind the recognition site.

Nuclease reactions

For BcgI, both the REase and the MTase require Mg2+ and AdoMet as cofactors [17]. Most mutations in the AdoMet-binding pocket prevented both MTase and REase activities, presumably by blocking AdoMet binding, but one inhibited only the MTase [19]. This mutation may therefore still allow AdoMet binding while disallowing the methyl-transfer reaction. The nuclease activity thus seems to depend on a conformational effect induced by AdoMet binding to the protein. Within an R-M system, an REase that needs AdoMet possesses a potential safeguard in that, should the AdoMet concentration fall too low for full methylation of the host DNA, then the endonuclease stops functioning [18]. Conversely, concomitant methylation can prevent complete cleavage by the REase [18,23,24]. The REases in the other two-polypeptide IIB systems require AdoMet in addition to Mg2+ (Table 2). In contrast, none of the one-polypeptide enzymes needs AdoMet for nuclease activity, except for CjeI [33].

During their reactions on plasmids with one or more copies of the relevant recognition site, the Type IIB REases produce virtually no nicked DNA cut in just one strand [16,36]. Instead, they proceed directly to products carrying one or more double-strand breaks, but this leaves open the question of whether the double-strand break(s) occur on one or on both sides of the site and, if on one, whether a particular side is preferred over the other. The HaeIV [34] and the NmeDI [22] REases act sequentially. They produce first an intermediate cut on one (random) side of the recognition site before generating the final product cut on both sides. In contrast, BcgI generated virtually no DNA cut on only one side and instead gave the final product directly, the excised 32 bp fragment carrying the recognition site (J.J.T. Marshall and S. Ganguly, unpublished work). BcgI thus seems to cut four phosphodiester bonds at each recognition site in a single highly concerted reaction.

Long-range communications

The Type I and many of the Type II REases need to interact with two copies of their cognate recognition sequence before cutting the DNA [37], particularly the Type IIE and IIF systems (Table 1). One method to determine whether an
enzyme needs two sites is to compare the rate at which it cleaves a DNA with two copies of its recognition site with that on a DNA with one target site [14,22,27,30,31].

An enzyme that requires two sites is likely to cleave the two-site substrate more rapidly than the one-site DNA, as interactions between sites in cis, on the same molecule of DNA, are favoured over interactions in trans, between sites on separate molecules of DNA [37].

When BcgI was tested for a fixed time against plasmids with one or two BcgI sites, more of the two-site plasmid was cleaved than the one-site plasmid [21]. In subsequent experiments, the rates of BcgI reactions on plasmids constructed with one or two identical BcgI sites (with respect to flanking and spacer sequences) were compared from the declines in the concentration of each substrate as a function of time [16]. The plasmid with two BcgI sites was cut far more rapidly (∼6-fold) than that with one site. Hence BcgI displays its full activity only after interacting with two copies of its recognition sequence, and its slow reaction on a DNA with one site almost certainly involves an interaction in trans, with the protein bridging sites on separate DNA molecules [36]. Moreover, during its reaction on the two-site plasmid, BcgI liberated virtually none of either the nicked DNA with single-strand breaks at one or both sites, nor the intermediate with one or more double-strand breaks at an individual site. Instead, the first species formed in any significant yield was the DNA with at least one double-strand break at both sites [16]. BcgI therefore acts in a highly concerted fashion, not only on both sides of an individual site, but also at two separate sites, in the manner of a Type IIF restriction enzyme [30,31]. But, although a normal Type II enzyme such as SfiI binds two copies of its recognition site and then proceeds to cut four phosphodiester bonds before dissociating from the DNA [30], BcgI appears to cut a total of eight phosphodiester bonds per reaction.

Several other Type IIB nucleases were tested on plasmids with one or two copies of their target sequence: AlhI, AolI, BaeI, BpiI, BsaXI, CspCI, FalI, PpiI, PsrI and NmeDI [16,22]. This series includes some that share with BcgI a two-polypeptide organization (BaeI and BpiI), some in the one-polypeptide group (AolI and PpiI), and one with separate REase and MTase proteins (NmeDI). With one exception, they all cleaved their two-site substrates more rapidly than their one-site substrates, sometimes very much more rapidly: in the case of CspCI, 40 times more rapidly [16]. In contrast, BpiI cleaved both substrates at the same rate under all conditions tested [16]. Although most Type IIB enzymes have asymmetric recognition sites, BpiI has a symmetric site, but AlhI, NmeDI (Table 2) and FalI (not shown) also have symmetrical sites, yet these all need two sites. BpiI does, however, appear to have a unique organization: two peptides, but in a 1:1 rather than 2:1 ratio [24]. Nevertheless, regardless of the reason that BpiI differs from the others, the Type IIB enzymes generally need to interact with two cognate sites before cutting DNA.

Communications between distant sites in DNA can occur either in one dimension, by tracking along the DNA between the sites, or through 3D (three-dimensional) space, upon the juxtaposition of the sites as a result of the random motion of the DNA chain [37]. The one-dimensional pathway can be distinguished from the 3D pathway by analysing the reaction on a DNA catenane containing two topologically interlinked rings of DNA with one target site on each ring. In a catenane, communications between the sites in the two separate rings cannot occur by one-dimensional tracking along the DNA from one site to the other. For example, two Type I restriction enzymes bound to separate rings of a catenane can never meet each other as they translocate around their individual rings, and thus fail to cleave the catenane [38]. On the other hand, the two sites in the interlinked rings of the catenane are held together almost as closely as they would be if the sites were present in cis on the same ring of DNA. Consequently, the Type IIF restriction enzyme SfiI cleaved a catenane with one SfiI site in each ring as readily as the two-site plasmid, at a much higher rate than the plasmid with one SfiI site [39].

When tested against catenanes with one copy of the requisite target sequence in each ring, BcgI [16], BaeI and CspCI [36] all acted like Type IIF enzymes and differed from Type I enzymes. All three Type IIB enzymes cleaved the catenane at similar rates to the two-site plasmid, i.e. much more rapidly than the one-site DNA. The Type IIB systems thus communicate between two DNA sites through 3D space, by bridging the sites and looping out the intervening DNA, and not by tracking along the DNA in the manner of a Type I enzyme.

**Gas guzzlers or solar power**

With the exception of NmeDI, the Type IIB systems are organized like the Type I systems in that both carry either separate peptides, or two or more separate domains within a single peptide, for their restriction, their modification and their DNA recognition activities, all within a single multifunctional protein. The main difference is that the Type IIB enzymes lack the DNA motor domain found in the HsdR subunit of the Type I enzymes, but a motor domain is not essential for R-M. The Type IIB systems fulfill the same biological function as the Type I systems, by using the endonuclease component of the protein assembly to restrict foreign DNA as it enters the cell while protecting the host DNA by using its MTase component to modify hemimethylated DNA [2,6].

Yet the Type IIB enzymes carry out these reactions far more efficiently in energetic terms than the Type I systems. The Type I REases consume hundreds or thousands of molecules of ATP in their helicase-like motor domains as they translocate along the DNA before acting at a random location often kilobases away from the recognition site [7], even though the reaction they actually catalyse, the hydrolysis of a DNA phosphodiester bond, has a massively negative ΔG°. In purely thermodynamic terms, there is no reason that DNA restriction should utilize any ATP. The Type I systems can thus be described as the ‘gas guzzlers’ of the restriction world in that they consume large amounts of energy for no apparent purpose. On the other hand, the Type IIB systems are, in effect, ‘solar-powered’ in that they do not require any
external supply of energy, but rely instead on the exchange of thermal energy with the environment to both find their target sites in the DNA (presumably by 3D diffusion [40]) and to carry out their negative ΔG° reactions.

Despite their inefficiency, the Type I systems may have survived on account of their ability to develop new sequence specificities. Although the evolution of a conventional Type II system to a new recognition sequence demands parallel changes in both the REase and the MTase proteins [2], the specificities of both REase and MTase functions in a Type I protein can evolve together by mutations in the HsdS subunit [6]. But in this respect, Type I has no advantages over Type IIB, as changes to the DNA-recognition domain in a IIB system can also affect in parallel both REase and MTase specificities [35].

The exception to the general rule, NmeDI [22], bears no resemblance to a Type I system, but may be akin to a Type IIS enzyme. The Type IIS REases recognize asymmetric sequences, to which they must bind in a fixed orientation set by the sequence, and then cleave both DNA strands at fixed locations upstream of the site [4,14,25–28]. But if there was a Type IIS REase that recognized a symmetrical sequence, then it would be able to bind to that site with its catalytic functions located upstream of the site, and also, with equal probability, to the same site in the opposite orientation, which would leave its catalytic functions downstream of the site. NmeDI has a symmetrical site (Table 2) so it may be just such an enzyme. Moreover, NmeDI is a tetrameric protein that needs to interact with two copies of its recognition site for full activity, but these properties are shared with several Type IIS endonucleases [41].

Acknowledgements

We thank Sumita Ganguly, Darren Gowers, Alistair Jacklin and Susan Rutter for their contributions to this work, and New England Biolabs for advice and materials.

Funding

The work in our laboratory was funded by the Biotechnology and Biological Sciences Research Council and by the Wellcome Trust.

References

14 Laganavicius, A., Sasaunaskas, G., Halford, S.E. and Siksnys, V. (2003) The metal-independent type IIS restriction enzyme Bli is a dimer that binds two DNA sites but has only one catalytic centre. J. Mol. Biol. 326, 1051–1064
22 Kwiatek, A. and Piekarowicz, A. (2007) The restriction endonuclease R.NmeDI from Neisseria meningitidis that recognizes a palindromic sequence and cuts the DNA on both sides of the recognition sequence. Nucleic Acids Res. 35, 6539–6546

©The Authors Journal compilation ©2010 Biochemical Society
38 Szczelkun, M.D., Dillingham, M.S., Janscak, P., Firman, K. and Halford, S.E. (1994) Repercussions of DNA tracking by the type IC restriction endonuclease EcoR124I on linear, circular and catenated substrates. EMBO J. 15, 6335–6347
41 Gormley, N.A., Hillberg, A.L. and Halford, S.E. (2002) The type II restriction endonuclease BspMI is a tetramer that acts concertedly at two copies of an asymmetric DNA sequence. J. Biol. Chem. 277, 4034–4041

Received 5 August 2009
doi:10.1042/BST0890410