Generation of DNA cleavage specificities of type II restriction endonucleases by reassortment of target recognition domains

Sonata Jurėnaitė-Urbanavičienė[†], Jurgita Šerkšnaitė[‡], Edita Kriukienė[†], Jolanta Giedrienė[†], Česlovas Venclovas[†], and Arvydas Lubys[‡§]

[†]Institute of Biotechnology, V. Graičiūno 8, LT-02241, Vilnius, Lithuania; and [‡]Fermentas UAB, V. Graičiūno 8, LT-02241, Vilnius, Lithuania

Edited by James A. Wells, University of California, San Francisco, CA, and approved May 8, 2007 (received for review November 23, 2006)

Type II restriction endonucleases (REases) cleave double-stranded DNA at specific sites within or close to their recognition sequences. Shortly after their discovery in 1970, REases have become one of the primary tools in molecular biology. However, the list of available specificities of type II REases is relatively short despite the extensive search for them in natural sources and multiple attempts to artificially change their specificity. In this study, we examined the possibility of generating cleavage specificities of REases by swapping putative target recognition domains (TRDs) between the type IIB enzymes AloI, PpiI, and TstI. Our results demonstrate that individual TRDs recognize distinct parts of the bipartite DNA targets of these enzymes and are interchangeable. Based on these properties, we engineered a functional type IIB REase having previously undescribed DNA specificity. Our study suggests that the TRD-swapping approach may be used as a general technique for the generation of type II enzymes with predetermined specificities.

hybrid | AloI | PpiI | TstI

Restriction endonucleases (REases) are parts of restriction-modification (R-M) systems, whose primary biological function is the protection of bacterial cells from incoming foreign DNA molecules (1). There are three main groups of restriction enzymes (types I, II, and III), which differ in enzyme composition, cofactor requirements, and mode of action (2). The best-studied are type II REases, which in general recognize specific DNA targets of 4–8 bp and cleave DNA at or close to these sequences (1, 2). The exquisite accuracy of type II enzymes (3, 4) has made them indispensable tools for DNA manipulations. Although almost 3,700 type II REases with 262 different specificities have been characterized to date (5), there still is a demand for enzymes recognizing new DNA targets.

During the past two decades, numerous efforts have been undertaken to engineer type II REases with altered specificities. Both rational protein design and random mutagenesis, followed by various selection procedures, have been tried (6), and several mutant enzymes with some preference for cleavage of altered DNA targets were isolated (7–10). However, projects concerned with orthodox type II REases so far have been largely unsuccessful mainly for two reasons: (*i*) difficulty of dealing with the observed tight coupling between DNA recognition and cleavage and (*ii*) absence of an efficient system for selecting enzymes with changed specificities. In this regard, unorthodox type II enzymes, such as the type IIG REase Eco57I (11), have shown more promise. Type IIG enzymes combine the catalytic centers of endonuclease and methyltransferase in one polypeptide chain, and the ability of Eco57I to methylate recognized DNA targets has been applied to isolate mutants having previously undescribed specificity (12).

The discovery of AloI-like REases (13–15), classified as type IIB enzymes, opened up new opportunities for the engineering of type II REases with altered specificities. AloI-like REases are large polypeptides having both DNA endonuclease and methyltransferase activities. This group of enzymes recognize bipartite DNA targets and cleave DNA on both sides of recognition sequences (13–15). It turned out that C-terminal regions of AloI-like REases share sequence similarity with specificity (HsdS) subunits of some type I R-M systems. Based on this observation, it was predicted that C-terminal regions of AloI-like REases, just like HsdS subunits, have two TRDs, each recognizing an individual part of the bipartite DNA target (13).

It was recognized >20 years ago that the recombinational reassortment of TRDs between HsdS subunits of type I proteins may result in REases with changed specificities (16–19), which, however, have no practical value, because type I REases cut DNA at random positions away from their recognition sequences (20). In this study, similarities between HsdS subunits and putative specificity regions (SRs) of type IIB enzymes encouraged us to test whether the reassortment of TRDs could also be used for rational engineering of type IIB REases enzymes with previously undescribed specificities.

Results

Characterization of REases Used for Domain-Swapping Experiments. The AloI (13), PpiI, and TstI REases used in this study are classified as type IIB REases because of their characteristic cleavage of DNA on both sides of the recognition sequence (2). Type IIB REases are bifunctional enzymes, which either cleave or methylate DNA, depending on the nature of cofactors added. DNA targets recognized by AloI, PpiI, TstI (Fig. 1*A*), and other type IIB enzymes are composed of two specific components separated by a nonspecific linker of varying length.

Genes for PpiI and TstI were cloned and expressed in *Escherichia coli* [see *Cloning of the Gene Encoding the PpiI Restriction-Modification System* and *Cloning of the Gene Encoding the TstI Restriction-Modification System* in supporting information (SI) *Text*], and the cloning and expression of the *aloI* gene were conducted as described (13). The primary structure analysis revealed very similar organization among AloI, PpiI, and TstI. As depicted in Fig. 1*A*, all three polypeptides can be subdivided into three structural-functional regions. A putative catalytic sequence motif AD. . . ECK is found near the N-terminal part of

Author contributions: S.J.-U. and Č.V. designed research; S.J.-U., J.Š., E.K., and J.G. performed research; S.J.-U. and Č.V. analyzed data; and S.J.-U., Č.V., and A.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: REase, restriction endonuclease; TRD, target recognition domain, R-M, restriction-modification; SR, specificity region.

Data deposition: The sequences reported in this paper have been deposited in the European Molecular Biology Laboratory database (accession nos. AM410094 and AM410095).

[§]To whom correspondence should be addressed. E-mail: lubys@fermentas.lt.

This article contains supporting information online at www.pnas.org/cgi/content/full/0610365104/DC1.

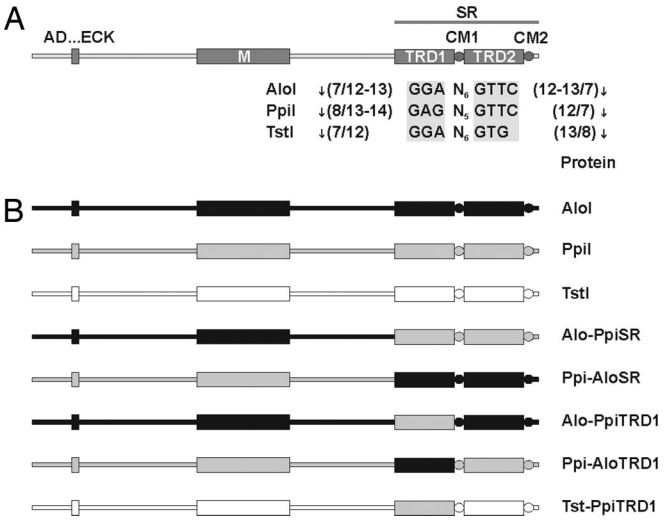

Fig. 1. Schematic representation of progenitor enzymes AloI, PpiI, TstI, and their hybrids. (*A*) The structure–function organization and DNA recognition sequences of AloI, PpiI, and TstI R-M systems. (*B*) Progenitor enzymes and hybrids constructed in this work. AD. . . ECK, catalytic motif of endonucleolytic activity; M, protein region encompassing DNA methyltransferase motifs; CM1 and CM2, conserved sequence motifs. Specific components of DNA targets predicted to be recognized by TRD1 and TRD2 are depicted in rectangles; cleavage positions are indicated by arrows.

each enzyme, indicating the potential involvement of this region in the DNA hydrolysis reaction (13). The central portions of the polypeptides are made up of conserved sequence motifs characteristic for amino methyltransferases (21) and presumably participate in AdoMet binding and the DNA methylation reaction. Similarly to AloI (13), the C-terminal SRs of PpiI and TstI harbor two putative target recognition domains TRD1 and TRD2 (see *Homology Searches of PpiI and TstI Target Recognition Domains* in *SI Text*) and two internally repeated conserved motifs CM1 and CM2 (Fig. 1*A*). Comparative analysis of SRs revealed intriguing similarity between the specificity determinants of type IIB and type I R-M enzymes. First, the individual TRDs of PpiI, TstI, and AloI (13) resemble the TRDs with identical or similar specificity found in HsdS subunits of type I enzymes (see *Homology Searches of PpiI and TstI Target Recognition Domains* in *SI Text*). Second, the conserved motifs of PpiI, TstI, and AloI (13) are similar to the conserved motifs found in HsdS subunits of type I enzymes (data not shown). Third, the mutual location of the TRDs and conserved motifs is the same in both enzyme types. Furthermore, DNA targets recognized by these enzymes share the same bipartite structure (13, 20). Taken together, these observations suggest that the SRs of type IIB enzymes share the same fold with HsdS, in which two TRDs form globular structures separated by a pair of two antiparallel α-helices, corresponding to conserved regions (22, 23).

Swapping of SRs Between AloI and PpiI. In initial experiments, we have swapped the entire C-terminal regions between AloI and PpiI (Fig. 1*B*). The goal of these experiments was 2-fold: (*i*) to test whether the presumed SRs are indeed involved in DNA target recognition and (*ii*) to find out whether these regions function as fairly independent structural modules and could be interchanged. The chosen swapping point for hybrid construction coincided with the appearance of reliable amino acid sequence similarities between the proximal (first) TRDs of AloI and PpiI and their equivalents of the same specificity in type I proteins. Thus, the SR of AloI (906–1,262 aa) was replaced with the corresponding segment of PpiI (906–1,289 aa) generating an Alo-PpiSR hybrid. The reciprocal swap resulted in Ppi-AloSR*, which was expressed in *E. coli* and was found to be soluble but had no REase activity in crude cell extracts (data not shown). Therefore, we decided to move the swapping point slightly upstream of the first putative TRD into the preceding conserved region. A Ppi-AloSR chimera, made by replacing the extended region of PpiI (856–1,289 aa) with the corresponding region of AloI (856–1,262 aa), was fully active (Table 1).

Table 1. Specific activities of AloI, PpiI, TstI, and hybrid REases

Enzyme	Recognition sequence*	Specific activity†
AloI	↓(7/12–13)GGAN$_6$GTTC(12–13/7)↓	4,700
PpiI‡	↓(8/13–14)GAGN$_5$GTTC(12/7)↓	18,600
TstI	↓(7/12)GGAN$_6$GTG(13/8)↓	18,900
Alo-PpiSR	↓(8/13–14)GAGN$_5$GTTC(12–13/7)↓	4,600
Alo-PpiTRD1	↓(8/13–14)GAGN$_5$GTTC(12–13/7)↓	4,600
Ppi-AloSR	↓(7/12–13)GGAN$_6$GTTC(12/7)↓	9,500
Ppi-AloTRD1	↓(7/12–13)GGAN$_6$GTTC(12/7)↓	18,900
Tst-PpiTRD1-Gly1006Leu§	↓(8/14–15)GAGN$_5$GTG(13–14/8)↓	<1,500

*Purification and cleavage position determination of hybrid proteins is described in *Methods* in *SI Text*.
†Specific activity is given as units per milligram of enzyme. One unit is defined as the amount of enzyme required to cleave 1 μg of BamHI-linearized DNA of pSEAd-7 in 1 h at optimal temperature in a reaction volume of 50 μl.
‡Cleavage positions may vary depending on the DNA environment.
§Detailed description of cleavage position determination is provided in *Cleavage Positions of the New Specificity REase* in *SI Text* and SI Fig. 7.

Because type IIB enzymes are able to methylate and cleave target DNA, both functions of the hybrid proteins were tested. To evaluate whether hybrids modify AloI and/or PpiI targets *in vivo*, recombinant plasmids pAlo-PpiSR and pPpi-AloSR coding for Alo-PpiSR and Ppi-AloSR, respectively, were digested with either AloI or PpiI. As shown in Fig. 2*A*, DNA of pAlo-PpiSR is linearized by AloI (lane 2) but is not cleaved by PpiI (lane 4), although there are four PpiI targets on it. In contrast, pPpi-AloSR is cleaved by PpiI (two DNA fragments of 8.42 and 0.84 kb are observed as expected for PpiI cleavage) but resistant to AloI despite the presence of two AloI targets (lanes 14 and 12, respectively). Cleavage of supplementary λ phage DNA in control reactions (lanes 3, 5, 13, and 15) clearly demonstrated that both REases performed well. Taken together, these results indicate that Alo-PpiSR modifies DNA targets recognized by PpiI, whereas Ppi-AloSR methylates AloI recognition sequences. To test the endonucleolytic activity of hybrid proteins as well as their specificity, the BamHI-linearized pSEAd-7 DNA was incubated with purified chimeric enzymes. Data presented in Fig. 3*A* demonstrate that specificities of Ppi-AloSR (lane 3) and Alo-PpiSR (lane 6) are identical to those of AloI (lane 2; as expected for the AloI cleavage, three DNA fragments of 4.94, 2.14, and 1.32 kb are observed) and PpiI (lane 5; as expected for PpiI digestion, five DNA fragments of 4.59, 1.85, 0.86, 0.76, and 0.35 kb are observed), respectively. Furthermore, Alo-PpiSR and Ppi-AloSR cleave DNA on both sides of the recognition sequences, and their specific DNA cleavage activities are comparable to those of their progenitors AloI and PpiI (Table 1). To summarize, the functional activity of hybrids demonstrate the exchangeability of SRs, and their involvement in DNA target recognition. Next, we investigated whether swapping of individual TRDs between the related type IIB enzymes AloI and PpiI could produce functional REases.

TRD1 Swapping Between AloI and PpiI. Proximal TRDs of AloI (13) and PpiI were assumed to recognize trinucleotide components of their bipartite DNA targets, GGA and GAG, respectively (see

BIOCHEMISTRY

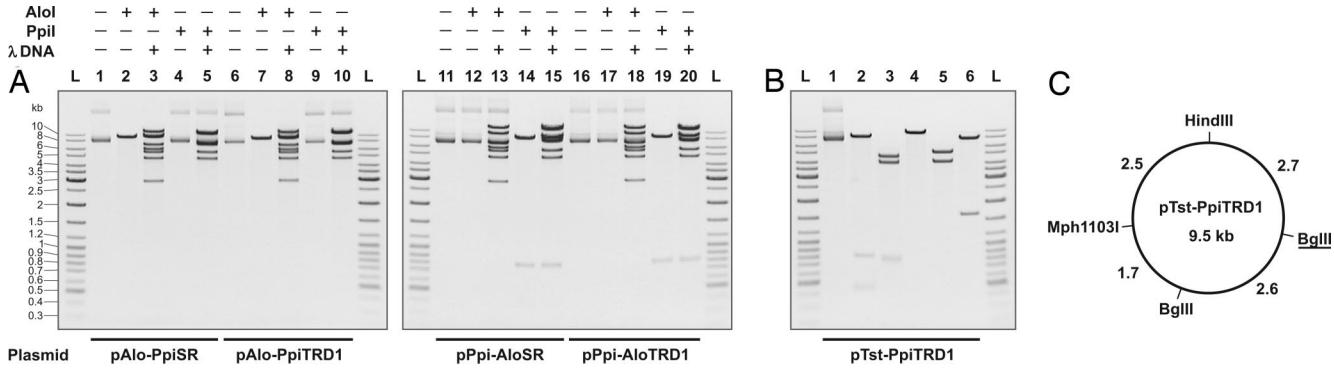

Fig. 2. The *in vivo* methylation specificity of hybrid proteins. (*A*) The *in vivo* methylation specificity of hybrids between AloI and PpiI. Recombinant plasmids were digested with AloI or PpiI. (*B*) The *in vivo* methylation specificity of Tst-PpiTRD1. DNA of plasmid pTst-PpiTRD1 was digested with: lane 1, no enzyme; lane 2, PpiI; lane 3, TstI; lane 4, BglII; lane 5, BglII + HindIII; lane 6, BglII + Mph1103I. (*C*) Restriction map of the pTst-PpiTRD1 plasmid. Modified BglII target is underlined. L, DNA ladder.

Homology Searches of PpiI and TstI Target Recognition Domains in *SI Text*). To investigate whether these TRDs maintain their function as part of another polypeptide, we designed two chimeric proteins. In the Alo-PpiTRD1 chimera encoded by the plasmid pAlo-PpiTRD1, the TRD1 region of AloI (906–1,019 aa) was replaced with an equivalent segment of PpiI (906–1,047 aa). In the Ppi-AloTRD1 hybrid encoded by plasmid pPpi-AloTRD1, the extended TRD1 fragment of PpiI (856–1,047 aa) was substituted for an equivalent fragment of AloI (856–1,019 aa) (Fig. 1*B*).

The ability of hybrid proteins to modify AloI and/or PpiI targets *in vivo* was evaluated by digesting DNA of purified pAlo-PpiTRD1 and pPpi-AloTRD1 plasmids with either AloI or PpiI. Data presented in Fig. 2*A* demonstrate that pAlo-PpiTRD1 is linearized by AloI (lane 7) but resistant to PpiI (lane 9) despite its four PpiI recognition sequences. On the contrary, the pPpi-AloTRD1 plasmid DNA is not cleaved by AloI (lane 17) despite the presence of three AloI targets but is cleaved by PpiI (lane 19; DNA fragments of 8.73 and 0.84 kb are produced as expected for the PpiI cleavage). Thus, Alo-PpiTRD1 and Ppi-AloTRD1 are functionally active *in vivo* and methylate PpiI and AloI DNA targets, respectively. To test whether hybrid enzymes maintained their endonucleolytic function, purified proteins were incubated with BamHI-linearized pSEAd-7 DNA. Digestion patterns of pSEAd-7 demonstrate that Alo-PpiTRD1 cleaves PpiI recognition sequences (Fig. 3*A*, compare lane 7 with lane 5), whereas Ppi-AloTRD1 possesses the specificity of AloI (Fig. 3*A*, compare lane 4 with lane 2). As shown in Table 1, both chimeras cleave DNA on both sides of their recognition sequences and demonstrate specific DNA cleavage activities that are similar to those of their progenitors.

Data presented above demonstrate that TRDs of analyzed type IIB enzymes are independent exchangeable structures. Furthermore, specificities of the analyzed hybrids clearly show that regions identified as the proximal TRDs of AloI and PpiI not only are involved in the recognition of trinucleotide segments of DNA targets, but they also determine the length of a nonspecific linker of bipartite DNA sequences. Because of this peculiarity and the fact that tetranucleotide parts of DNA targets of AloI and PpiI are identical (GTTC), the hybrid enzymes Alo-PpiTRD1 and Ppi-AloTRD1 acquired specificities of their progenitors.

TRD1 Swapping Between PpiI and TstI. Results of TRD swapping described above have disclosed principles of bipartite DNA target recognition by type IIB enzymes. Accordingly, we reasoned that the replacement of TRD1 in TstI with that from PpiI (Fig. 1*B*) should yield a hybrid enzyme of mixed specificity $GAGN_5GTG$, which has not yet been found in nature. However, it should be pointed out that the SRs of PpiI and TstI share only ≈19% identical residues, making the task of selecting optimal swapping points far from trivial. We considered that the sequence similarity-based approach used in designing chimeras made of AloI and PpiI (≈44% identity between their SRs) in this case would be insufficient. Therefore, to design the Tst-PpiTRD1 hybrid, we first generated models for the SRs of PpiI and TstI using recently determined crystal structure of the specificity subunit of the type I R-M system from *Methanococcus*

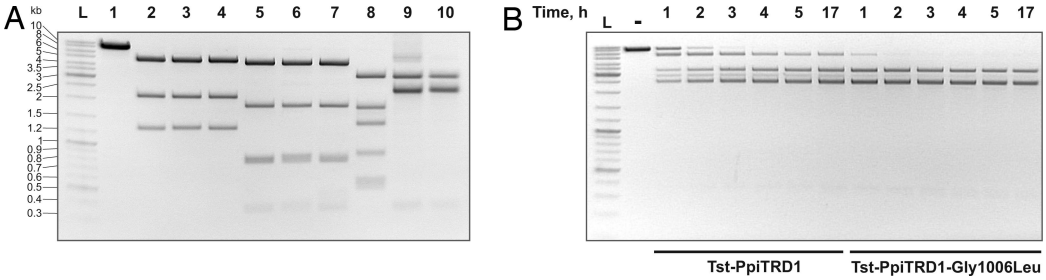

Fig. 3. Digestion of BamHI-linearized DNA of pSEAd-7 with progenitor and hybrid REases. (*A*) Specificity of progenitor and hybrid enzymes. Lane L, DNA ladder. DNA (3.7 nM) was incubated with the indicated enzymes: lane 1, no enzyme; lane 2, AloI (29.6 nM); lane 3, Ppi-AloSR (14.8 nM); lane 4, Ppi-AloTRD1 (7.4 nM); lane 5, PpiI (7.2 nM); lane 6, Alo-PpiSR (29.6 nM); lane 7, Alo-PpiTRD1 (29.6 nM); lane 8, TstI (7.4 nM); lane 9, Tst-PpiTRD1 (111 nM); lane 10, Tst-PpiTRD1-Gly1006Leu (29.6 nM). Reactions with AloI, Ppi-AloSR, Ppi-AloTRD1, PpiI, Alo-PpiSR, and Alo-PpiTRD1 continued for 1 h at 30°C, DNA digestion with TstI was performed for 1 h at 37°C, DNA was incubated with Tst-PpiTRD1 overnight at 37°C, DNA incubation with Tst-PpiTRD1-Gly1006Leu was carried out for 3 h at 37°C. (*B*) DNA digestion by Tst-PpiTRD1 and its mutant Tst-PpiTRD1-Gly1006Leu for various times at 37°C. DNA (3.7 nM) was incubated with the indicated enzymes at 29.6 nM.

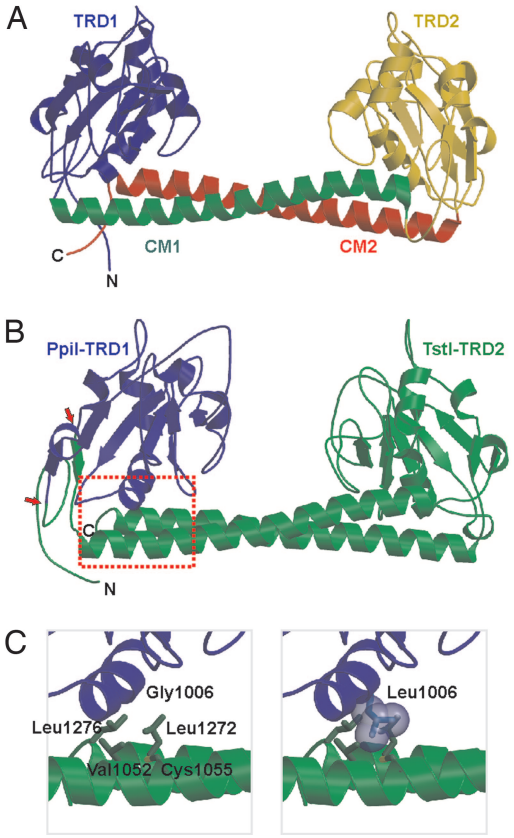

Fig. 4. 3D structures of the HsdS subunit of the type I R-M system from *M. jannaschii* (Protein Data Bank ID code 1YF2) and the computationally derived model of the SR of Tst-PpiTRD1. (*A*) The domain architecture of the HsdS subunit of the type I R-M system from *M. jannaschii*. Individual structural domains/motifs are colored from blue to red according to the progression of the polypeptide chain from the N to the C terminus. TRD1, TRD2, TRDs; CM1, CM2, conserved sequence motifs. (*B*) A model of the SR of Tst-PpiTRD1. TRD1 of PpiI is colored blue, the TstI moiety is shown in green. Red arrows indicate swapping points. A segment of the SR of TstI-PpiTRD1 enlarged in *C* is marked by a dotted square. (*C*) A closeup view of the interface between TRD1 of Ppi and the conserved helices of TstI. The initial Tst-PpiTRD1 construct is on the left, and the Gly1006Leu mutant is on the right. Several side chains at the interface are shown as sticks with space-filling contour added for Leu-1006.

jannaschii (23) as a template. Corresponding alignments are presented in SI Fig. 5. As shown in Fig. 4*A*, the specificity subunit is comprised of two globular TRDs separated by two antiparallel α-helices encompassing residues of conserved motifs. For TRD1 swapping between TstI and PpiI, we selected residue positions, which would preserve native contacts within the TRD1 structure of PpiI as much as possible and would also minimize the number of novel contacts between this domain and the TstI part of the hybrid. Thus, the Tst-PpiTRD1 hybrid encoded by pTst-PpiTRD1 was constructed by replacing the 894- to 1,010-aa segment of TstI with the 904- to 1,048-aa fragment of PpiI (Fig. 4*B* and SI Fig. 5). As shown in Fig. 2*B*, the hybrid enzyme does not methylate *in vivo* DNA targets recognized by its progenitor enzymes, because pTst-PpiTRD1 DNA is cleaved with both PpiI (lane 2; PpiI generates three DNA fragments of expected size, 8.24, 0.84, and 0.45 kb) and TstI (lane 3; as expected for the complete TstI cleavage, three DNA fragments of 4.68, 4.08, and 0.77 kb are observed). To test whether the hybrid enzyme modifies presumable targets GAGN$_5$GTG *in vivo*, DNA of pTst-PpiTRD1 was cleaved with BglII. There are two BglII recognition sites AGATCT on this plasmid, and one of them overlaps with GAGN$_5$GTG at the underlined positions (Fig. 2*C*). It is known that BglII cleavage is impaired if the external adenine of its recognition sequence is methylated (5). Considering that PpiI modifies adenine in the upper DNA strand of its target GAGN$_5$GTTC (data not shown), we anticipated that the hybrid enzyme should methylate the same A of the hybrid site GAGN$_5$GTG. This modification should interfere with the cleavage of the overlapping BglII target. As shown in Fig. 2*B*, BglII cuts pTst-PpiTRD1 only at one position (lane 4). Mapping of the cleavage site by double digestions with BglII-HindIII (lane 5; cleavage results in DNA fragments of 5.3 and 4.2 kb) and BglII-Mph1103I (lane 6; cleavage results in DNA fragments of 7.8 and 1.7 kb) revealed that the plasmid is cut by BglII at the site which does not overlap with the predicted DNA target for Tst-PpiTRD1, showing that the hybrid enzyme indeed modifies GAGN$_5$GTG DNA sequences *in vivo*. The DNA cleavage specificity of the purified hybrid enzyme was evaluated by incubating Tst-PpiTRD1 with BamHI-linearized DNA of pSEAd-7. Although complete cleavage of the substrate was not achieved even after overnight incubation with 30-fold excess of Tst-PpiTRD1 over DNA, the pattern of generated DNA fragments (Fig. 3*A*, lane 9) was identical to that calculated for a REase of GAGN$_5$GTG specificity (3.2, 2.46, 2.39, and 0.34 kb). Based on these observations, we concluded that the hybrid enzyme recognizes the hybrid DNA target GAGN$_5$GTG. On the other hand, the very slow DNA cleavage rate suggested that the 3D structure of Tst-PpiTRD1 might have flaws and therefore needs optimization. Indeed, upon inspection of the Tst-PpiTRD1 model, we have detected a cavity at the interface between TRD1 and conserved helices. This cavity was absent in models of progenitor enzymes, where it was filled with the combination of either a small-large or a large-small pair of residues, the first one coming from TRD1 and the second from one of the conserved α-helices. In contrast, the Tst-PpiTRD1 hybrid had a small-small residue pair (Gly-1006 and Cys-1055) at the corresponding positions (Fig. 4*C*). Based on this finding, we attempted to improve the structure of Tst-PpiTRD1 by mutating Gly-1006 to a bulky residue, which could fill the cavity with its large hydrophobic side chain, thereby increasing the stability of the protein. Thus, Gly-1006 was replaced with Leu as both visual analysis and computational energy estimation (24) suggested this to be the optimal choice for substitution (Fig. 4*C*). A stereoview of interfaces between proximal TRDs and the conserved helices in models of Tst-PpiTRD1-Gly1006Leu chimera, PpiI and TstI are presented in SI Fig. 6. Data in Fig. 3*B* demonstrate the significant activity enhancement resulting from Gly1006Leu substitution: with the enzyme/substrate ratio = 8 for each enzyme, DNA digestion with Tst-PpiTRD1-Gly1006Leu is complete after 3 h of incubation whereas cleavage with Tst-PpiTRD1 is partial even after overnight incubation (see also Fig. 3*A*, lanes 10 and 9, respectively). Although the specific activity of the Tst-PpiTRD1-Gly1006Leu mutant is still reduced compared with that of the parental enzyme TstI (Table 1), the mutant is sufficiently active to be useful in practical applications. We believe that additional minor adjustments of the Tst-PpiTRD1-Gly1006Leu tertiary structure could further increase its specific activity. However, here our general goal was to establish proof of principle rather than to produce a superactive enzyme. Thus, our results have demonstrated that the modular structure of type IIB REases SRs can be successfully exploited to generate functionally active artificial enzymes with predetermined specificities. We also have shown that minimal structural modification might be sufficient to dramatically improve the performance of such chimeric REases.

Discussion

The exchange of functional domains between existing proteins is one way to obtain hybrid enzymes with desired activities and properties (25). To date, the construction of hybrid endonucle-

ases using this approach has been reported for several groups of these proteins. It was demonstrated that the specificity of type I R-M enzymes may be changed by recombinational reassortment of TRDs between HsdS subunits of enzymes attributed to the same type I enzyme family (16–19). However, because of their complexity and property of cleaving DNA at undefined positions, these enzymes are not used as tools for molecular biology. The modular structure of type IIS REases with separate domains for DNA target recognition and cleavage (26) also made it possible to generate hybrid restriction enzymes. The DNA cleavage domain of the type IIS REase FokI was fused with DNA binding proteins such as the *Drosophila Ubx* homeodomain (27), the yeast Gal4 protein (28) and zinc-finger proteins (29) to engineer chimeric endonucleases. Furthermore, it was demonstrated that domains from unrelated homing endonucleases or mutated subdomains of an individual homing enzyme can be fused to create active chimeric enzymes of altered specificity (30–33). However, the hybrids mentioned above are extremely rarely cutting enzymes, which have specific applications including gene therapy (34, 35), but their potential to be used as analytical tools like conventional type II REases is limited.

In this study, we have demonstrated that the domain-swapping technique can also be used to generate type II REases with previously undescribed specificities. The technique works with type IIB REases, which recognize bipartite DNA sequences and combine DNA cleavage and methylation activities in a single polypeptide. By engineering active hybrids with swapped proximal TRDs among AloI, PpiI, and TstI, we have demonstrated a straightforward approach to generate type II REases of predictable specificity. In addition, we have shown that computational protein structure modeling and evaluation methods are effective means for both designing and improving functional properties of hybrid type IIB enzymes. The significant increase in Tst-PpiTRD1 activity resulting from a single amino acid substitution suggested by the computational analysis indicates both the accuracy of the generated protein models and the prospects for further improvements. It is worth mentioning that the domain-swapping technique is not limited only to TRD1 swaps, because TRD2 exchange between PpiI and TstI enzymes can also generate REase of changed specificity (S.J.-U., unpublished data).

Currently, there are 18 type IIB REases in REBASE (5), providing 21 unique TRDs. Theoretically, all these TRDs could be used in domain-swapping experiments. The sum of their combinations (>400), where each combination represents different DNA recognition specificity, exceeds the number of specificities (262) of type II enzymes known today. Obviously, not all of the TRD combinations might produce functionally active enzymes. Nonetheless, based on the results presented in this work, we believe that the approach of combinatorial reassortment of TRDs of type IIB REases has the potential to greatly expand the list of available REase specificities. In addition, the potential pool of TRDs for constructing type IIB REases perhaps could be enlarged by borrowing some of these domains from HsdS subunits of type I R-M systems.

Methods

Bacterial Strains and Plasmids. *E. coli* strain ER2566 (New England Biolabs, Ipswich, MA) was used for the expression of recombinant proteins. *E. coli* cells were grown in LB medium containing ampicillin (50 μg/ml). *E. coli* transformations were carried out by using the $CaCl_2$-heat-shock method (36). AloI, PpiI, TstI, and hybrid proteins were overproduced by using the expression vector pET21b(+) (Novagen, Madison, WI). BamHI-linearized DNA of pSEAd-7 was provided by Fermentas UAB and used for the determination of DNA cleavage specificity and the specific activity of hybrids.

Hybrid Construction. Hybrid genes were constructed on the backbone of plasmids pET-Alosup9 (provided by Fermentas UAB), pET-PpiI, and pET-TstIx (see *Cloning of the Gene Encoding the PpiI Restriction-Modification System* and *Cloning of the Gene Encoding the TstI Restriction-Modification System* in *SI Text*). All three plasmids are derived from pET21b(+) and harbor genes for AloI, PpiI, and TstI REases, respectively, under the control of the T7 promoter.

The pPpi-AloSR plasmid coding for the Ppi-AloSR hybrid was constructed by replacing the Bsp119I-Bsp120I fragment of pET-PpiI with the Bsp119I-Bsp120I fragment from pET-Alosup9. Hybrids described below were constructed by applying primers used in an overlap extension technique (37). The 5′-end sequences of these primers are complementary to sequences of other genes allowing joining DNA sequences from different genes at a particular position. To construct pPpi-AloSR* coding for the Ppi-AloSR* chimera, PCR was performed by using pET-Alosup9 as a template and the primers 5′-gtccgtcacgaagtggccgcaagtgaaggttggaagtatttgtagctt-3′/5′-cgagtgcggccgcaagcttgg-3′. The PCR product was cleaved with AdeI and NotI and then ligated to the large AdeI-NotI fragment from pET-PpiI. To obtain pAlo-PpiSR coding for the Alo-PpiSR hybrid REase, the PCR fragment was produced by using the primers 5′-tcatatcgccagcaaatggccacagatgccaatccgacaagttgcggtg-3′/5′-cgagtgcggccgcaagcttgg-3′ and pET-PpiI as a template. The resulting purified PCR fragment (megaprimer) and a new primer 5′-gaggcgaagcggaaggtagc-3′ were used for the second PCR with pET-Alosup9 as a template. A resulting PCR fragment was BcuI-NotI cleaved and ligated to the large BcuI-NotI fragment from pET-Alosup9. To construct pAlo-PpiTRD1 coding for Alo-PpiTRD1, the TRD1 of PpiI was PCR-amplified by using a pair of primers 5′-tcatatcgccagcaaatggccacagatgccaatccgacaagttgcggtg-3′ and 5′-tattggcggcagaggcatcggaaggtcggcatagaacgattcagg-3′ and pET-PpiI as a template. The purified PCR product and a new primer 5′-cgagtgcggccgcaagcttgg-3′ were used for the second PCR round with pET-Alosup9 as a template. The third PCR round was performed by using pET-Alosup9 as a template, the purified fragment obtained after the second round of PCR and the primer 5′-gaggcgaagcggaaggtagc-3′. The resulting PCR product was cleaved with BcuI-Eco105I and ligated to the large BcuI-Eco105I fragment of pET-Alosup9. To construct pPpi-AloTRD1 coding for the Ppi-AloTRD1 hybrid, the PCR product obtained by using primers 5′-gaggcgaagcggaaggtagc-3′/5′-ggggcggtaccgggattctcagt-tcgtgcgcttcgttacgattaag-3′ and pET-Alosup9 as a template was Bsp119I-Acc65I cleaved and ligated to the large Bsp119I-Acc65I fragment of pET-PpiI. Construction of pTst-PpiTRD1 coding for the Tst-PpiTRD1 chimera was done by producing a PCR fragment by using pET-PpiI as a template and the primers 5′-caacgcaggtg-cagagcaaatggccgcaaatgccaatccgacaag-3′/5′-cagcggtgggacggggatttt-tcaggtcggcatagaacgattcag-3′. The resulting purified PCR fragment and the primer 5′-gctagttattgctcagcggtg-3′ were used for the second round of PCR with pET-TstIx as a template. Then, a purified PCR fragment and the primer 5′-tgaccgagatggcgtctcag-3′ were used for the third PCR round with the same plasmid as a template. The resulting PCR product was Mph1103I-HindIII cleaved and ligated to the large Mph1103I-HindIII fragment of pET-TstIx. To construct pTst-PpiTRD1-Gly1006Leu coding for the mutant enzyme Tst-PpiTRD1-Gly1006Leu, the PCR product was obtained by using the primers 5′-cctccgccctattcaagtacaggaatacaaagtcaggaa-3′/5′-tgtcccgtgattgtggtgc-3′ and pTst-PpiTRD1 as a template. The second PCR round was performed by using the resulting purified PCR fragment, the primer 5′-gctagttattgctcagcggtg-3′ and pTst-PpiTRD1 as a template. The PCR product was BshTI-MlsI cleaved and ligated to the large BshTI-MlsI fragment of pTst-PpiTRD1.

Isolation of DNA and Recombinant DNA Techniques. Plasmids were prepared by the alkaline lysis procedure (38) and purified additionally as described (39) or isolated by using GeneJET Plasmid Miniprep Kit (Fermentas UAB). Standard techniques

(36) were used for recombinant plasmid construction. The QIAquick Gel Extraction Kit (Qiagen, Chatsworth, CA) was used for isolation of DNA fragments from agarose gels. DNA sequencing was carried out on an ABI PRISM 377 sequencer by using The BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). All enzymes used for DNA manipulations, their buffers, dNTPs, and GeneRuler DNA Ladder Mix were from Fermentas UAB.

Endonuclease and Methyltransferase Assays. REase activity was assayed in 50 μl of Fermentas UAB buffer R+ [10 mM Tris·HCl (pH 8.5 at 37°C)/10 mM $MgCl_2$/100 mM KCl/0.1 mg/ml BSA] containing 1 μg of BamHI-linearized pSEAd-7 DNA at 30°C or 37°C. Reaction products were resolved on a 1% agarose gel and stained with ethidium bromide.

For determination of DNA modification specificity *in vivo*, cells were grown at 37°C, and protein synthesis was induced by adding isopropyl β-D-thiogalactoside to a final concentration of 1 mM when the optical density of culture reached 0.6 at 600 nm. After additional growth of cells for 3 h at 30°C or 37°C, plasmid DNA was isolated and incubated with an excess of AloI, PpiI, TstI, or BglII. Reaction products were resolved on a 0.7% agarose gel and stained with ethidium bromide.

Modeling of the SRs of PpiI and TstI. Initially, the SRs of both PpiI and TstI were searched for related proteins of known 3D structure that could be used as modeling templates. Searches were conducted with PSI-BLAST against the National Center for Biotechnology Information nonredundant protein sequence database with the last iteration against Protein Data Bank (PDB) sequences (sequences of known structure). These searches for both sequences detected highly significant (E value $< 10^{-40}$) matches to the structure of the HsdS subunit of the type I R-M system from *M. jannaschii* (PDB ID code 1YF2) (23). Despite the high significance of the homologous relationship, matching sequences share only ≈16% sequence identity. At this level of sequence similarity, it is not uncommon to have misaligned regions (40). Therefore, reliable sequence-structure alignment regions were derived with the PSI-BLAST intermediate sequence search (PSI-BLAST-ISS) tool (41). In unreliable regions, possible alternative alignments were explored further by building and assessing corresponding 3D models. Evaluation of models representing underlying alternative alignments was done by visual inspection with an emphasis on significant structural flaws such as buried uncompensated charges or hydrogen donors/acceptors and severe steric clashes. In addition, the models were assessed with Prosa2003 (24) energy profiles and Z scores as well as Verify3D (42) profiles. The consensus of evaluation results was used to select the best alignment variant. Prosa2003 Z scores for the models and experimental structures are presented in SI Table 2. Coordinate files for the models of SRs of PpiI, TstI, and TstPpiTRD1-Gly1006Leu are available at www.ibt.lt/bioinformatics/models/ppi_tst.

Three-dimensional models were constructed automatically from given sequence-structure alignments with MODELLER (43). The amino acid side chains for the resulting models were positioned by using a backbone-dependent rotamer library implemented in SCWRL (44).

We thank Fermentas UAB for the gift of enzymes, kits, and BamHI-linearized DNA of pSEAd-7. This work was supported by Lithuanian State Science and Studies Foundation Grants T-549 and B-10. Č.V. was supported in part by the International Research Scholar grant from the Howard Hughes Medical Institute.

1. Wilson GG, Murray NE (1991) *Annu Rev Genet* 25:585–627.
2. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev S, Dryden DT, Dybvig K, *et al.* (2003) *Nucleic Acids Res* 31:1805–1812.
3. Alves J, Selent U, Wolfes H (1995) *Biochemistry* 34:11191–11197.
4. Thielking V, Alves J, Fliess A, Maass G, Pingoud A (1990) *Biochemistry* 29:4682–4691.
5. Roberts RJ, Vincze T, Posfai J, Macelis D (2007) *Nucleic Acids Res* 35:D269–D270.
6. Pingoud AM (2004) *Restriction Endonucleases* (Springer, Heidelberg).
7. Skowron PM, Swaminathan N, McMaster K, George D, Van Etten JL, Mead DA (1995) *Gene* 157:37–41.
8. Waugh DS, Sauer RT (1994) *J Biol Chem* 269:12298–12303.
9. Whitaker RD, Dorner LF, Schildkraut I (1999) *J Mol Biol* 285:1525–1536.
10. Samuelson JC, Xu SY (2002) *J Mol Biol* 319:673–683.
11. Janulaitis A, Vaisvila R, Timinskas A, Klimasauskas S, Butkus V (1992) *Nucleic Acids Res* 20:6051–6056.
12. Rimseliene R, Maneliene Z, Lubys A, Janulaitis A (2003) *J Mol Biol* 327:383–391.
13. Cesnaviciene E, Petrusyte M, Kazlauskiene R, Maneliene Z, Timinskas A, Lubys A, Janulaitis A (2001) *J Mol Biol* 314:205–216.
14. Piekarowicz A, Golaszewska M, Sunday AO, Siwinska M, Stein DC (1999) *J Mol Biol* 293:1055–1065.
15. Vitor JM, Morgan RD (1995) *Gene* 157:109–110.
16. Fuller-Pace FV, Bullas LR, Delius H, Murray NE (1984) *Proc Natl Acad Sci USA* 81:6095–6099.
17. Gann AA, Campbell AJ, Collins JF, Coulson AF, Murray NE (1987) *Mol Microbiol* 1:13–22.
18. Gubler M, Braguglia D, Meyer J, Piekarowicz A, Bickle TA (1992) *EMBO J* 11:233–240.
19. Nagaraja V, Shepherd JC, Bickle TA (1985) *Nature* 316:371–372.
20. Murray NE (2000) *Microbiol Mol Biol Rev* 64:412–434.
21. Malone T, Blumenthal RM, Cheng X (1995) *J Mol Biol* 253:618–632.
22. Calisto BM, Pich OQ, Pinol J, Fita I, Querol E, Carpena X (2005) *J Mol Biol* 351:749–762.
23. Kim JS, DeGiovanni A, Jancarik J, Adams PD, Yokota H, Kim R, Kim SH (2005) *Proc Natl Acad Sci USA* 102:3248–3253.
24. Sippl MJ (1993) *Proteins* 17:355–362.
25. Nixon AE, Ostermeier M, Benkovic SJ (1998) *Trends Biotechnol* 16:258–264.
26. Wah DA, Hirsch JA, Dorner LF, Schildkraut I, Aggarwal AK (1997) *Nature* 388:97–100.
27. Kim YG, Chandrasegaran S (1994) *Proc Natl Acad Sci USA* 91:883–887.
28. Kim YG, Smith J, Durgesha M, Chandrasegaran S (1998) *Biol Chem* 379:489–495.
29. Kim YG, Cha J, Chandrasegaran S (1996) *Proc Natl Acad Sci USA* 93:1156–1160.
30. Chevalier BS, Kortemme T, Chadsey MS, Baker D, Monnat RJ, Stoddard BL (2002) *Mol Cell* 10:895–905.
31. Epinat JC, Arnould S, Chames P, Rochaix P, Desfontaines D, Puzin C, Patin A, Zanghellini A, Paques F, Lacroix E (2003) *Nucleic Acids Res* 31:2952–2962.
32. Steuer S, Pingoud V, Pingoud A, Wende W (2004) *ChemBioChem* 5:206–213.
33. Smith J, Grizot S, Arnould S, Duclert A, Epinat JC, Chames P, Prieto J, Redondo P, Blanco FJ, Bravo J, *et al.* (2006) *Nucleic Acids Res* 34:e149.
34. Gouble A, Smith J, Bruneau S, Perez C, Guyot V, Cabaniols JP, Leduc S, Fiette L, Ave P, Micheau B, *et al.* (2006) *J Gene Med* 8:616–622.
35. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) *Nature* 435:646–651.
36. Sambrook J, Fritsch EF, Maniatis TM (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY), 2nd Ed.
37. McPherson MJ (1991) *Directed Mutagenesis: A Practical Approach* (Oxford Univ Press, New York).
38. Birnboim HC, Doly J (1979) *Nucleic Acids Res* 7:1513–1523.
39. Marko MA, Chipperfield R, Birnboim HC (1982) *Anal Biochem* 121:382–387.
40. Kryshtafovych A, Venclovas Č, Fidelis K, Moult J (2005) *Proteins* 61(Suppl 7):225–236.
41. Margelevičius M, Venclovas Č (2005) *BMC Bioinformatics* 6:185.
42. Luthy R, Bowie JU, Eisenberg D (1992) *Nature* 356:83–85.
43. Sali A, Blundell TL (1993) *J Mol Biol* 234:779–815.
44. Canutescu AA, Shelenkov AA, Dunbrack RL, Jr. (2003) *Protein Sci* 12:2001–2014.