Chimeras of the Homing Endonuclease PI-SceI and the Homologous *Candida tropicalis* Intein: A Study to Explore the Possibility of Exchanging DNA-Binding Modules to Obtain Highly Specific Endonucleases with Altered Specificity

Shawn Steuer, Vera Pingoud, Alfred Pingoud, and Wolfgang Wende*[a]

Homing endonucleases are extremely specific endodeoxyribonucleases. In vivo, these enzymes confer mobility on their genes by inducing a very specific double-strand cut in cognate alleles that lack the cooling sequence for the homing endonuclease; the cellular repair of the double-strand break with the endonuclease-containing allele as a template leads to integration of the endonuclease gene, completing the homing process. As a result of their extreme sequence specificity, homing endonucleases are promising tools for genome engineering. For this purpose, it is desirable to design enzymes with defined new specificities. To analyse which DNA-binding elements are potential candidates for use in the design of enzymes with modified or even new specificity, we produced several chimeric proteins derived from the Saccharomyces cerevisiae *VMA1 intein (PI-SceI) and the related* Candida tropicalis *VMA1 intein. Although the mature* Candida *intein is devoid of endonucleolytic activity, the exchange of two DNA-binding modules of PI-SceI with the homologous elements from the* Candida *intein results in an active endonuclease. The low sequence homology in these modules indicates that different protein – DNA contacts are responsible for the recognition of related DNA sequences. This flexibility in DNA recognition should, in principle, allow endonucleases to be produced with new specificities useful for genome engineering.*

Introduction

Homing endonucleases are remarkably site-specific endonucleases typically found as inteins or encoded by introns.[1–4] In vivo, the activity of these enzymes results in a duplicative and unidirectional transfer of their own intron/intein into the host gene with the intron/inteinless allele, a process called homing.[5, 6]

In the first step of this process, homing endonucleases recognise an extended sequence of up to 40 bp in length that spans both flanks of the corresponding intron/intein insertion site, and cleave the DNA in both strands in the presence of Mg^{2+} ions. Genes that already contain introns/inteins are resistant to their own intron/intein-encoded endonucleases because the recognition sequence is interrupted by the intron/intein. Subsequent repair of the cleaved homing site by the cellular double-strand-break repair machinery with the homologous intron/intein-containing allele as a template results in the integration of the intron/intein into the cleaved allele. The activity of the homing endonuclease leads to occupation of nearly all putative insertion sites of a population by the homing endonuclease genes, which have often been described as selfish mobile genetic elements.

Most homing endonucleases confer no obvious benefits on the host organism. Therefore, without a selective pressure, the homing endonuclease genes accumulate mutations that result in an inactive variant of the homing endonuclease[7–9] or, occasionally, evolve an enzyme with a new, physiologically relevant function, such as the mating type switch endonuclease F-SceI, formerly called HO-endonuclease.[10, 11]

The remarkably long recognition sequences of homing endonucleases guarantee that only the specific insertion site in the genome is cleaved. Nevertheless, these enzymes tolerate some base variation at their homing site, which ensures their propagation despite evolutionary drift of their target sequences, and allows for the invasion of new host genomes by horizontal transmission.[7, 12]

Homing endonucleases are phylogenetically widespread: they occur in bacteria, archaea and eukaryotes. By considering conserved sequence motifs, four homing endonuclease families have been identified: LAGLIDADG, GIY-YIG, H-N-H and His-Cys box. The last two families belong to the ββαMe-finger family of

[a] *Dipl.-Biol. S. Steuer, Dr. V. Pingoud, Prof. Dr. A. Pingoud, Dr. W. Wende*
Justus-Liebig-Universität
Institut für Biochemie FB08
Heinrich-Buff-Ring 58, 35392 Giessen (Germany)
Fax: (+49) 641-9935409
E-mail: wolfgang.wende@chemie.bio.uni-giessen.de

  DOI: 10.1002/cbic.200300718

endonucleases.[4] Members of the largest and best-characterised family contain two copies of the conserved LAGLIDADG sequence motif, present in two identical subunits or in one polypeptide chain with internal structural homology. Crystallographic analyses show that this sequence motif is localised on an α helix and that two of these helices form the core interface between the two catalytic subdomains of the endonuclease, an architecture which has been found in all structures of LAGLIDADG homing endonucleases so far elucidated.[13–16] The catalytic domain of the monomeric homing endonucleases, such as PI-SceI, PI-PfuI or I-DmoI, contains two copies of the LAGLIDADG motif arranged in a pseudosymmetrical manner that reflects the structure of the homodimeric homing endonucleases.

Mutational analysis and genetic studies have revealed that the LAGLIDADG sequence is involved in catalysis; amino acid substitutions, in particular those of the conserved acidic amino acid residue in the penultimate position of this motif, lead to an inactive endonuclease.[17–19] It has been shown that this amino acid residue is involved in coordinating the divalent metal ion cofactor essential for the endonucleolytic reaction.[20–22]

The common structural architecture of the LAGLIDADG homing endonucleases, which consists of defined subdomains, makes these enzymes ideal targets for the generation of new artificial endonucleases by protein design. Two groups have already succeeded in producing a chimeric endonuclease by fusing subdomains of I-CreI and I-DmoI.[23, 24] These engineered enzymes cleave a composite site derived from the I-CreI and I-DmoI recognition sites, with only a slightly reduced catalytic activity compared to the parent enzymes.

One of the best-characterised LAGLIDADG homing endonucleases is PI-SceI from *Saccharomyces cerevisiae*, an intein located in the catalytic subunit of the vacuolar H^+-ATPase (VMA1 intein). The free endonuclease is produced from the VMA1 preprotein by an autocatalytic protein splicing reaction.[25] Since the mature vacuolar H^+-ATPase is essential for the cell, dysfunction of the protein splicing reaction is lethal. Thus PI-SceI combines two catalytic functions: a protein-splicing and an endonucleolytic activity. This dual function is reflected in the bipartite structure of PI-SceI, which consists of two separate domains with very different architecture and function, as shown by the crystal structure (Figure 1).[14, 26] The protein-splicing domain (Domain I) is structurally related to the GyrA mini-intein from *Mycobacterium xenopi* and to the *Drosophila* hedgehog protein autoprocessing domain.[27–29] In addition to the protein-splicing core, Domain I of PI-SceI harbours two regions involved in specific DNA binding: the DNA-recognition region (DRR) and the loop comprising amino acids residues 53 – 70. The two catalytic centers, which include both LAGLIDADG motifs and other residues also responsible for specific DNA binding, are located in the endonucleolytic domain (Domain II).

PI-SceI shares 35% of its amino acid sequence with VMA1 intein from *Candida tropicalis* (Ctr), a pathogenic yeast.[30] The sequence homology of the two inteins is mainly observed in the protein-splicing domain (Figure 2). However, comparison of the two insertion sites reveals that both inteins have to cleave nearly identical target sites to initiate the homing process (Figure 3).[31]

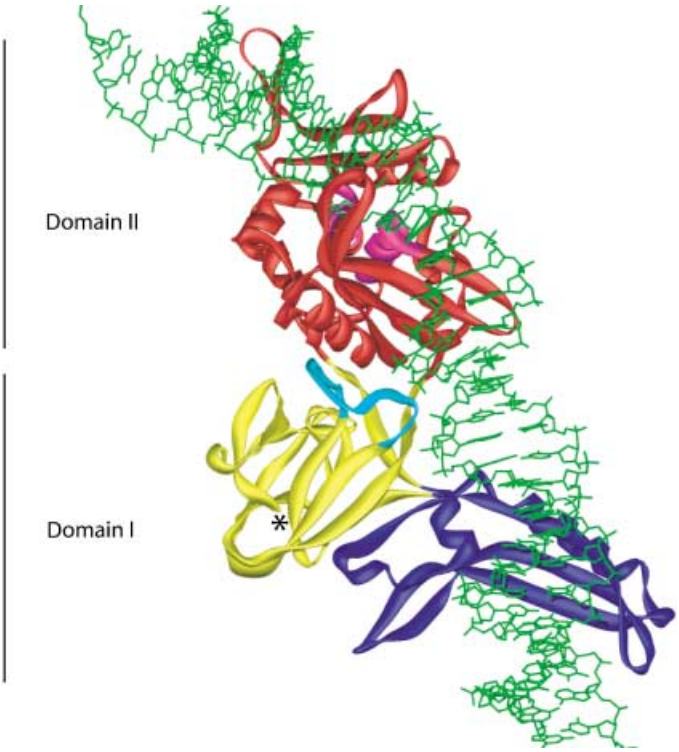

Figure 1. *Cocrystal structure of the homing endonuclease PI-SceI in complex with its specific substrate.[26] The ribbon presentation shows the endonucleolytic domain (Domain II, red), which contains the two LAGLIDADG motifs (magenta). Domain I of PI-SceI contains the protein-splicing core (yellow), the DNA-recognition region (blue) and the PI-SceI DNA-binding loop, which consists of amino acid residues 53 – 70 (cyan). The specific DNA target is coloured green. The position of the N and C termini of PI-SceI, which lie close together, is indicated by an asterisk.*

This observation indicates that both inteins recognise nearly the same DNA sequence by a different set of amino acid residues, as recently described for I-CreI and I-MsoI.[15] This flexibility in the protein – DNA interface, which is necessary to maintain specific contacts, the tolerance to single-base substitutions in the recognition sequence and the fact that parts of the specific DNA-recognition region are separated from the catalytic subdomain make homing endonucleases derived from a protein precursor ideal targets for use in the design of highly site-specific endonucleases with new specificities.

In this study, we engineered different chimeric endonucleases composed of parts from the PI-SceI homing endonuclease and the Ctr intein to investigate whether DNA-recognition modules can be exchanged between related homing endonucleases.

Results

Cloning and characterisation of the *C. tropicalis* VMA1 intein

One of the best-characterised homing endonucleases is PI-SceI from *S. cerevisiae*. The enzyme is autocatalytically excised from the translation product of the VMA1 gene. It has been shown that other yeast strains have an intein in their VMA1 gene product, but little is known about whether these proteins show endonuclease activity.[7, 8, 31] The Ctr intein was chosen in addition

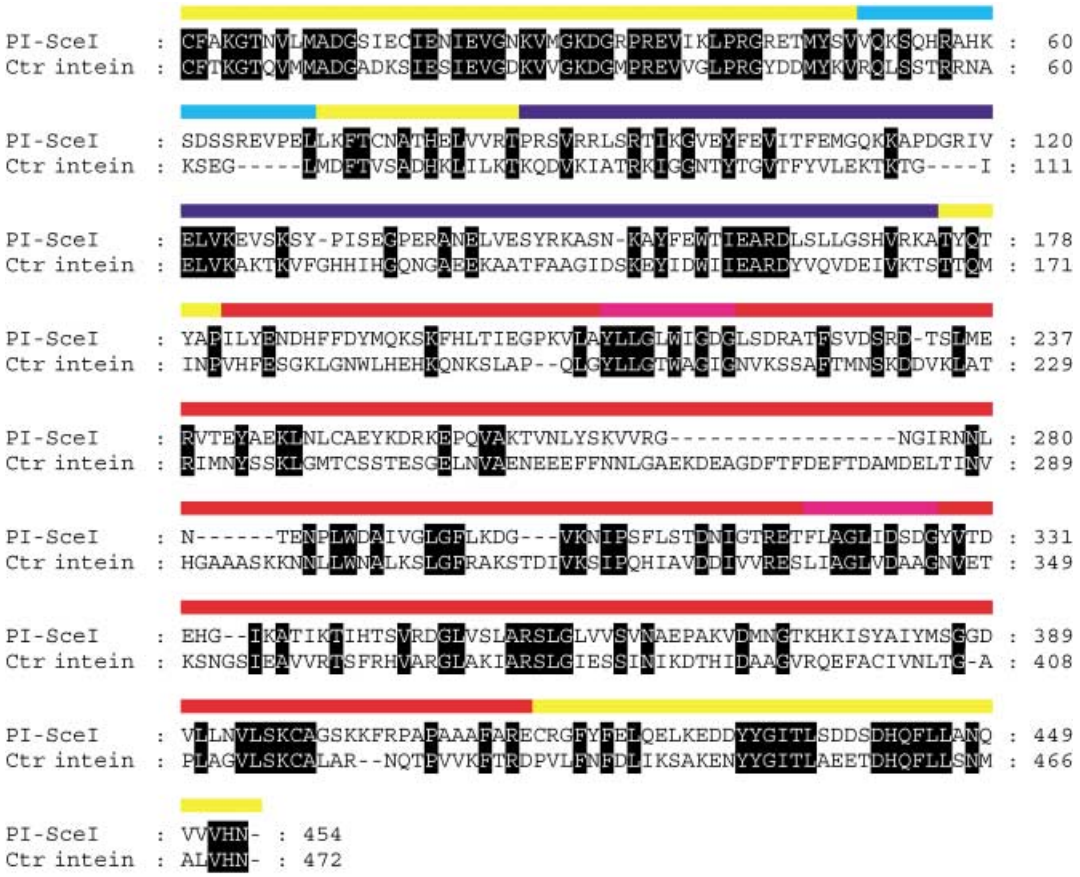

Figure 2. *Amino acid alignment of PI-SceI and the Ctr intein. The alignment shows that 35 % of the amino acids in the sequences are identical (black background). The coloured bars above the sequences indicate the protein architecture of PI-SceI as shown in Figure 1: protein-splicing domain, yellow; DNA-binding loop, cyan; DNA-recognition region, blue; endonuclease domain, red with the LAGLIDADG motifs coloured magenta.*

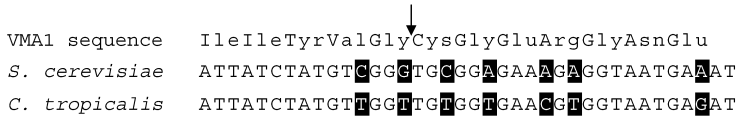

Figure 3. *Alignment of the VMA1 intein insertion site of* S. cerevisiae *with that of* C. tropicalis. *The insertion sites of the inteins are indicated by an arrow. The bases that differ between the sequences are highlighted; these differences do not change the amino acid sequence of the VMA1 protein (top).*

to PI-SceI for studying the feasibility of exchanging DNA-recognition modules between these related homing endonucleases because *C. tropicalis* shows only a weak phylogenetic relation to *S. cerevisiae*, as determined from rRNA sequences.[7]

The VMA1 intein from *C. tropicalis* was cloned and expressed in *Escherichia coli*. The His-tagged version of the Ctr intein was purified by metal-affinity chromatography to near homogeneity, as judged by SDS-PAGE. The isolated Ctr intein turned out to be inactive with regard to cleaving supercoiled or linear plasmid DNA containing the Ctr intein insertion site in the presence of either Mg^{2+} or Mn^{2+} ions. This result was not unexpected because the highly conserved acidic amino acid residues (underlined) in the two LAGLIDADG motifs are replaced by hydrophobic residues in the Ctr intein sequence (I209, A344). In an effort to reactivate the Ctr intein, we produced the Ctr intein variant I209D/A344D. However, this variant displayed neither endonuclease activity nor specific DNA binding to the Ctr intein insertion site, as shown by an electrophoretic mobility shift assay (data not shown).

Construction of PI-SceI/Ctr intein chimeras

To minimise the risk that the substitution of DNA-recognition elements in PI-SceI with corresponding elements from the Ctr intein might disturb the coupling of recognition and catalysis in PI-SceI, only the DNA-binding modules in the protein-splicing domain were considered for recombination. The composition of each chimeric protein is shown in schematic form in Figure 4.

In the PI-SceI/DRR Ctr intein construct, the coding region for the DNA-recognition region of PI-SceI (amino acid residues 86 – 174; shown in blue in Figures 1 and 2) was exchanged with the homologous region of the Ctr intein sequence (encoding amino acid residues 89 – 174). A similar exchange was made with the PI-SceI DNA-binding loop (amino acid residues 51 – 70; shown in cyan in Figures 1 and 2) to give the PI-SceI/loop Ctr intein variant, which harbours the loop from the Ctr intein (residues 51 – 64). To investigate whether this loop has a significant DNA-binding

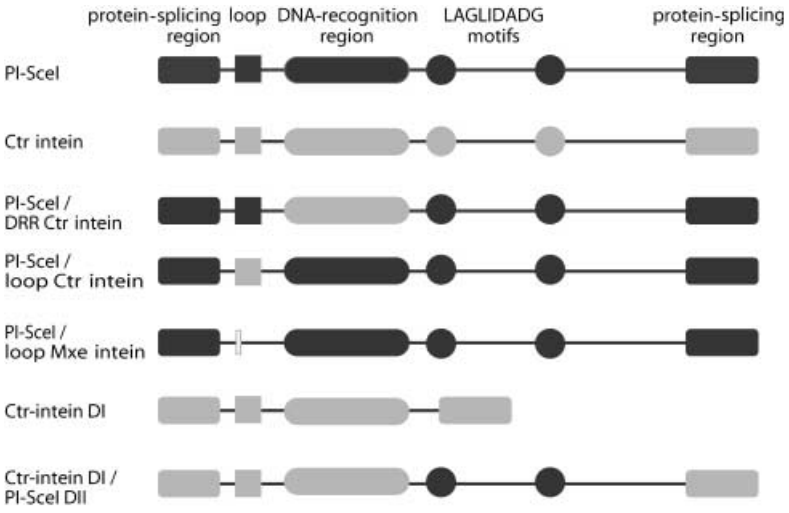

Figure 4. *Overview of the chimeric homing endonucleases based on PI-SceI and the Ctr intein constructed in this work.*

function, the PI-SceI/loop Mxe intein construct was generated. This chimera contains the very short loop (RTVEG) of the *Mycobacterium xenopi* GyrA mini-intein, a protein that displays protein splicing but no endonuclease activity.

It has been shown previously that the isolated protein-splicing domain of PI-SceI binds specifically to the recognition sequence.[32] To verify whether this is also the case for the Ctr intein, the protein-splicing domain of the Ctr intein (Ctr-intein DI) was isolated. Since the mature Ctr intein is inactive, we also fused the active endonucleolytic domain of PI-SceI with the protein-splicing domain of the Ctr intein (Ctr-intein DI/PI-SceI DII).

Production of the chimeric proteins

All chimeric proteins were cloned and the integrity of the coding sequences was checked by sequencing. The proteins were overexpressed in *E. coli* and purified to near homogeneity by metal-affinity chromatography. Circular dichroism spectra were recorded to check whether the chimeric proteins were properly folded. These CD spectra (Figure 5) indicate that the proteins are folded and have an overall secondary structure not too different from that of PI-SceI. In particular, the PI-SceI/Ctr intein chimeras have similar CD spectra to that of PI-SceI, which suggests that the chimeric proteins have comparable structures to that of PI-SceI.

Binding activity of the chimeric proteins

The DNA-binding ability of the protein constructs was tested by electrophoretic mobility shift assays with DNA substrates containing the PI-SceI homing site or the Ctr intein insertion site. Two variants, PI-SceI/DRR Ctr intein and PI-SceI/loop Ctr intein, show strong specific DNA-binding activity (Figure 6). No difference between the two DNA substrates was detectable. Surprisingly, the PI-SceI/loop Ctr intein displays fivefold stronger binding than wild-type PI-SceI and also binds more strongly than the PI-SceI/DRR Ctr intein variant. The PI-SceI/loop Mxe intein construct displays only very weak specific DNA binding ability that becomes apparent at high protein concentrations (100 nM). The other variants did not bind the insertion site of PI-SceI or that of Ctr intein specifically.

Cleavage activity of the chimeric proteins

The cleavage activity of the variants was examined with supercoiled plasmid DNA or linear DNA substrates containing the PI-SceI or the Ctr intein insertion site. The results of the cleavage assays with the linear DNA fragments are shown in Table 1. Only the PI-SceI/DRR Ctr intein and the PI-SceI/loop Ctr

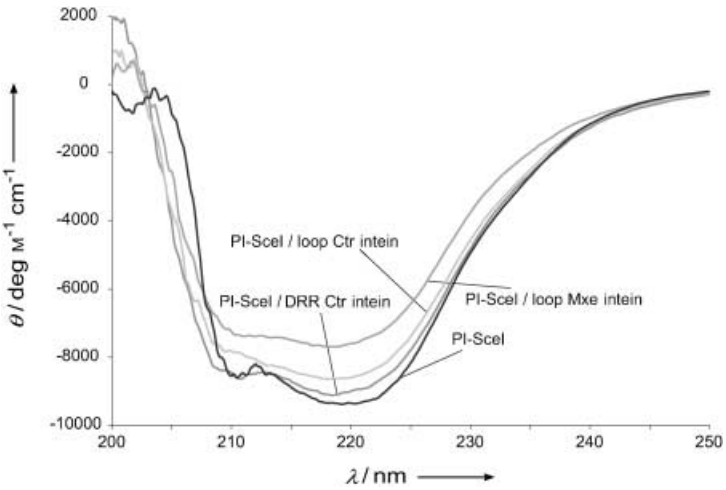

Figure 5. *Circular dichroism spectra of the chimeric homing endonucleases. The PI-SceI/Ctr intein chimeras all show very similar CD spectra. The PI-SceI variant with the loop from the* Mycobacterium xenopi *GyrA mini-intein shows less α-helical structure, which is indicative of some structural rearrangement.*

Table 1. *Summary of the results of the endonuclease cleavage assays.[a]*

Protein	*Saccharomyces cerevisiae* homing site	*Candida tropicalis* homing site
PI-SceI	100[b]	100[b]
Ctr intein	0	0
Ctr intein (I209D/A343D)	0	0
PI-SceI/DRR Ctr intein	98	120
PI-SceI/loop Ctr intein	99	96
PI-SceI/loop Mxe intein	<1	<1
Ctr-intein DI	0	0
Ctr-intein DI/PI-SceI DII	0	0

[a] Endonuclease activity is shown as a percentage of the wild-type (PI-SceI) activity. [b] Cleavage rate constant $= (0.7 \pm 0.08) \times 10^{-3}$ min^{-1}.

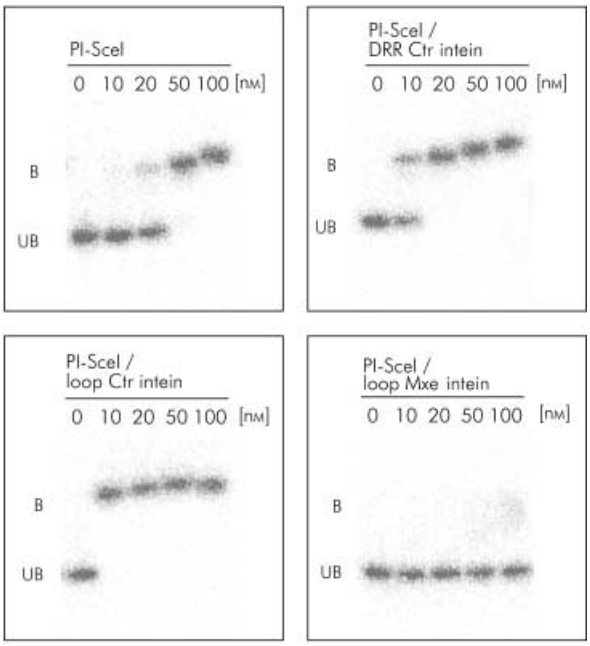

Figure 6. *DNA binding of the chimeric homing endonucleases. Autoradiograms of electrophoretic mobility shift gels are shown to demonstrate the binding of PI-SceI and its variants to the PI-SceI homing site (311-mer). UB, free 311-mer; B, protein/DNA complex.*

intein fusion proteins show measurable specific cleavage activity (Figure 7). The PI-SceI/DRR Ctr intein enzyme has a slight preference for the *C. tropicalis* homing site. No endonucleolytic activity could be detected for the other variants, even in the presence of Mn^{2+} as a cofactor, which stimulates the activity of PI-SceI (Table 1).[33]

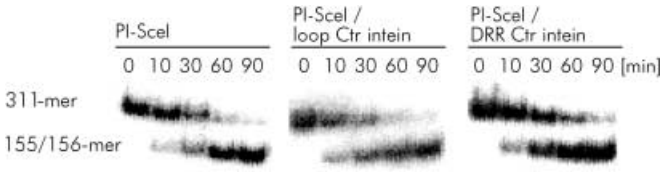

Figure 7. *Cleavage activities of the chimeric homing endonucleases: time course of DNA digestion. Autoradiogram of a gel showing the time course of the digestion of a 311-mer containing the PI-SceI homing site to give two cleavage products of 155 and 156 bp.*

Discussion

Homing endonucleases have a remarkably high specificity for DNA cleavage (for example, PI-SceI cleaves the yeast genome only once)[34, 35] and therefore could be of use for genome engineering.[36] To facilitate gene targeting, endonucleases with new, defined specificities are required. Homing endonucleases are of particular interest in this respect. Two groups have already succeeded in producing artificial endonucleases by fusing domains of the LAGLIDADG homing endonucleases I-CreI and I-DmoI.[23, 24] One of the new engineered enzymes, E-DreI, cleaves a chimeric recognition site with a comparable rate to that obtained with the parent enzymes. The approach presented herein takes advantage of the modular architecture of the LAGLIDADG intein homing endonucleases. A hypothetical scheme of evolution of these enzymes is shown in Figure 8. In

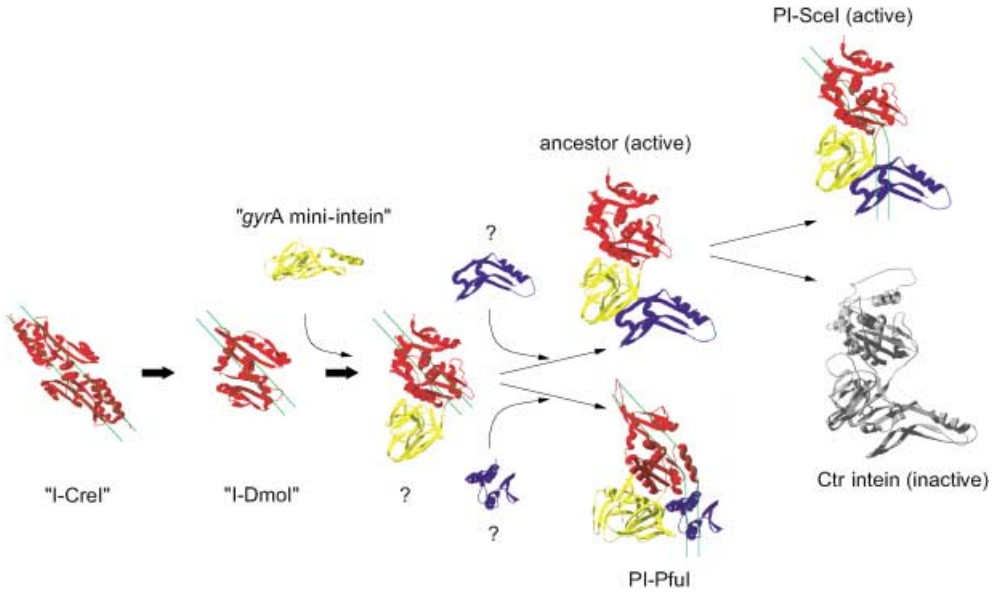

Figure 8. *A hypothetical scheme for the evolution of the LAGLIDADG family of homing endonucleases. The scheme is based on the structures of various homing endonucleases and structural models (Ctr intein) representing putative ancestor proteins. Starting from a homodimeric enzyme (red) such as I-CreI, a monomeric enzyme, such as I-DmoI, evolved. By acquisition of a protein-splicing domain (yellow) like the gyrA mini-intein, a homing endonuclease was generated, which allowed invasion of protein-encoding open reading frames. Fusion with other DNA-binding (blue) domains expanded the range of possible target sequences to give enzymes like PI-SceI and PI-PfuI. The lack of selection pressure on the endonucleases in the host organism may result in an accumulation of mutations that inactivate the endonucleases but preserve the protein-splicing activity, as observed for the Ctr intein (grey). Another possibility is that the mutations allow the alteration of the DNA recognition site to enable the homing endonuclease to invade new hosts (not shown).*

contrast to the intron-encoded homing endonucleases, in which the DNA-recognition regions are directly associated with the catalytic centers (I-CreI and I-DmoI in Figure 8), the intein homing endonucleases acquired additional DNA-binding modules located in the protein-splicing domain during evolution (PI-SceI and PI-PfuI in Figure 8). All known structures of the endonucleolytic domains and protein-splicing core domains are conserved, but the additional DNA-binding modules (Figure 8, shown in blue) display no structural or sequence homology, which suggests that they are interchangeable to a certain extent. The long-term goal of our approach is to utilise this modular architecture to generate endonucleases with new specificities.

To investigate whether these modules can be exchanged between related homing endonucleases, we chose to use the DNA-binding modules of PI-SceI, the VMA1 intein from *Saccharomyces cerevisiae*. Many other VMA1 inteins have sequence identities with PI-SceI of greater than 95% and identities are particularly high for inteins from other *Saccharomyces* strains.[8] We therefore selected the VMA1 intein from *Candida tropicalis*, a yeast distantly related to *Saccharomyces cerevisiae*, for our studies. The *S. cerevisiae* and *C. tropicalis* VMA1 inteins have only 35% sequence identity (Figure 2).

The cloned, expressed and purified Ctr intein was found to be inactive for DNA cleavage. This result was not unexpected because selection in the host population only retains the protein-splicing capacity of the intein, as has been demonstrated for the Ctr intein.[30] The sequence alignment in Figure 2 shows that the protein-splicing core is highly conserved; other regions have accumulated mutations that have inactivated the endonucleolytic function of the enzyme. In the Ctr intein, two catalytically essential acidic amino acids from the LAGLIDADG motif are replaced by hydrophobic residues. To test whether we could reactivate the enzyme, we reintroduced the functional acidic amino acid residues, but the Ctr intein variant I209D/A344D is still inactive with respect to both DNA cleavage and binding. We conclude that additional mutations must have inactivated the enzyme. A likely candidate region is a loop (residue 258 – 285) mainly composed of acidic amino acids that may block DNA binding to the catalytic domain, as indicated by the structure model of the Ctr intein[37].

The reason for the inactivity of the Ctr intein is not only the alteration in the endonucleolytic domain. This fact is shown by the binding-deficiency of the variant generated by the fusion of the active endonucleolytic domain of PI-SceI and the protein-splicing domain of the Ctr intein (Ctr intein DI/PI-SceI DII). To find out whether this binding deficiency was caused by lack of the proper interdomain contacts, we produced the protein-splicing domain of the Ctr intein in isolation, since the isolated homologous PI-SceI domain binds specifically to the homing site.[32] No DNA binding of the Ctr intein protein-splicing domain could be detected.

We therefore wondered whether it would be possible to obtain an active homing endonuclease if individual DNA-binding modules in the protein-splicing domain were exchanged between PI-SceI and the Ctr intein. Two different regions in PI-SceI were exchanged for the structurally homologous parts from the Ctr intein: the DNA-recognition region and the loop containing amino acids residues 53 – 70 (see Figure 1). The replacement of the loop by the Ctr intein loop did not significantly affect the activity and specificity of the endonuclease. In contrast, the insertion of the short loop found in the GyrA mini-intein from *Mycobacterium xenopi* into PI-SceI leads to a nearly inactive enzyme. This observation emphasises the importance of this flexible loop in DNA binding and is consistent with results obtain by site-directed mutagenesis[38] and with the structural analysis of PI-SceI in complex with the recognition site.[26]

The most promising results were obtained by the exchange of the DRR region. The PI-SceI variant containing the DNA-recognition region from the Ctr intein cleaves the PI-SceI homing site with a rate comparable to that achieved by wild-type PI-SceI. This chimeric endonuclease displays significantly higher binding and cleavage activity for the putative *C. tropicalis* homing site (Table 1). The high specificity and activity of this variant is remarkable when one considers that of the 93 amino acid residues exchanged less than 25% are conserved residues. Results from mutagenesis experiments[38] and the structure of PI-SceI in complex with the specific substrate[26] indicate that most of the conserved residues, with the exception of R94 in PI-SceI, are not vital for the function of PI-SceI. This underlines the flexible site-recognition strategy of homing endonucleases.

Conclusion

The results presented herein show that the additional DNA-binding modules in the protein-splicing domain, that is, the DRR and loop regions, are ideal targets for use in the generation of highly specific endonucleases with new specificities. Such endonucleases can probably be most effectively produced by random mutagenesis coupled to effective selection procedures.[39, 40] It remains to be seen whether it is possible to exchange unrelated DNA-binding modules from different inteins, such as PI-PfuI and PI-SceI, and thereby to achieve a change in specificity.

Experimental Section

Cloning of the Ctr intein and chimeric proteins: DNA from *Candida tropicalis* was obtained from the Institute for Medical Microbiology, Justus-Liebig University, Giessen, Germany. By using this DNA as a template, a 1437-bp PCR fragment containing the Ctr intein was generated with the primers 5′-CGCGGATCCTGTTTCACTAAAGGTACTCAAGTCATG-3′/5′-CCGGCGTCGACGTCAGTTGTGCACCAAGGCCATGTTGGA-3′. After purification, the PCR product was digested with BamHI and SalI and ligated into pHisPI-SceI,[41] which was cleaved with the same enzymes. The resulting plasmid, pHisCtr, codes for the Ctr intein with an additional N-terminal affinity tag, $Met(His)_6GlySerAla$. The Ctr intein obtained differs in three positions (M28V, V371A, H375R) from the sequence given by Gu et al.[30] The Ctr intein variant I209D/A344D was produced by a PCR-based site-directed mutagenesis technique essentially as described by Kirsch and Joly[42] by using the primer 5′-GTTGGGTACTTGGGCCGGCGATGGAAATGTTAAATC-3′ to introduce the mutation I209D, and subsequently 5′-

GGTTTCAACATTACCATCAGCGTCAACTAAACCGGCAATC-3′ to create the A344D mutation.

To obtain the plasmid coding for the protein-splicing domain of the Ctr intein, two PCR reactions were performed with pHisCtr as a template and the primers 5′-TTGAATCCGGACCAGTTGTCAAATTTAC-CAG-3′/5′-AATAGGCGTATCACGAGGCCCTTTC-3′ (primerI) to obtain one PCR product and 5′-ACTGGTCCGGATTCAAAATGAACTGGGTTG′-3′/5′-TCAACAGGAGTCCAAGCTCAGCTAA-3′ (primerII) to obtain the second PCR product. Both gel-purified PCR products were digested with BspEI and the fragments were subsequently ligated with each other. The resulting ligation product was cleaved with BamHI and SalI and inserted into the plasmid pHisPI-SceI as described before.

To obtain the PI-SceI/loop Ctr intein construct, two PCR fragments derived from the plasmid pHisPI-SceI were cleaved with ScaI, ligated and inserted into pHisPI-SceI. The two PCR fragments were produced by using the following primers: primerI/5′-CGTCTAGTACTAGAAAGTT-GACGGACACTGTACATAGTTTCACTTCC-3′ and primerII/5′-TTCTAG-TACTAGACGTAATGCTAAATCCGAAGGCCTCAAGTTCACGTGTAATGC-GAC-3′. Essentially the same technique was used to generate the plasmid coding for the PI-SceI/DRR Ctr intein construct, with the exception that three PCR products were ligated. Products A and C were generated with pHisPI-SceI as a template by using the primers primerI/5′-ACGAGGCGTACGAACAACCAACTCATGAGTCGC-3′ (product A) and primerII/5′-CGCGAACTAGTACCTACCAGACTTACGCTCCA-3′(product C). PCR product B was obtained from pHisCtr with 5′-AAGATTCGTACGCGTAAAATTGGTGGCAACAC-3′/5′-AGTGGTAC-TAGTCTTGACAATTTCATCAAC-3′ as primers. Products A and B were digested with SpII and subsequently ligated with each other. Product C was ligated through a SpeI site to the A – B fragment. The resulting fragment was inserted into the plasmid pHisPI-SceI as described before. To create the Ctr-intein DI/PI-SceI DII construct, a PCR was performed on pHisPI-SceI with 5′-CACCACTCAAATGATCAACC-CAATTCTTTATGAGAATGACCACTTTTTC-3′/5′-TCAAATCAAAGTTGAA-CAAAACTGGCTCACGTGCAAAAGCAGCGGCGG-3′ as primers. The resulting PCR product was subsequently used as a megaprimer[42] with pHisPI-CtrI as a template. To generate the plasmid coding for the PI-SceI/loop Mxe intein protein, a megaprimer was used on pHisPI-SceI that was produced from the primers 5′-ACTATGTA-CAGTGTCCGTACTGTTGAAGGCCTCAAGTTCACGTGTAATG-3′/primerII on the same template. All the proteins are His_6-tagged and were expressed and purified as described by Wende et al.[41]

Circular dichroism spectroscopy: Circular dichroism spectra of PI-SceI and its variants were recorded at a protein concentration of 25 μM in a buffer containing 2-[4-(2-hydroxyethyl)-1-piperazinyl]e-thanesulfonic acid-NaOH (10 mM, pH 7.0) and NaCl (30 mM). The measurements were performed in a JASCO J-710 spectrophotometer at ambient temperature.

Cleavage assays: Cleavage assays with the PI-SceI homing site were performed by using a ^{32}P-labelled 311-bp DNA fragment[41] with a central PI-SceI cleavage site, or supercoiled or linearised pBSVDEX[33] plasmid DNA as substrate. For the cleavage analysis of the putative homing site of the Ctr intein, a double-stranded oligonucleotide containing the specific site was inserted into the BamHI site of pAT153. A ^{32}P-labelled 350-bp fragment was obtained by a standard PCR reaction with the primers 5′-ATCGCCAGTCACTATGGCGTGC-3′/5′-TGGCGCCCAACAGTCCCCCGGCC-3′ and ^{32}P-α-dATP. In the cleavage assay, the substrate (8 nM) was incubated with the appropriate proteins in cleavage buffer (10 mM tris(hydroxymethyl)aminomethane (Tris)-HCl, pH 8.5, 100 mM KCl, 1 mM 1,4-dithiothreitol (DTT), 100 μg mL^{-1} bovine serum albumin, 2.5 mM $MgCl_2$) at 37 °C for a measured period of time. The reactions were terminated by the addition of ethylenediaminetetraacetate (EDTA; 20 mM final concentration). The cleavage reactions were analysed on 10% polyacrylamide gels in TPE buffer (89 mM Tris-phosphate, pH 8.3, 2 mM EDTA). After electrophoresis, the gels were dried and the radioactive bands were visualised by autoradiography with an instant imager system (Canberra Packard).

DNA binding assays: Electrophoretic mobility shift assays were performed on the same proteins and DNA constructs as used in the endonuclease assays. The substrate (4 nM) was incubated with binding buffer (10 mM Tris-HCl, pH 7.5, 50 mM KCl, 5 mM $CaCl_2$ 1 mM DTT, 0.05% nonfat dry milk, 5% (v/v) glycerol), the nonspecific DNA carrier poly(dI-dC) (0.1 mg) and protein at ambient temperature for 10 minutes in a volume of 10 μL. Loading buffer (2.5 μL; 10% (w/v) Ficoll, 10% (v/v) glycerol, 0.2% (w/v) bromophenolblue, 0.2% (w/v) xylene cyanol) was then added. Complex formation was analysed by electrophoretic mobility shift assays on 8% polyacrylamide gels in TBE buffer (70 mM Tris-borate, 2 mM EDTA). After electrophoresis, the gels were dried and the radioactive bands visualised by autoradiography with an instant imager system (Canberra Packard).

Acknowledgements

We thank Nadine Böcher and Nadine Thome for expert technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft (Grant no. Pi 151/2-2) and the Fonds der Chemischen Industrie.

Keywords: DNA recognition · endonucleases · gene targeting · genome engineering · protein design

[1] M. Belfort, R. J. Roberts, *Nucleic Acids Res.* **1997**, *25*, 3379 – 3388.
[2] J. E. Mueller, M. Bryk, N. Loizos, M. Belfort, R. S. Lloyd, R. J. Roberts in *Nucleases* (Ed.: S. M. Linn), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **1993**, pp. 111 – 143.
[3] M. S. Jurica, B. L. Stoddard, *Cell. Mol. Life Sci.* **1999**, *55*, 1304 – 1326.
[4] B. S. Chevalier, B. L. Stoddard, *Nucleic Acids Res.* **2001**, *29*, 3757 – 3774.
[5] B. Dujon, M. Belfort, R. A. Butow, C. Jacq, C. Lemieux, P. S. Perlman, V. M. Vogt, *Gene* **1989**, *82*, 115 – 118.
[6] A. M. Lambowitz, M. Belfort, *Annu. Rev. Biochem.* **1993**, *62*, 587 – 622.
[7] V. Koufopanou, M. R. Goddard, A. Burt, *Mol. Biol. Evol.* **2002**, *19*, 239 – 246.
[8] Y. Okuda, D. Sasaki, S. Nogami, Y. Kaneko, Y. Ohya, Y. Anraku, *Yeast* **2003**, *20*, 563 – 573.
[9] M. R. Goddard, A. Burt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13 880 – 13 885.
[10] E. Nahon, D. Raveh, *Nucleic Acids Res.* **1998**, *26*, 1233 – 1239.
[11] R. Kostriken, J. N. Strathern, A. J. Klar, J. B. Hicks, F. Heffron, *Cell* **1983**, *35*, 167 – 174.
[12] F. S. Gimble, *FEMS Microbiol. Lett.* **2000**, *185*, 99 – 107.
[13] K. Ichiyanagi, Y. Ishino, M. Ariyoshi, K. Komori, K. Morikawa, *J. Mol. Biol.* **2000**, *300*, 889 – 901.
[14] X. Duan, F. S. Gimble, F. A. Quiocho, *Cell* **1997**, *89*, 555 – 564.
[15] B. Chevalier, M. Turmel, C. Lemieux, R. J. Monnat, Jr., B. L. Stoddard, *J. Mol. Biol.* **2003**, *329*, 253 – 269.
[16] G. H. Silva, J. Z. Dalgaard, M. Belfort, P. Van Roey, *J. Mol. Biol.* **1999**, *286*, 1123 – 1136.
[17] K. Komori, K. Ichiyanagi, K. Morikawa, Y. Ishino, *Nucleic Acids Res.* **1999**, *27*, 4175 – 4182.
[18] L. M. Seligman, K. M. Stephens, J. H. Savage, R. J. Monnat, *Genetics* **1997**, *147*, 1653 – 1664.
[19] F. S. Gimble, B. W. Stephens, *J. Biol. Chem.* **1995**, *270*, 5849 – 5856.
[20] S. Schoettler, W. Wende, V. Pingoud, A. Pingoud, *Biochemistry* **2000**, *39*, 15 895 – 15 900.
[21] B. S. Chevalier, R. J. Monnat, Jr., B. L. Stoddard, *Nat. Struct. Biol.* **2001**, *8*, 312 – 316.
[22] J. Lykke-Andersen, R. A. Garrett, J. Kjems, *EMBO J.* **1997**, *16*, 3272 – 3281.

[23] J. C. Epinat, S. Arnould, P. Chames, P. Rochaix, D. Desfontaines, C. Puzin, A. Patin, A. Zanghellini, F. Paques, E. Lacroix, *Nucleic Acids Res.* **2003**, *31*, 2952 – 2962.
[24] B. S. Chevalier, T. Kortemme, M. S. Chadsey, D. Baker, R. J. Monnat, B. L. Stoddard, *Mol. Cell* **2002**, *10*, 895 – 905.
[25] P. M. Kane, C. T. Yamashiro, D. F. Wolczyk, N. Neff, M. Goebl, T. H. Stevens, *Science* **1990**, *250*, 651 – 657.
[26] C. M. Moure, F. S. Gimble, F. A. Quiocho, *Nat. Struct. Biol.* **2002**, *9*, 764 – 770.
[27] T. M. T. Hall, J. A. Porter, K. E. Young, E. V. Koonin, P. A. Beachy, D. J. Leahy, *Cell* **1997**, *91*, 85 – 97.
[28] S. Pietrokovski, *Protein Sci.* **1998**, *7*, 64 – 71.
[29] E. V. Koonin, *Trends Biochem. Sci.* **1995**, *20*, 141 – 142.
[30] H. H. Gu, J. Xu, M. Gallagher, G. E. Dean, *J. Biol. Chem.* **1993**, *268*, 7372 – 7381.
[31] F. S. Gimble, *Nucleic Acids Res.* **2001**, *29*, 4215 – 4223.
[32] W. Grindl, W. Wende, V. Pingoud, A. Pingoud, *Nucleic Acids Res.* **1998**, *26*, 1857 – 1862.
[33] F. S. Gimble, J. Thorner, *J. Biol. Chem.* **1993**, *268*, 21 844 – 21 853.
[34] M. C. D. Bremer, F. S. Gimble, J. Thorner, C. L. Smith, *Nucleic Acids Res.* **1992**, *20*, 5484.
[35] F. S. Gimble, J. Thorner, *Nature* **1992**, *357*, 301 – 306.
[36] M. Jasin, *Trends Genet.* **1996**, *12*, 224 – 228.
[37] W. Wende, unpublished results.
[38] W. Wende, S. Schoettler, W. Grindl, F. Christ, S. Steuer, A. J. Noel, V. Pingoud, A. Pingoud, *Mol. Biol. (Mosk)* **2000**, *34*, 1054 – 1064.
[39] A. Pingoud, A. Noel, V. Pingoud, S. Steuer, W. Wende in *Enzyme Functionality: Design, Engineering and Screening* (Ed.: A. Svendsen), Marcel Dekker, New York, **2003**, pp. 325 – 352.
[40] M. Gruen, K. Chang, I. Serbanescu, D. R. Liu, *Nucleic Acids Res.* **2002**, *30*, e29.
[41] W. Wende, W. Grindl, F. Christ, A. Pingoud, V. Pingoud, *Nucleic Acids Res.* **1996**, *24*, 4123 – 4132.
[42] R. D. Kirsch, E. Joly, *Nucleic. Acids Res.* **1998**, *26*, 1848 – 1850.

Received: July 18, 2003 [F 718]

