Nucleic Acids Research, 1998, Vol. 26, No. 8 **1857–1862**

The protein splicing domain of the homing endonuclease PI-*Sce*I is responsible for specific DNA binding

Wolfgang Grindl, Wolfgang Wende, Vera Pingoud and Alfred Pingoud*

Institut für Biochemie, Justus-Liebig-Universität, Heinrich-Buff-Ring 58, D-35392 Gießen, Germany

Received February 9, 1998; Accepted February 23, 1998

ABSTRACT

The homing endonuclease PI-*Sce*I consists of a protein splicing domain (I) and an endonucleolytic domain (II). To characterize the two domains with respect to their contribution to DNA recognition we cloned, purified and characterized the isolated domains. Both domains have no detectable endonucleolytic activity. Domain I binds specifically to the PI-*Sce*I recognition sequence, whereas domain II displays only weak non-specific DNA binding. In the specific complex with domain I the DNA is bent to a similar extent as observed with the initial complex formed between PI-*Sce*I and DNA. Our results indicate that protein splicing domain I is also involved in recognition of the DNA substrate.

INTRODUCTION

The highly specific homing endonucleases are encoded by introns or inteins and have been found in all three phylogenetic kingdoms. *In vivo* their activity leads to a gene conversion process, called homing, resulting in site-directed integration of the homing endonuclease gene in an intronless or inteinless allele (reviewed in 1–3).

The homing endonuclease PI-*Sce*I is encoded by the VMA1 intein of *Saccharomyces cerevisiae* (4–7). Excision of the intein from the preprotein and ligation of the two remaining exteins is an autocatalytic protein splicing process (8,9). The resulting free endonuclease binds specifically to a 35–45 bp recognition sequence, distorts the DNA and cleaves the substrate in the presence of Mg^{2+} to produce a 4 bp 3′ overhang (10–12). Thus PI-*Sce*I combines two catalytic functions: a protein splicing and an endonucleolytic activity. This dual function is reflected in the bipartite structure of PI-*Sce*I (13), which consists of two separate domains with very different architectures.

Domain I (amino acids 1–182 and 411–454; Fig. 1) is presumably responsible for proteolytic excision of PI-*Sce*I from the preprotein. It contains the closely adjacent N- and C-termini of the homing endonuclease PI-*Sce*I and all amino acid residues identified so far to be essential for protein splicing (14,15). Furthermore, deletion experiments have shown that domain I is sufficient to excise the intein from a precursor (16,17) and mutations that abolish endonucleolytic activity do not affect protein splicing (18). The structure of domain I is elongated and dominated by seven β-sheets (13). This is in good agreement with a tertiary structure prediction of a common protein splicing domain based on hidden Markov models (19). The recently published crystal structure of the GyrA intein (20), which lacks endonucleolytic activity, is very similar to the structure determined for the protein splicing domain of PI-*Sce*I, suggesting that there is a common protein splicing domain architecture. In corroboration of that, the conserved amino acids essential for protein splicing and the essential features of the GyrA structure are also found in other proteins that undergo a self-processing reaction, as shown by sequence alignments (19,21,22) and crystal structure analysis of the *Drosophila* hedgehog protein autoprocessing domain (23).

Domain II, comprising amino acids 183–410, is presumably the endonucleolytic domain of PI-*Sce*I. It harbours two copies of the LAGLIDADG motif, which is characteristic for one family of homing endonucleases (3). Mutagenesis experiments have shown that the last aspartate residue in the LAGLIDADG motif of homing endonucleases is essential for endonucleolytic activity (18,24). Other amino acids involved in catalysis and/or substrate binding have been found proximal to the LAGLIDADG motif (25–28; V.Pingoud, unpublished results). Amino acid substitutions which abolish protein splicing in the preprotein do not interfere with the endonucleolytic activity of PI-*Sce*I (W.Grindl and W.Wende, unpublished results). For PI-*Sce*I it has been discussed that both LAGLIDADG motifs together form one catalytic centre (11,13,24) or, alternatively, that each LAGLIDADG motif is part of a separate catalytic centre (12). The structure of domain II consists of two α/β motifs each harbouring one LAGLIDADG motif related by a quasi-2-fold symmetry (13; Fig. 1). This architecture of domain II is very similar to that of the homodimeric I-*Cre*I homing endonuclease (29). I-*Cre*I is an intron-encoded homing endonuclease and lacks the protein splicing domain. It has only one LAGLIDADG motif per subunit. Recently published data based on a multiple alignment of 130 LAGLIDADG family members suggest that the main structural features as seen in the I-*Cre*I and PI-*Sce*I structures are present in all LAGLIDADG motif-containing domains (30).

The bipartite structure in PI-*Sce*I and other intein-encoded endonucleases make it likely that these homing endonucleases

*To whom correspondence should be addressed. Tel: +49 641 9935400; Fax: +49 641 9935409; Email: wolfgang.wende@chemie.bio.uni-giessen.de

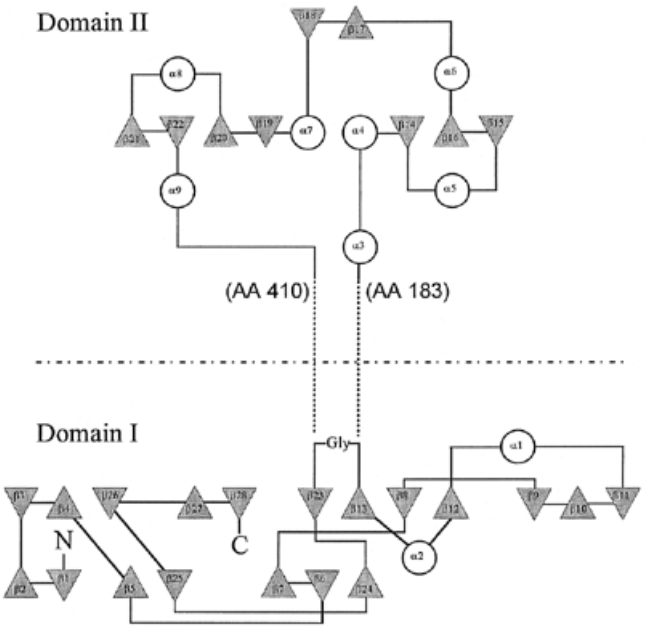

Figure 1. Topology of the individual PI-*Sce*I domains (adapted from Duan *et al.*; 13). PI-*Sce*I consists of two separate domains: protein splicing domain I and endonucleolytic domain II. The connection of both domains in PI-*Sce*I is shown by a dotted line. The additional glycine residue inserted into domain I is indicated. α-Helices are indicated in sequential order by α, β-sheets by β.

evolved by invasion of a functional endonuclease gene into a formerly autonomous protein splicing element (13,16,23).

Here we address the question whether the endonucleolytic activity of PI-*Sce*I is independent of the protein splicing domain and whether both domains are necessary for substrate binding, as indicated by the structure (13). To this end we have cloned, purified and characterized the two domains of PI-*Sce*I.

MATERIALS AND METHODS

Preparation of PI-*Sce*I domains I and II

The plasmids pHisPI-*Sce*I-DI and pHisPI-*Sce*I-DII coding for PI-*Sce*I domains I and II (13) were cloned by PCR. The domain II gene fragment was prepared using primers 5′-GCGCGCGGATC-CCTTTATGAGAATGACCACTTTTTC-3′ and 5′-GCGCGCGT-CGACGATATCAGGCGGGAGCAGGCCTGAATTTTTTAG-3′ in a PCR reaction with pHisPI-*Sce*I as template using *Pfu* DNA polymerase (Stratagene) essentially as described by Wende *et al.* (12). The resulting domain II gene fragment codes for amino acids 183–410 of wild-type PI-*Sce*I. To prepare the domain I gene fragment two PCR reactions were performed using pHisPI-*Sce*I as template. The first reaction was carried out with primers 5′-GCGGATCCGCATGCTTTGCCAAGGGTACCAATG-3′ and 5′-GCAAAAGCAGCACCAATTGGAGCGTAAGTCTGG-TAGG-3′, the second with primers 5′-CGCTCCAATTGGTGC-TGCTTTTGCACGTGAGTGCCGCGG-3′ and 5′-CCGGCGT-CGACGTCAGCAATTATGGACGACAACCTGG-3′. These primers introduce a *Mun*I restriction site at the 3′-end of the first PCR product and at the 5′-end of the second PCR product. Both PCR fragments were purified, digested with *Mun*I and ligated head-to-tail using T4 DNA ligase (MBI Fermentas). The resulting domain I gene fragment codes for amino acids 1–182 and 411–454 of wild-type PI-*Sce*I with an additional glycine connecting the protein fragments.

The purified gene fragments of domains I and II were cleaved with *Bam*HI and *Sal*I and inserted into plasmid pHisPI-*Sce*I which had previously been digested with the same restriction enzymes, to give pHisPI-*Sce*I-DI and pHisPI-*Sce*I-DII. The sequence of the genes coding for domains I and II were confirmed by sequencing. *Escherichia coli* cells were transformed with these plasmids and after fermentation the N-terminally His_6-tagged PI-*Sce*I domains were purified as described before for the full-length protein (12). The protein concentration was determined from absorbance at 280 nm. The extinction coefficients of the PI-*Sce*I domains were calculated according to Pace *et al.* (31).

DNA cleavage assay

Cleavage reactions were performed with 200 nM domain I or 500 nM domain II and 7 nM ^{32}P-labelled 311 bp DNA fragment with a central PI-*Sce*I cleavage site (12) or 8 nM supercoiled or linearized plasmid DNA (pBSVDEX; 10) as substrate. The reactions were carried out at 37°C in cleavage buffer (10 mM Tris–HCl, pH 8.5, 100 mM KCl, 1 mM DTT, 100 µg/ml BSA) containing 2.5 mM $MgCl_2$ or 2.5 mM $MnCl_2$. For competition experiments 50 nM PI-*Sce*I and different amounts of domain I (0, 50, 100 or 250 nM) were incubated in cleavage buffer with 2.5 mM $MgCl_2$ for 15 min. Aliquots of 7 nM ^{32}P-labelled 311 bp DNA fragment were added and the reaction mixture was incubated at 37°C. After defined time intervals aliquots were withdrawn and mixed with stop buffer [100 mM EDTA, pH 8.0, 25% Ficoll, 0.1% (w/v) bromophenol blue, 0.1% xylene cyanol]. The substrates and products were separated by electrophoresis in TPE buffer (80 mM Tris–phosphate, pH 8.0, 2 mM EDTA) either on 7% (w/v) polyacrylamide gels or 0.8% agarose gels, which were stained after electrophoresis with ethidium bromide or analysed using an Instant Imager (Canberra Packard).

Electrophoretic mobility shift assay

For electrophoretic mobility shift assays 7 nM ^{32}P-labelled 201 bp DNA fragment [pBend2(F)/*Sma*I, see below] were incubated with increasing amounts of domain I or domain II in shift buffer [10 mM Tris–HCl, pH 7.5, 1 mM EDTA, 50 mM NaCl, 0.05% (w/v) non-fat dry milk, 5% (v/v) glycerol, 10 mM DTT] for 1 h at room temperature. A sample of 10 µl of this binding mixture contains in addition 0.1 µg poly(dI·dC) (Pharmacia). After incubation 3 µl gel loading buffer (10% Ficoll, 15% glycerol, 50% binding buffer, 40 mM EDTA, 0.2% bromophenol blue, 0.1% xylene cyanol) were added and electrophoresis was performed on 6% (w/v) polyacrylamide gels. After electrophoresis gels were analysed using an Instant Imager.

Circular permutation assay

Substrates for the circular permutation assay were obtained from plasmid pBend2(F), which harbours the whole PI-*Sce*I recognition sequence (pBend2VMAΔvde in 12), and pBend2(FII), which contains the right half, downstream of the position of cleavage. pBend2(FII) was prepared by insertion of oligodeoxynucleotide FII (12) into the *Sal*I site of plasmid pBend2 (32) after the resulting sticky ends had been converted to blunt ends by the Klenow fragment of DNA polymerase I. Each plasmid was used

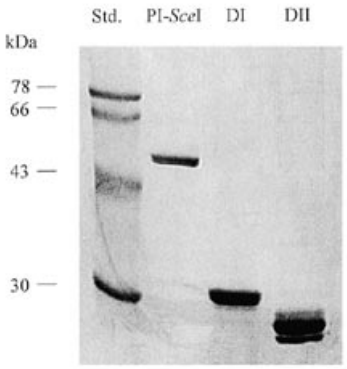

Figure 2. SDS–PAGE analysis of the purified PI-*Sce*I domains. The molecular weights of the products are: PI-*Sce*I, 52 kDa; domain I (DI), 26 kDa; domain II (DII), 27 kDa. Std, Merck IV protein standard.

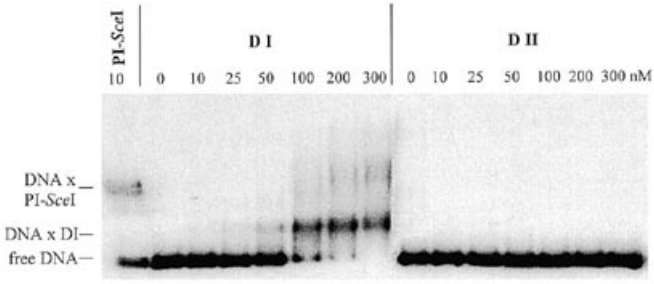

Figure 3. Electrophoretic mobility shift experiments with domain I or domain II and a 201 bp fragment derived from pBend2(F). Aliquots of 7 nM ^{32}P-labelled substrate were incubated with increasing amounts of domain I (DI), domain II (DII) or 10 nM PI-*Sce*I as a control. Complex formation was analysed on a 7% polyacrylamide gel. While domain II does not display specific binding, domain I does. At high concentration of domain I additional bands are seen, presumably due to non-specific binding of domain I to the DNA.

as template in a PCR with primers 5′-GAGGCCCTTTCGTCTT-CAAGAATTC-3′ and 5′-GTGATAAACTACCGCATTAAAG-CTT-3′ in the presence of [α-^{32}P]dATP to produce radioactively labelled substrates. The actual substrates used in bending experiments were generated by digestion of the labelled PCR products with the restriction enzyme *Mlu*I, *Sma*I or *Bam*HI, resulting in 201 bp fragments from pBend2(F) and 160 bp fragments from pBend2(FII). Electrophoretic mobility shift assays were performed on 20 × 20 cm 6% (w/v) polyacrylamide gels under the conditions as described above.

RESULTS

Cloning and expression of domains I and II

As recently described, PI-*Sce*I consists of two separate domains (Fig. 1). Domain I has been suggested to be involved in protein splicing, domain II to be responsible for the endonucleolytic activity and both domains to be required for DNA binding (13). To address the question whether both domains are involved in specific binding and cleavage of the recognition site we constructed two expression vectors coding for the separated domains I and II. Domain I (residues 1–182 and 411–454) contains amino acids identified to be involved in protein splicing of the VMA1 protozyme (14–17). In our construct of domain I we inserted an additional glycine residue to connect residues 182 and 411 and to prevent formation of a helix predicted by a secondary structure prediction program (33). Domain II (residues 183–410) harbours the two endonucleolytically important LAGLIDADG motifs (24). Domains I and II were expressed in *E.coli* and purified by metal affinity chromatography to near homogeneity as judged by SDS–PAGE (Fig. 2). Domain I has a molecular weight of 26 kDa but migrates with a mobility characteristic for a protein of 30 kDa. This anomalous retardation may reflect the elongated non-spherical structure of this domain. Analysis of the domain II preparation by SDS–PAGE showed two bands. The upper major band represents a protein with M_r ~ 27 kDa, the estimated size of domain II. The lower minor band is presumably a proteolytic fragment of domain II, with the intact N-terminal His_6-tag and five to six residues removed from the C-terminus. Correct folding of the individual PI-*Sce*I domains was verified by circular dichroism spectroscopy, which demonstrated that the isolated domains had the expected secondary structure composition (data not shown).

Cleavage activity of domains I and II

Domain II of PI-*Sce*I harbours two LAGLIDADG sequences which are considered to be part of the active sites of the homing endonucleases containing this motif (18,24,34). The isolated domain II, however, turned out to be inactive in cleaving supercoiled or linear plasmid DNA with the PI-*Sce*I recognition sequence, even in the presence of Mn^{2+}, which with native PI-*Sce*I aleads to relaxation of cleavage specificity, or at excess enzyme over substrate. As expected, domain I also did not show any cleavage activity under these conditions, neither did an equimolar mixture of domain I and II (data not shown). Thus the isolated domains I and II as well as their combination are devoid of specific and non-specific DNA cleavage activity.

Binding of domains I and II to DNA

Despite the lack of nucleolytic activity of domains I and II we characterized these proteins in terms of DNA binding activity. As a specific substrate a 201 bp fragment of pBend2(F) was used. Binding experiments were carried out in the presence of poly(dI·dC) to suppress non-specific binding. The results of electrophoretic mobility shift assays show that domain I binds to the substrate (Fig. 3). The apparent equilibrium constant K_d for formation of this complex in the absence of Mg^{2+} is ~140 nM. Under these conditions full-length PI-*Sce*I binds to the pBend2(F) fragment seven times more strongly. Binding is specific, as verified by competition with specific and non-specific DNA (data not shown). Domain II does not show any detectable specific binding (Fig. 3). In gel shift experiments without poly(dI·dC) as a non-specific competitor domain II showed weak non-specific binding (data not shown).

If domain I binds specifically to the PI-*Sce*I recognition sequence it should compete with full-length PI-*Sce*I for substrate binding. Indeed, domain I inhibits cleavage of the specific 311 bp substrate by PI-*Sce*I in a concentration-dependent manner (Fig. 4). At equimolar concentrations of domain I and PI-*Sce*I the activity of PI-*Sce*I was found to be reduced by 50%, suggesting that

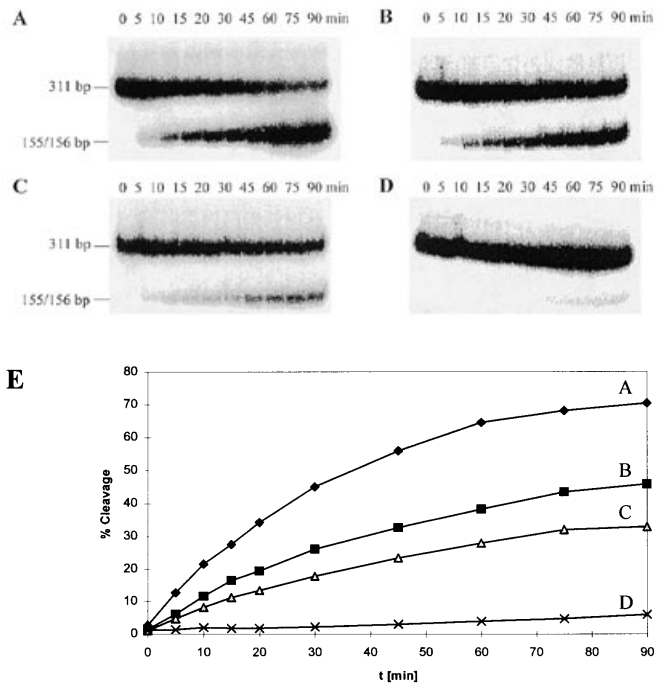

Figure 4. DNA cleavage experiments with PI-*Sce*I in competition with domain I. Aliquots of 7 nM specific 311 bp substrate derived from pBSVDEX were incubated with 50 nM PI-*Sce*I and various concentrations of domain I; (**A**) 0, (**B**) 50, (**C**) 100 or (**D**) 250 nM. The reaction products were analysed on a 6% polyacrylamide gel. (**E**) Reaction progress curves for the experiments shown in (A)–(D), from which initial rates were determined to obtain an estimate for the relative affinity of domain I and PI-*Sce*I to DNA.

domain I binds under cleavage conditions in the presence of Mg^{2+} with nearly the same affinity to DNA as full-length PI-*Sce*I.

Bending of DNA by domain I

As shown previously (11,12), PI-*Sce*I forms two complexes with specific DNA which differ in their electrophoretic mobilities. The 'lower' complex is characterized by a bend of 45° and the 'upper' complex by a bend of 75° (12). It has been suggested that the lower complex reflects initial binding to the right half of the recognition sequence downstream of the cleavage position (11,12).

We have now analysed, in gel shift experiments and circular permutation assays, whether domain I bends the DNA and whether two different complexes are formed upon specific binding to the PI-*Sce*I recognition sequence. As shown in Figure 5B, domain I binds to and bends DNA fragments derived from pBend2(F) which contain the whole recognition sequence, albeit at different positions. The degree of bending, however, is much less pronounced than with full-length PI-*Sce*I. Furthermore, in contrast to PI-*Sce*I, domain I only forms one complex, corresponding with respect to bend angle to the lower complex produced with full-length PI-*Sce*I. It was shown previously that PI-*Sce*I binds firmly to the right half of its recognition sequence (11,12); this binding is accompanied by slight bending of the DNA (11). Domain I also binds to DNA fragments derived from pBend2(FII) which contain the right half of the recognition sequence, albeit at different positions. This permutation analysis suggests that domain I bends this DNA substrate to a similar extent as PI-*Sce*I (Fig. 5C). There is no detectable binding of domain II, neither to pBend2(F) nor to the pBend2(FII) DNA fragment. Furthermore, an equimolar mixture of domain I and domain II displays only the binding and bending characteristics of domain I. There is no supershift, which one would expect if the isolated domains formed a stable complex (Fig. 5B and C).

DISCUSSION

We have cloned, expressed and purified the two domains of the homing endonuclease PI-*Sce*I. Neither the separate domains nor an equimolar mixture of both exhibit any nucleolytic activity. In the case of other nucleases, the class II restriction enzyme *Fok*I or the homing endonuclease I-*Tev*I, both of which are structurally organized in two domains, one responsible for specific DNA binding, the other for phosphodiester bond cleavage, the isolated catalytic domains display non-specific nucleolytic activity (35,36). The lack of activity of PI-*Sce*I domain II, which harbours the LAGLIDADG motifs, indicates that in the case of PI-*Sce*I the isolated catalytic domain is not sufficient to form a catalytically competent complex and that the activity is dependent on cooperation between domain I and domain II. The necessity for a precise juxtaposition and conformation of both domains is supported by the observation that PI-*Sce*I subjected to a denaturation/renaturation cycle binds specifically to the recognition site but displays no endonucleolytic activity (V.Pingoud, unpublished results). Domain II of PI-*Sce*I, which structurally resembles the homodimeric homing endonuclease I-*Cre*I (13,29), which is also a member of the LAGLIDADG family, only shows weak non-specific binding. In contrast, domain I of PI-*Sce*I, the protein splicing domain, displays specific DNA binding comparable in strength with that of PI-*Sce*I. Strong binding is not only observed for the recognition sequence but also for the right half of the recognition sequence, corresponding to one of the cleavage products of PI-*Sce*I. On the basis of a docking model for B-DNA bound to PI-*Sce*I it was suggested that domain I is involved in substrate binding by covering ~16 bp downstream of the cleavage site (13). Our experiments support this suggestion and indicate in addition that this binding is specific for the PI-*Sce*I recognition site. It must be emphasized that the absence of specific DNA binding by domain II does not mean that this part of PI-*Sce*I is not involved in DNA binding, because it might well be that subtle conformational differences between the isolated domain II and domain II in the context of intact PI-*Sce*I may preclude specific binding.

In previous studies it has been shown that PI-*Sce*I also binds to the downstream cleavage product (11,12). To investigate whether domain I distorts the DNA in this complex we performed electrophoretic mobility shift assays with two different sets of circularly permuted fragments, one containing the full-length recognition site, the other containing the downstream cleavage product. While PI-*Sce*I binding to a DNA containing the recognition sequence leads to two distinct complexes, as reported before (11,12), in which the DNA is bent by 75° (upper complex) and 45° (lower complex) respectively, binding of domain I gives rise to formation of only one complex, corresponding to the lower complex with PI-*Sce*I. Both PI-*Sce*I and domain I bind to DNA which contains the right half of the recognition site. With this DNA, however, only one complex is formed. The DNA distortion in this complex is comparable with the distortion in the lower complex (~45°; 11,12) occurring upon binding of PI-*Sce*I to the full-length recognition site. These findings support our previous

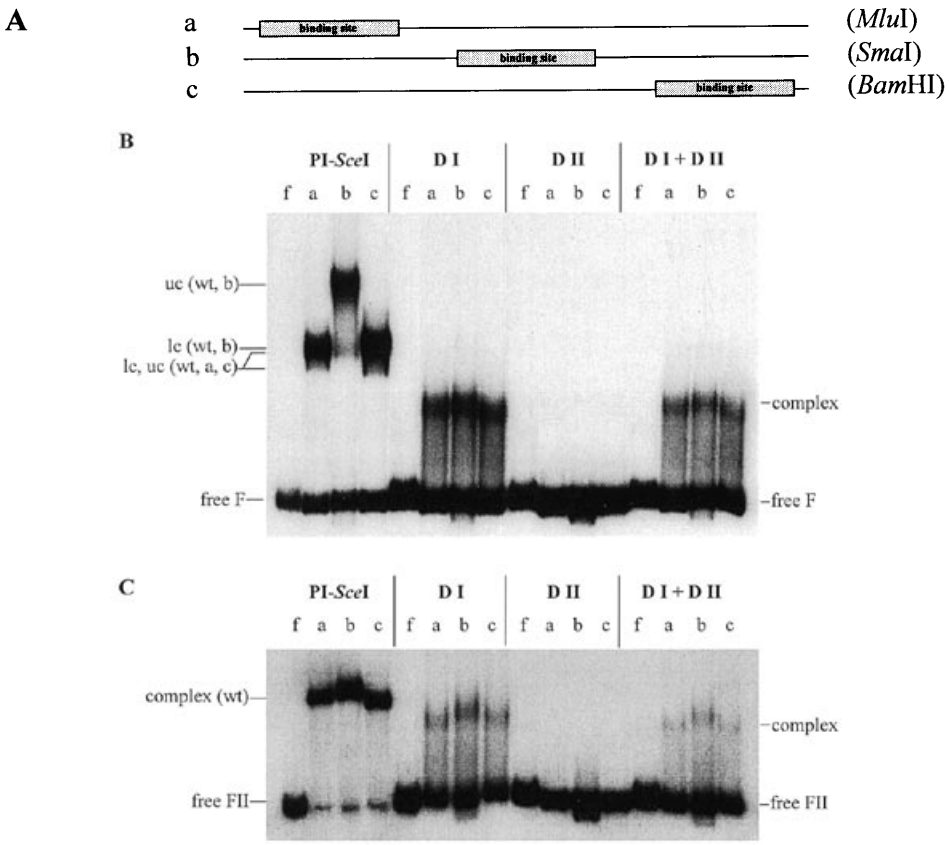

Figure 5. Bending experiments with PI-*Sce*I, domain I (DI), domain II (DII) and circularly permuted fragments. (**A**) Schematic diagram showing the DNA fragments used in this assay. pBend2(F) and pBend2(FII) were digested with (a) *Mlu*I, (b) *Sma*I and (c) *Bam*HI resulting in fragments with different positions of the binding site. (**B**) Electrophoretic mobility shift assay with PI-*Sce*I, domain I (DI), domain II (DII) and an equimolar mixture of both domains (DI + DII) with pBend2(F) fragments. The positions of the upper complex (uc), the lower complex (lc) for PI-*Sce*I, the complex with DI (complex) and free DNA (free F, lane f) are indicated. (**C**) Electrophoretic mobility shift assay with pBend2(FII) fragments. The positions of the various complexes and free DNA are indicated as above.

suggestion that the lower complex is formed by initial binding of PI-*Sce*I to the region downstream of the cleavage site. Based on the data presented here we conclude that this initial complex is dominated by interactions between domain I and the right half of the recognition sequence. Only when additional interactions between domain II and the left half of the recognition sequence are formed is the active centre activated and does the reaction proceed.

The domain in PI-*Sce*I containing the LAGLIDADG motifs, domain II, is not necessary for protein splicing and can be deleted (16,17). The analogous experiment carried out here shows that the protein splicing domain, domain I, cannot be removed without loss of endonucleolytic activity. Our results show that domain I is not only responsible for protein splicing but is also involved in specific binding of the recognition sequence, as recently suggested by Hall *et al.* (23). For the evolution of intein-embedded homing endonucleases this could mean that after fusion of the genes coding for these domains co-evolution may have led to generation of specific DNA contacts in domain I. Thereby new specificities could have evolved which led to expansion of this family of intein-encoded homing endonucleases.

ACKNOWLEDGEMENTS

We would like to thank Drs Marlene Belfort and Shmuel Pietrokovski for sending us manuscripts prior to publication, Ms Frauke Christ for discussions and Ms Anja Wahl for expert technical assistance. This work has been supported by the Deutsche Forschungsgemeinschaft (Pi 122/13-1), the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie and the Fonds der Chemischen Industrie.

REFERENCES

1 Lambowitz,A.M. and Belfort,M. (1993) *Annu. Rev. Biochem.*, **62**, 587–622.
2 Mueller,J.E., Bryk,M., Loizos,N. and Belfort,M. (1993) In Linn,S. M., Lloyd,R.S. and Roberts,R.J. (eds), *Nucleases*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 111–143.
3 Belfort, M. and Roberts, R.J. (1997) *Nucleic Acids Res.*, **25**, 3379–3388.
4 Hirata,R., Ohsumi,Y., Nakano,A., Kawasaki,H., Suzuki,K. and Anraku,Y. (1990) *J. Biol. Chem.*, **265**, 6726–6733.
5 Kane,P.M., Yamashiro,C.T., Wolczyk,D.F., Neff,N., Goebl,M. and Stevens,T.H. (1990) *Science*, **250**, 651–657.

6 Gimble,F.S. and Thorner,J. (1992) *Nature*, **357**, 301–306.
7 Perler,F.B., Davis,E.O., Dean,G.E., Gimble,F.S., Jack,W.E., Neff,N., Noren,C.J., Thorner,J. and Belfort,M. (1994) *Nucleic Acids Res*., **22**, 1125–1127.
8 Chong,S.R., Shao,Y., Paulus,H., Benner,J., Perler,F.B. and Xu,M.Q. (1996) *J. Biol. Chem*., **271**, 22159–22168.
9 Kawasaki,M., Satow,Y., Ohya,Y. and Anraku,Y. (1997) *FEBS Lett*., **412**, 518–520.
10 Gimble,F.S. and Thorner,J. (1993) *J. Biol. Chem*., **268**, 21844–21853.
11 Gimble,F.S. and Wang,J. (1996) *J. Mol. Biol*., **263**, 163–180.
12 Wende,W., Grindl,W., Christ,F., Pingoud,A. and Pingoud,V. (1996) *Nucleic Acids Res*., **24**, 4123–4132.
13 Duan,X., Gimble,F.S. and Quiocho,F.A. (1997) *Cell*, **89**, 555–564.
14 Xu,M.Q. and Perler,F.B. (1996) *EMBO J*., **15**, 5146–5153.
15 Kawasaki,M., Nogami,S., Satow,Y., Ohya,Y. and Anraku,Y. (1997) *J. Biol. Chem*., **272**, 15668–15674.
16 Derbyshire,V., Wood,D.W., Wu,W., Dansereau,J.T., Dalgaard,J.Z. and Belfort,M. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 11466–11471.
17 Chong,S.R. and Xu,M.Q. (1997) *J. Biol. Chem*., **272**, 15587–15590.
18 Hodges,R.A., Perler,F.B., Noren,C.J. and Jack,W.E. (1992) *Nucleic Acids Res*., **20**, 6153–6157.
19 Dalgaard,J.Z., Moser,M.J., Hughey,R. and Mian,I.S. (1997) *J. Comput. Biol*., **4**, 193–214.
20 Klabunde,T., Sharma,S., Telenti,A., Jacobs,W.R. and Sacchettini,J.C. (1998) *Nature Struct. Biol*., **5**, 31–36.
21 Koonin,E.V. (1995) *Trends Biochem. Sci*., **20**, 141–142.
22 Pietrokovski,S. (1998) *Protein Sci*., **7**, 64–71.
23 Hall,T.M.T., Porter,J.A., Young,K.E., Koonin,E.V., Beachy,P.A. and Leahy,D.J. (1997) *Cell*, **91**, 85–97.
24 Gimble,F.S. and Stephens,B.W. (1995) *J. Biol. Chem*., **270**, 5849–5856.
25 Lykke-Andersen,J., Garrett,R.A. and Kjems,J. (1996) *Nucleic Acids Res*., **24**, 3982–3989.
26 Szczepanek,T. and Lazowska,J. (1996) *EMBO J*., **15**, 3758–3767.
27 Lykke-Andersen,J., Garrett,R.A. and Kjems,J. (1997) *EMBO J*., **16**, 3272–3281.
28 Turmel,M., Otis,C., Cote,V. and Lemieux,C. (1997) *Nucleic Acids Res*., **25**, 2610–2619.
29 Heath,P.J., Stephens,K.M., Monnat,R.J. and Stoddard,B.L. (1997) *Nature Struct. Biol*., **4**, 468–476.
30 Dalgaard,J.Z., Klar,A.J., Moser,M.J., Holley,W.R., Chatterjee,A. and Mian,I.S. (1997) *Nucleic Acids Res*., **25**, 4626–4638.
31 Pace,C.N., Vajdos,F., Fee,L., Grimsley,G. and Gray,T. (1995) *Protein Sci*., **4**, 2411–2423.
32 Kim,J., Zwieb,C., Wu,C. and Adhya,S. (1989) *Gene*, **85**, 15–23.
33 Rost,B. and Sander,C. (1993) *J. Mol. Biol*., **232**, 584–599.
34 Henke,R.M., Butow,R.A. and Perlman,P.S. (1995) *EMBO J*., **14**, 5094–5099.
35 Derbyshire,V., Kowalski,J.C., Dansereau,J.T., Hauer,C.R. and Belfort,M. (1997) *J. Mol. Biol*., **265**, 494–506.
36 Li,L., Wu,L.P. and Chandrasegaran,S. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 4275–4279.