Nucleic Acids Research, 1997, Vol. 25, No. 19 **3767–3776**

Purification, biochemical characterization and protein–DNA interactions of the I-*Cre*I endonuclease produced in *Escherichia coli*

Jingshan Wang, Hyong-Ha Kim, Xiaoqin Yuan and David L. Herrin*

Department of Botany and Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78713, USA

Received July 22, 1997; Revised and Accepted August 11, 1997

ABSTRACT

I-*Cre*I is a member of the LAGLI-DADG family of homing nucleases; however, unlike most members of this family it contains only a single copy of this signature motif. I-*Cre*I was over-expressed in *Escherichia coli*, and a simple purification protocol developed that gave reasonably pure protein in high yield. Size-exclusion chromatography and chemical cross-linking indicated that the protein is a dimer in solution. DNA cleavage by I-*Cre*I was absolutely dependent on Mg^{2+} (or Mn^{2+}), and was inhibited by monovalent cations. I-*Cre*I displayed a surprisingly high temperature optimum (>50°C), with full activity occurring even at 70°C. Interestingly, SDS was needed for efficient release of the cleavage products from the protein, indicating formation of very stable DNA–protein complexes. In contrast to these robust characteristics, purified I-*Cre*I was unstable; however, it could be stabilized by the addition of either target or non-target DNA. Mobility shift assays revealed that I-*Cre*I binds to DNA in the absence of Mg^{2+}. Hydroxyl radical footprinting showed that I-*Cre*I strongly protected the backbone of a continuous stretch of at least 12 nt on each strand that were shifted, relative to each other, by 2 bp in the 3′ direction. Methylation protection and interference analyses were also performed, and together with the hydroxyl radical footprinting, indicate that I-*Cre*I binds in both the major and minor grooves of its target DNA.

INTRODUCTION

Group I introns from a variety of organisms have been found to contain long open reading frames (ORFs) that encode site-specific DNA endonucleases (1–3). Related ORFs in archael introns (4) and in-frame spacers in certain proteins (inteins) have also been shown to encode site-specific DNA endonucleases (reviewed in 5). The natural recognition sequences for these endonucleases are typically long, asymmetric, and unique for each enzyme. The principal function of these highly specific enzymes seems to be to promote the integration of their DNA into cognate sites via homologous recombination (5).

The group I intron-encoded endonucleases typically cleave intron-lacking DNA near the site of intron insertion (i.e. the exon–exon junction) creating a staggered double-strand break which facilitates intron invasion (1,2). This type of gene conversion, which is also known as 'intron homing', can be very efficient, and has been demonstrated in mitochondria, chloroplasts and the nucleus of eucaryotic cells, as well as with bacteriophage (3). The model suggested to explain this recombination process is based on the double strand break-repair mechanism suggested for yeast (1,2), although more than one recombination pathway may be operating in different systems (6,7). Because of their association with intron (and other DNA) homing, these site-specific endonucleases have sometimes been called 'homing endonucleases' (e.g. 5).

I-*Cre*I is an intron-encoded endonuclease (8,9) that promotes homing of the chloroplast 23S rRNA intron, *Cr.LSU*, of *Chlamydomonas reinhardtii* (7,8). It is a member of the family of homing endonucleases known as LAGLI-DADG enzymes, so named because they contain the LAGLI-DADG motif, a semi-conserved peptide first noted among mitochondrial intron ORFs (10). This motif typically occurs twice in each polypeptide, with the two non-identical copies (called P1 and P2, respectively; 10) being separated by 90–120 amino acids (5). The importance of these motifs has been demonstrated by mutagenesis; P1 was shown to be critical for the endonuclease activity of I-*Sce*II (11), PI-*Tli*I (12) and PI-*Sce*I (13). However, while P2 was also required for activity by the PI-*Sce*I endonuclease (13), it was found to be dispensable for the endonuclease activity of I-*Sce*II (11). Interestingly, however, P2 was needed for the maturase activity of I-*Sce*II, which is a bifunctional protein (11). Thus, these data indicate that I-*Sce*II and PI-*Sce*I are quite different enzymes.

Based on the above, it seems likely that the LAGLI-DADG family of proteins will prove to be a diverse group of endonucleases that may have evolved more than one way to specifically interact with and cleave DNA. Along this line, I-*Cre*I differs from the aforementioned enzymes in that it contains only one copy of the LAGLI-DADG motif (5). It is also correspondingly smaller than the two-motif enzymes, raising the possibility that it may have evolved from a two-motif enzyme by losing one of its domains. Alternatively, the two-motif enzymes could have evolved from a

*To whom correspondence should be addressed. Tel: +1 512 471 3843; Fax: +1 512 471 3878; Email: BOHL734@utxvms.cc.utexas.edu

one-motif protein like I-*Cre*I via a duplication event. The recognition sequence for I-*Cre*I is also atypical, because it contains partial dyad symmetry (5). Finally, it is interesting to note that the only other single-LAGLI-DADG endonuclease known, I-*Ceu*I, is also found in a *Chlamydomonas* chloroplast rRNA intron, although it is from a different species, namely *Chlamydomonas eugametos* (14).

We are interested in understanding how I-*Cre*I recognizes and cleaves its target DNA. Previously, using enzyme produced by *in vitro* translation, it was shown that cleavage by I-*Cre*I is staggered, occurring 5 nt downstream of the intron insertion site on the top strand, but only 1 nt downstream on the bottom strand (9,15). Hence, I-*Cre*I generates termini with 3′ extensions of 4 nt (9,15), similar to other LAGLI-DADG endonucleases (5). In addition, cleavage of sequence ladders indicated that the recognition sequence of I-*Cre*I spans the cleavage site and is 19–24 bp in length (9,15).

In order to carry out detailed biochemical studies of I-*Cre*I, it was necessary to produce the protein in *Escherichia coli*, since it occurs at a very low level in *C.reinhardtii* (8). Fortunately, chloroplasts use the universal genetic code, so ORFs such as I-*Cre*I do not have to be altered to efficiently produce the protein in *E.coli*. The results presented here show that I-*Cre*I can accumulate to high levels in *E.coli*. A simple purification protocol was also developed, and the purified protein was characterized with respect to its quaternary structure, activity and *in vitro* stability. Also, the binding of I-*Cre*I to its natural target DNA was assessed using footprinting and interference analyses.

MATERIALS AND METHODS

Solutions

Lysis solution (16): 20 mM Tris–HCl (pH 7.5), 0.5 M NaCl, 10% (v/v) glycerol, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), 5 mg/ml leupeptin, 1% (w/v) aprotinin and 0.1% (v/v) NP-40; solution A: 20 mM Tris–HCl pH 8.0, 5% glycerol; storage solution: 20 mM Tris–HCl pH 7.5, 10% glycerol; standard reaction solution: 20 mM Tris–HCl (pH 9.0), 10 mM $MgCl_2$; standard stop solution (10×): 0.1 M Tris–HCl pH 7.5, 0.25 M EDTA, 5% (w/v) SDS, 0.5 mg/ml proteinase K.

Plasmids

An expression plasmid was engineered to produce I-*Cre*I without extraneous amino acids. Briefly, a 622 bp fragment containing the ORF sequence (17), and flanked by *Nde*I and *Bam*HI restriction sites was prepared using the polymerase chain reaction (PCR); the oligos were d(5′-GCAGCCATATGAATACAAAATATAAT-3′), which contained the start codon, and d(5′-CGGGATCCAGGA-GTCGGCGTATTAT-3′), which annealed to nt 98–113 downstream of the ORF (17). *Taq* DNA polymerase was used according to the manufacturer's (Promega) suggestions, and the plasmid pGEM23S.1 (18) was used as template. The PCR product was cleaved with *Bam*HI and *Nde*I, purified by gel electrophoresis and cloned into *Nde*I–*Bam*HI-digested vector, pAII17 (obtained from W.Jack, New England BioLabs), using the Novablue (Novagen Inc.) strain as host. The new expression plasmid was 6809 bp, and was called pI-*Cre*I.

For DNA footprinting, a 161 bp PCR product that spanned the I-*Cre*I cleavage site (in the fused exons), and containing *Eco*RI and *Bam*HI linkers was generated and subcloned. The oligonucleotides for PCR were as follows: the 5′ oligo, d(5′-CGGAATTCTGTAG-TAGGTCCGAAGGGTT-3′), annealed to nt 47–66 upstream of the intron insertion site, while the 3′ oligo, d(5′-CGGGATCCTT-TCAGGTCCTCTCGTACTA-3′), was complementary to nt 60–79 downstream of the intron insertion site. The plasmid pGEM23S.E (9) was used as the template. The PCR product was cleaved with *Eco*RI and *Bam*HI, and the 151 bp product was gel-purified [4% Nusieve GTG (LMP) agarose], and subcloned into the vector pGEM3zf(+) digested with the same enzymes. The host strain was *E.coli* DH5αF′ (GIBCO/BRL Inc.), and the recombinant plasmid was called pGEM23S.E151 (3329 bp).

Expression and purification of I-*Cre*I

Escherichia coli strain BL21(DE3)pLysS (19) was used for the purification of I-*Cre*I. Clones transformed with the pI-*Cre*I plasmid were grown in 250 ml of Luria Broth containing 200 µg/ml ampicillin and 25 µg/ml chloramphenicol at 37°C with shaking. When the culture reached an OD_{600} of 0.7–1.0, expression was induced by adding IPTG to a final concentration of 1 mM, and after 2.5 h, the cells were harvested by centrifugation.

The following procedures were performed at 4°C unless stated otherwise. The harvested cells were resuspended in 20 ml of ice-cold lysis solution, and then broken in a pre-cooled French Pressure Cell (SLM Aminco, Urbana, IL) with a single passage at 20 000 p.s.i. The lysate was centrifuged at 3000 *g* for 15 min, the supernatant recovered, and then subjected to ultracentrifugation at 100 000 *g* for 50 min. This supernatant, which contained 90% of the protein, was called the 'soluble' fraction. To this fraction, solid ammonium sulfate was slowly added (with stirring) to 40% saturation, and the mixture left on ice for 1 h before centrifugation at 20 000 *g* for 20 min. The supernatant was recovered and additional ammonium sulfate was added to give a final saturation of 80%. After incubation on ice for 1 h, the mixture was centrifuged as before. The pellet was resuspended in 4 ml of solution A, and then dialyzed against 4 l of solution A overnight. The dialyzed fraction was applied to a 1 ml Mono S FPLC column (Pharmacia BioTech, Uppsala, Sweden) at a flow rate of 1 ml/min (Pharmacia FPLC apparatus). After washing the column with 10 ml of solution A, bound proteins were eluted with a 0–0.3 M linear NaCl gradient made up in solution A. Forty 2 ml fractions were collected, and the amounts of protein (20) and I-*Cre*I activity (see below) were determined. The column fractions were also analyzed by SDS–PAGE. The fractions containing most of the I-*Cre*I activity were pooled, dialyzed against 4 l of storage solution overnight, and stored in aliquots at –70°C.

SDS–polyacrylamide gel electrophoresis

SDS–polyacrylamide gel electrophoresis (SDS–PAGE) was performed as described by Laemmli (21) using 15% acrylamide gels, or with the 10–20% gradient polyacrylamide gel system described by Herrin *et al*. (22). Gels were stained with coomassie brilliant blue, or with silver as previously described (22).

Size-exclusion chromatography

Size-exclusion chromatography was performed on an analytical 7.8 × 300 mm BioSep-SEC-S3000 column (Phenomenex, Torrance, CA). The column was equilibrated with 50 mM Na_2HPO_4 pH 7.5, 100 mM NaCl, and calibrated with the following molecular mass standards (Bio-Rad): thyroglobulin (670 kDa), gamma globulin

(158 kDa), ovalbumin (44 kDa), myoglobulin (17 kDa) and vitamin B_{12} (1.35 kDa). An aliquot of 100 µl of purified I-*Cre*I, at a concentration of 7 µM monomer (0.125 mg/ml), was loaded onto the column, which was eluted at a flow rate of 0.5 ml/min (at room temperature), and monitored by UV absorbance. Fifty fractions of 0.25 ml were collected, and analyzed by SDS–PAGE.

Cross-linking

Purified I-*Cre*I was treated with dimethylsuberimidate·HCl (Pierce, Rockford, IL) in 50 mM borate buffer pH 9.5; the elevated pH favors the amidination of the protein over hydrolysis of the imidoester (23). The reactions were carried out at different protein concentrations, and the cross-linker was added to a final concentration of 18 mM (23). After incubation at 37°C for 1 h, equal volumes of each reaction were analyzed by SDS–PAGE.

I-*Cre*I endonuclease activity assays

Plasmid pGEM23S.E (3.9 kb), which was constructed previously (9), was used for the cleavage assays. It contains fused 23S *rrn* exon sequences, and was isolated by banding in a CsCl gradient, or with the Wizard™ Maxipreps DNA Purification System (Promega, Madison, WI). For most experiments, it was linearized with *Sca*I prior to the assay. Standard I-*Cre*I assays were performed at 37°C in the standard reaction solution (see above), stopped with 0.1 vol of 10× standard stop solution, and the products separated by electrophoresis in 0.8% agarose/ethidium bromide gels at room temperature (9). The fluorescence was photographed with Polaroid film (type 667) using a transilluminator (Foto-Prep I, Fotodyne). Quantitative analysis of fluorescent bands was performed with the aid of a FluorImager and ImageQuaNT software (Molecular Dynamics, Inc., Sunnyvale, CA). One unit of endonuclease activity (U) was defined as the amount of I-*Cre*I necessary to cleave 100 ng of target DNA in 20 min at 37°C.

Optimal conditions for DNA cleavage by I-*Cre*I were determined using 1250 ng of substrate DNA in a total reaction volume of 100 µl (5 nM). Aliquots (20 µl) were removed at different times (up to 60 min), and the reactions were terminated and analyzed as described above. The various changes to the reaction conditions (salt, pH, temperature, etc.) are described in the figure legends or in the text.

The effect of different conditions on the release of the DNA products from the protein was assessed by carrying out standard cleavage assays, and then varying either the composition of the 10× stop solution, or the post-reaction temperature, as described in the figure legend. The DNAs were then analyzed by agarose gel electrophoresis as described above.

Stability of I-*Cre*I

Purified I-*Cre*I was diluted with 20 mM Tris–HCl pH 9.0, and pre-incubated at 37°C for up to 70 min in the presence or absence of 10 mM $MgCl_2$, or in the presence of various DNAs. Aliquots were removed at selected times, and the pre-incubated protein was used in standard activity assays. The DNAs used in the *in vitro* stability tests were: (i) the substrate, (ii) the plasmid vector pGEM3zf(+) (Promega BioTech) and (iii) plasmid pSR1. Plasmid pSR1 contains a 7 kb *Bam*HI fragment of chloroplast DNA that maps 5′ to the *Bam*HI fragment that contains the *Cr.LSU* intron (e.g. 9). The vector used for construction of pSR1 was pGEM3zf(+), thus giving the plasmid a total size of 10 kb (A.Thompson, unpublished results).

Mobility shift assay

150 ng (1.5 pmol) of the 161 bp PCR product, or insert from pGEM23S.151 (see above), was incubated with different amounts of purified protein in 20 mM Tris–HCl pH 9.0 at room temperature for 20 min. After addition of native dye solution (i.e. glycerol + xylene cyanol), the protein–DNA complexes were analyzed by electrophoresis at room temperature (90 V) for 40 min in a 2% agarose/ethidium bromide gel (9).

Hydroxyl radical protection and interference analysis

To prepare target DNA with only one strand end-labeled, plasmid pGEM23S.E151 was digested with *Eco*RI or *Bam*HI, followed by dephosphorylation with alkaline phosphatase, and end-labeling with T4 polynucleotide kinase and [γ-^{32}P]ATP (>4500 Ci/mmol). The end-labeled DNA was then subjected to a second restriction (with *Bam*HI or *Eco*RI, respectively), and the 151 bp singly-end-labeled DNA was purified by electrophoresis on a 2% agarose (FMC Inc.) gel as above. Care was taken to avoid damaging the DNA with UV light.

Hydroxyl radical footprinting was performed according to Tullius and Dombroski (24). 120 pmol of I-*Cre*I was incubated with 5 pmol of the end-labeled target DNA in 60 µl of 25 mM Tris–HCl pH 9.0 in a total volume of 60 µl for 20 min at room temperature. To effect cleavage, the mixture was adjusted to contain 0.1 mM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$, 0.2 mM EDTA, 10 mM sodium ascorbate and 0.15% H_2O_2. After 1 min, the reaction was quenched with excess thiourea, and the bound and unbound DNA fractions were separated by electrophoresis on a 2% agarose gel as described above, except the time of electrophoresis was increased to 70 min. The fluorescent bands were recovered from the gel using centrifugal filters (Millipore), and then analyzed by denaturing electrophoresis on an 8% polyacrylamide gel. Hydroxyl radical cleavage of the target DNA in the absence of I-*Cre*I was also performed as a control.

Hydroxyl radical interference analysis was performed as for the protection footprinting, except that the end-labeled DNA was treated with the scission reagents first (for 45 s), ethanol precipitated, redissolved in water, and then incubated with I-*Cre*I as above. The bound and unbound DNA bands were gel-purified, and then analyzed by denaturing PAGE (8% poyacrylamide).

A Maxam–Gilbert G+A reaction of the end-labeled DNA was performed to provide a marker for PAGE, and it was carried out according to the modifications of Bencini *et al*. (25).

Methylation protection and interference analysis

The target DNA (5 pmol), end-labeled on only one strand (see above), was incubated with 120 pmol of I-*Cre*I in 20 mM Tris–HCl, pH 9.0, in a volume of 60 µl at room temperature for 20 min. An aliquot of 2 µl of freshly made 10% dimethylsulfate (DMS) was added, and after 1 min at room temperature the reaction was quenched with 9 µl of 1 M β-mercaptoethanol and immediately loaded on a 2% agarose gel. After electrophoresis for 40 min at room temperature (90 V) and ethidium bromide staining, bound DNA was recovered from the gel, and the strand-scission reaction performed as described (26). After two precipitations with butanol, the DNA was dissolved in 90% formamide and

analyzed on an 8% denaturing polyacrylamide gel. The methylation reaction was also performed on the target DNA in the absence of I-*Cre*I, for comparison.

The methylation interference analysis was performed similarly to the protection analysis, except that the DNA was methylated first, then incubated with I-*Cre*I. After native gel electrophoresis, the bound and unbound DNAs were recovered from the gel, cleaved and analyzed on an 8% denaturing gel.

Denaturing polyacrylamide gel electrophoresis (PAGE) and autoradiography

Denaturing polyacrylamide gel electrophoresis (PAGE) of DNA was performed as described previously (26) using 8% acrylamide gels run at 50°C. The gels were dried, and then exposed to X-ray film (Kodak XAR) with or without intensifying screens. The relevant portions of the autoradiograms of the methylation reactions were scanned with a densitometer (Model 1312, ISCO) connected to a UA-5 detector/recorder (ISCO). The recorder outputs were imaged using a RealTech scanner and Adobe Photoshop (v3.0) on a PowerMac 6100 computer (Apple).

RESULTS

Expression and purification of I-*Cre*I

Preliminary data indicated that I-*Cre*I was toxic to *E.coli* (unpublished results), so the T7 expression system (19) with a modified pET (12) vector was used. Several different host strains were tried, and BL21(DE3)pLysS gave the best yield. Conditions for induction and accumulation of I-*Cre*I were optimized by varying cell density, IPTG concentration and the time of treatment with IPTG (up to 4 h). The level of induction was estimated using activity assays and SDS–PAGE (monomer I-*Cre*I is 18 kDa; 9). Under optimized conditions, I-*Cre*I accumulated to 5–10% of the total protein, and >90% was recovered in the soluble fraction (27). Ammonium sulfate fractionation and cation exchange chromatography were used to purify I-*Cre*I in high yield from the soluble fraction. Figure 1A shows an SDS–PAGE analysis of whole cell proteins from control and induced cells, as well as fractions enriched for the protein. Figure 1B shows the elution profile for the Mono S cation-exchange column; I-*Cre*I eluted as a single peak at 0.15–0.2 M NaCl (fractions 25–30). Pooled fractions were >90% homogeneous for the 18 kDa polypeptide (Fig. 1A, lane 7). SDS–PAGE analysis of all fractions also showed that the 18 kDa band correlated well with DNA cleavage activity (data not shown). The strategy used for purifying I-*Cre*I from the crude soluble fraction is summarized in Table 1. The protocol gave a 10-fold purification of I-*Cre*I with 20% recovery. The preparations specifically cleaved the substrate DNA, and without non-specific degradation of the substrate (see below).

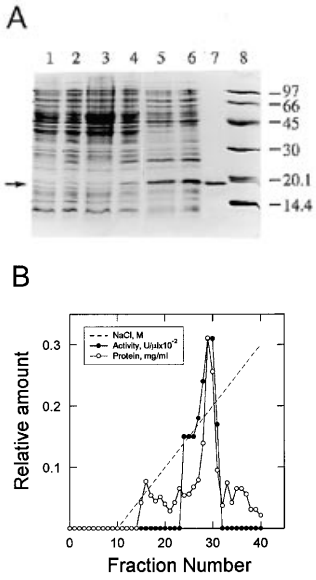

Figure 1. Purification of I-*Cre*I. (**A**) SDS–PAGE analysis of *E.coli* whole cell proteins and various fractions containing I-*Cre*I. The polypeptides were separated on a 15% acrylamide gel, which was stained with coomassie blue. Lane 1, uninduced cells transformed with the pAII17 vector; lane 2, uninduced cells transformed with the pI-*Cre*I plasmid; lane 3, induced cells transformed with the pAII17 vector; lane 4, induced cells transformed with the pI-*Cre*I plasmid; lane 5, soluble fraction from induced cells transformed with pI-*Cre*I; lane 6, 40–80% ammonium sulfate cut (precipitate); lane 7, pooled Mono S eluate; lane 8, molecular mass standards (in descending order: phosphorylase B, bovine serum albumin, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor and cytochrome *c*). The sizes of the standards are given in kDa; the arrow points to I-*Cre*I. (**B**) Mono S FPLC column elution profile. Each fraction was assayed for I-*Cre*I activity and protein content; the dashed line indicates the salt (NaCl) gradient. Fractions 25–30 were pooled, and an aliquot was applied to lane 7 of the gel in (A).

Size of native I-*Cre*I

The molecular weight of I-*Cre*I in solution was estimated by size-exclusion chromatography of the purified protein. The results of the sizing column are summarized in Figure 2A. The column fractions were analyzed by SDS–PAGE, and the 18 kDa band eluted primarily in fractions 37 and 38, which corresponded to a molecular mass of 36 kDa. Thus, this analysis indicated that I-*Cre*I is a dimer.

Table 1. Purification of I-*Cre*I

Fraction	Volume (ml)	Protein (mg)	Activity ($U \times 10^{-3}$)	Specific activity ($U/mg \times 10^{-3}$)	Fold purification	Yield (%)
Soluble	18	63.1	2898	46	1	100
40–80% $(NH_4)_2SO_4$	4	38.8	2020	52	1.13	69.7
Mono S	12	1.3	600	462	10	20.1

A unit of activity (U) was defined as the amount of enzyme needed to cleave 100 ng of pGEM23S.E DNA in 20 min at 37°C.

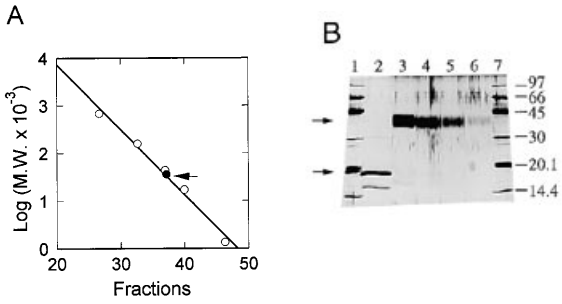

Figure 2. Size of native I-*Cre*I determined by size-exclusion chromatography and chemical cross-linking. (**A**) Size-exclusion chromatography of I-*Cre*I. Purified I-*Cre*I (filled circle) was eluted from a BioSep-SEC-S3000 column, which was calibrated with thyroglobulin (670 kDa), gamma globulin (158 kDa), ovalbumin (44 kDa), myoglobulin (17 kDa) and vitamin B_{12} (1.35 kDa); the markers are indicated as open circles. The arrow indicates the position of I-*Cre*I, which was determined by SDS–PAGE analysis of column fractions. (**B**) SDS–PAGE analysis of I-*Cre*I treated with the cross-linker dimethylsuberimidate at different protein concentrations. Equal volumes of the reactions were analyzed by SDS–PAGE on a 15% acrylamide gel, which was stained with silver nitrate. Lane 1, molecular mass standards (same as Fig. 1A); lane 2, untreated I-*Cre*I; lanes 3–6, I-*Cre*I treated with cross-linker at protein (monomer) concentrations of 1 μM (0.018 mg/ml), 0.5 μM (0.009 mg/ml), 0.25 μM (0.0045 mg/ml) and 0.125 μM (0.0023 mg/ml), respectively; lane 7, molecular mass standards. The arrows indicate the position of monomeric I-*Cre*I (18 kDa), and the major cross-linked product of ~38 kDa. The sizes of the markers are given to the right in kDa.

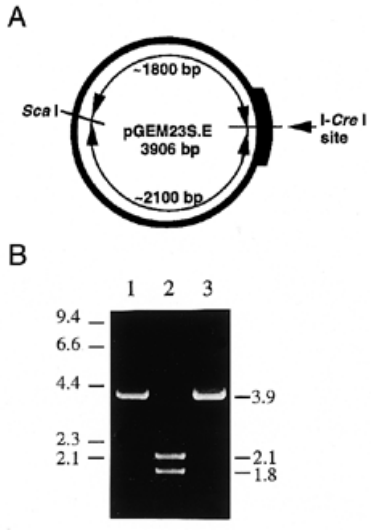

Figure 3. I-*Cre*I activity assay and Mg^{2+} dependence. (**A**) Organization of plasmid pGEM23S.E, which was used as the substrate. Construction of pGEM23S.E was described previously (9); it contains a portion of the fused exons of the 23S *rrn* gene of *C.reinhardtii* (thickened area) in the pGEM3zf(+) plasmid vector. The position of the I-*Cre*I cleavage site is indicated. The pGEM23S.E DNA was linearized with *Sca*I prior to digestion with I-*Cre*I, which produced DNA fragments of 2.1 and 1.8 kb, respectively. (**B**) Cleavage of the substrate DNA using the standard reaction solution: Mg^{2+} dependence. 500 ng of substrate was incubated with purified I-*Cre*I in the standard reaction solution at 37°C for 1 h (lane 2), or in the same solution minus Mg^{2+} (lane 3). The reactions were terminated with standard stop solution prior to analysis by agarose gel electrophoresis. Lane 1 contained untreated substrate DNA. The sizes (in kb) and positions of molecular mass markers, which were *Hin*dIII-digested λ DNA, are indicated to the left of the gel. The substrate and product DNAs are similarly indicated to the right.

The native size of I-*Cre*I was also assessed using chemical cross-linking with dimethylsuberimidate, which cross-links polypeptides covalently via lysine residues (23). I-*Cre*I was treated with dimethylsuberimidate, separated by SDS–PAGE, and then stained with silver. Figure 2B shows that the major protein in the treated samples migrated as a diffuse band with a relative molecular mass of 38 kDa; no band was observed remaining at the position of the monomer. Figure 2B (lanes 3–6) also shows that the cross-linking pattern did not change when the concentration of I-*Cre*I in the reaction was decreased from 1 μM, in a 2-fold series, to 0.125 μM (0.0023 mg/ml). The cross-linking pattern was also not affected by the addition of Mg^{2+} (data not shown). Thus, these data also indicate that I-*Cre*I is a dimer in solution. The minor band of 15 kDa in lane 2 is a polypeptide that contaminated some I-*Cre*I preparations, and was only visible with silver staining. It does not seem to be associated with I-*Cre*I, since its presence was variable. As to its fate in the presence of dimethylsuberimidate, a faint band migrating at 16 kDa in the treated samples (only visible in lane 3 as lanes 4–6 have too little protein) may be a modified form of this polypeptide; modification with dimethylsuberimidate is known to reduce slightly the mobility of polypeptides in SDS–PAGE (23).

I-*Cre*I endonuclease assays: optimization of activity

The plasmid pGEM23S.E, which contains the I-*Cre*I cleavage site (Fig. 3A), was used as the substrate; the plasmid is 3.9 kb. Although I-*Cre*I cuts supercoiled pGEM23S.E, the substrate DNA was linearized with *Sca*I (Fig. 3B, lane 1), so that cleavage with I-*Cre*I would generate two bands of 2.1 and 1.8 kb, respectively. Figure 3B (lane 2) shows the results of a typical cleavage assay; the figure also shows that no cleavage occurred when Mg^{2+} was omitted (lane 3).

At least four factors are known to affect the activity of homing endonucleases; temperature, pH, divalent cations and monovalent cations (29,30). In order to determine the optimal conditions for DNA cleavage by I-*Cre*I, these parameters were varied systematically, and the effects on cleavage assessed.

Mg^{2+} (or Mn^{2+}) was found to be essential for DNA cleavage by I-*Cre*I; in the absence of Mg^{2+}, no cleavage was detected (Fig. 3B, lane 3). As shown in Figure 4A, maximum activity was observed at 5–10 mM of $MgCl_2$, and higher concentrations of Mg^{2+} were somewhat inhibitory (e.g. 50 mM $MgCl_2$ produced only 50% relative cleavage). Other divalent cations were tested for their ability to support cleavage. In the presence of 5–50 mM $MnCl_2$, the substrate was cleaved by I-*Cre*I; however, no activity was observed in the presence of Ca^{2+} or Zn^{2+} (data not shown).

The effect of varying temperature over the range of 0–70°C on I-*Cre*I activity is shown in Figure 4B. Cleavage was more efficient at the higher temperatures, with complete cleavage of the substrate DNA occurring even at 70°C. At 50–70°C, most of the substrate was cleaved within the first 5 min, but at 37°C, it took 20 min for 100% of the substrate to be cleaved (data not shown). Interestingly, only 40% cleavage was observed at the physiological temperature range of 20–30°C, and at 0–10°C, no cleavage was detected even after 1 h.

To examine the effect of pH, cleavage assays were performed under standard conditions, except for varying the pH from 6 to 10.5. Figure 4C shows that I-*Cre*I exhibited a broad pH optimum,

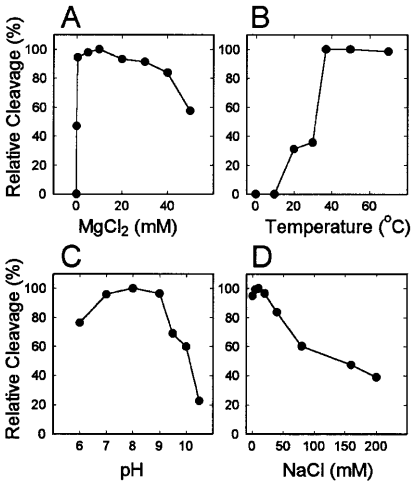

Figure 4. Determination of optimum conditions for DNA cleavage by I-*Cre*I. The graphs show the effects of (**A**) $MgCl_2$ concentration, (**B**) temperature, (**C**) pH and (**D**) NaCl concentration on DNA cleavage by purified I-*Cre*I. The other conditions were as described in Materials and Methods, except that diethanolamine was substituted for Tris to buffer the pH 9.5, 10 and 10.5 reactions in (C). In (A), (B) and (D), the ordinate values (% relative cleavage) were the same as the % substrate cleaved; in (C), however, the data was normalized using the reaction with the greatest fraction of substrate cleaved (which was >80%) set to 100% relative cleavage.

ranging from 7 to 9. The pH–activity profile also shows that the DNA-cleavage activity was reduced substantially above pH 9.5.

The effect of monovalent salts on I-*Cre*I activity was also examined. Figure 4D shows that cleavage was not stimulated by NaCl, and in fact was inhibited above 25 mM NaCl; similar results were obtained with KCl (data not shown). When the chloride salts were replaced with NaOAc and KOAc, inhibition occurred with concentrations >40 and 80 mM, respectively (data not shown).

Finally, the reducing agent, dithiothreitol, had no effect on I-*Cre*I activity.

Release of products following cleavage by I-*Cre*I

The standard stop solution used to terminate I-*Cre*I reactions contained SDS and proteinase K, in addition to EDTA to chelate the Mg^{2+} (Materials and Methods and Fig. 3). However, as shown in Figure 5 (lane 3), when SDS and proteinase K were omitted from the stop solution, the majority of the DNA migrated like uncleaved DNA (lane 1). This result indicates that most of the cleavage products are not released from the protein under the standard reaction conditions (the mobility of the DNA–protein complex is not shifted compared to naked DNA, because its expected size is only 1.5% larger). Since DNA cut with I-*Cre*I can be re-ligated with T4 ligase (9), the protein does not become covalently attached to the DNA ends. Moreover, Figure 5 (lane 2) shows that SDS alone was effective in promoting product release, but that proteinase K alone (lane 5) was relatively ineffective, at least under the conditions used here. Finally, a 70°C post-reaction heat-treatment (lane 4) was almost as effective as SDS in promoting release of the DNA products. With product release being so inefficient, one might expect that cleavage of DNA by I-*Cre*I would be predominantly stoichiometric, at least under the standard reaction conditions. Indeed, the amount of substrate cleaved was directly proportional to the amount of enzyme added (data not shown), and a molar excess of I-*Cre*I was needed to cleave all of the substrate DNA under standard conditions.

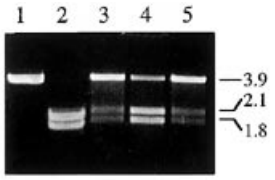

Figure 5. Efficient release of DNA products after digestion with I-*Cre*I requires SDS. A standard cleavage assay with purified I-*Cre*I was performed as described in Materials and Methods, except that at the end of the incubation period, equal volume aliquots were removed, and the reaction terminated with the different stop solutions described below (lanes 2–5). Lane 1, untreated substrate DNA; lane 2, the reaction was terminated with standard stop solution minus proteinase K (room temperature); lane 3, the reaction was terminated with standard stop solution minus proteinase K and SDS (room temperature); lane 4, the reaction was terminated with standard stop solution minus proteinase K and SDS, but heated to 70°C for 10 min; lane 5, the reaction was terminated with standard stop solution minus SDS, and then incubated for 10 min at room temperature. The samples were frozen immediately after the post-reaction treatments, subsequently thawed on ice, and then electrophoresed at room temperature. Sizes (in kb) of the substrate and products are indicated to the right of the gel.

Stability of I-*Cre*I

The stability of purified I-*Cre*I was examined by pre-incubating the protein under various conditions, followed by standard cleavage assays to estimate the fraction of activity remaining. Figure 6 shows that incubation of the protein in reaction buffer (Tris–HCl, pH 9.0) at the standard reaction temperature of 37°C resulted in a fairly rapid loss of I-*Cre*I activity, with only 20% remaining after 40 min. Adding Mg^{2+} to the pre-incubation buffer increased slightly the rate of loss, such that only 10% of the activity remained after 40 min. When pGEM23S.E (i.e. substrate) DNA was added (instead of Mg^{2+}), 90% of the starting activity remained even after 70 min of pre-incubation. Since the protein was in molar excess of the substrate, it is likely that most of the protection involved non-specific interactions. To test this, the protein was pre-incubated with plasmid DNAs that are not substrates for I-*Cre*I. Figure 6 shows that both pSR1 and pGEM3zf(+) effectively stabilized the activity; plasmid pSR1 was slightly more effective than either pGEM3zf(+) or pGEM23S.E, although the reason for this is not clear. In conclusion, these data show that purified I-*Cre*I is unstable in the absence of DNA, but can be stabilized by the addition of non-target DNA. Finally, it should be noted that the purified protein was unstable even at –70°C (in the absence of DNA and presence of 10% glycerol), losing activity at the rate of ~50% per year.

Binding of I-*Cre*I to target DNA in the absence of Mg^{2+}

A mobility shift assay was used to determine if I-*Cre*I would bind to its target DNA in the absence of Mg^{2+}. Figure 7 shows that the DNA shifts to a form with slower electrophoretic mobility in the presence of I-*Cre*I. The figure also shows that at the molar ratio of 4:1 (I-*Cre*I:DNA) essentially all of the 161 bp target DNA was shifted. Subsequent addition of a 15-fold excess of a 400 bp competitor DNA fragment (i.e. containing the same target site) did not displace the 161 bp DNA (data not shown), further indicating that the binding in the absence of Mg^{2+} was tight.

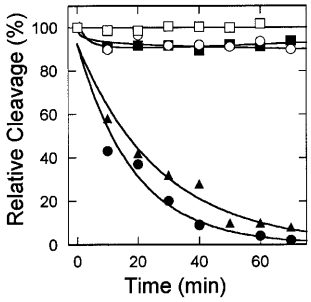

Figure 6. Stabilization of I-*Cre*I by DNA. Purified I-*Cre*I (10 pmol) was pre-incubated at 37°C in 100 µl of the solutions indicated below. At selected intervals, 10 µl aliquots were removed, and the fraction of activity remaining was assessed using the standard cleavage assay. The data were normalized using the 0 pre-incubation time point as 100% relative cleavage. The pre-incubation mixtures contained the standard buffer (Tris–HCl pH 9.0), plus one of the following: no further additions (▲), 10 mM $MgCl_2$ (●), 0.1 pmol of substrate DNA (pGEM23S.E) (■), 0.1 pmol pGEM3zf vector DNA (○) or 0.1 pmol pSR1 plasmid DNA (□).

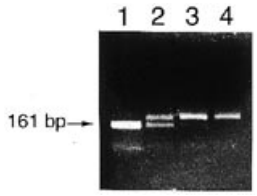

Figure 7. Mobility shift assay for I-*Cre*I binding to target DNA in the absence of Mg^{2+}. 150 ng (1.5 pmol) of the 161 bp PCR product containing the I-*Cre*I cleavage site was incubated with different amounts of I-*Cre*I in the absence of Mg^{2+}, and then analyzed by electrophoresis in a 2% agarose gel containing ethidium bromide. Lane 1, no I-*Cre*I was added; lane 2, with 108 ng (3 pmol) of I-*Cre*I; lane 3, with 216 ng (6 pmol) of I-*Cre*I; lane 4, with 324 ng (9 pmol) of I-*Cre*I. The arrow indicates the position of naked target DNA.

Hydroxyl radical protection and interference analysis

Hydroxyl radical cleavage (24) was used for high resolution footprinting and interference analysis of I-*Cre*I binding. Figure 8 (lanes 3–5 and 8,9) shows that I-*Cre*I strongly protected a single, continuous, 12 nt stretch of DNA backbone on each strand. There was also evidence of partial protection, or 'shadowing' (24), of one to three additional residues flanking the strongly protected regions. The strongly protected regions spanned the cleavage sites asymmetrically, and were slightly out of register, with a 2 nt shift towards the 3′ direction (Fig. 10). Figure 8 also shows the results of hydroxyl radical interference analysis; the same DNA fragments protected from cleavage by bound I-*Cre*I were diminished in the protein-bound fraction, and enriched in the unbound fraction (lanes 6,7 and 10,11). This result confirms the importance of these residues for binding.

Methylation protection and interference analysis

Methylation with DMS was used to examine the interaction of I-*Cre*I with the major and/or minor grooves of DNA. DMS methylates guanine at the N-7 position in the major groove, and adenine at the N-3 position in the minor groove (26); however, bound protein can protect these sites against methylation. Figure 9A shows the results of protection analysis. On the top strand, there

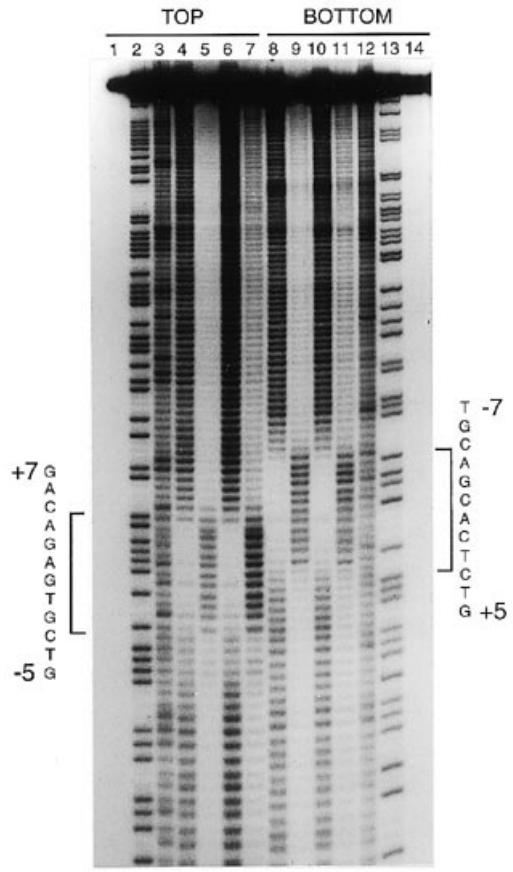

Figure 8. Hydroxyl radical protection and interference analysis of I-*Cre*I binding to target DNA. The reactions were carried out as described in Materials and Methods using singly-end-labeled target DNA, and the products analyzed on an 8% denaturing gel followed by autoradiography. Where appropriate, the bound and unbound DNA fractions were recovered from a native gel prior to their analysis on the denaturing gel. The strandedness (top or bottom) is indicated above the lanes. **Top strand**: lane 1, untreated DNA; lane 2, G+A reaction (as a marker); lane 3, reaction in the absence of protein; lane 4, protection reaction, bound DNA; lane 5, protection reaction, unbound DNA; lane 6, interference reaction, bound DNA; lane 7, interference reaction, unbound DNA. **Bottom strand**: lane 8, protection reaction, bound DNA; lane 9, protection reaction, unbound DNA; lane 10, interference reaction, bound DNA; lane 11, interference reaction, unbound DNA; lane 12, reaction in the absence of protein; lane 13, G+A reaction (marker); lane 14, untreated DNA. The affected regions and their sequence are indicated; they are numbered from the center of the cleavage site.

was strong protection of guanines at positions –5, +1, +3 and +7, and weak protection at –2. There was also weak protection of adenines +2 and +4; however, methylation of adenine at +6 was enhanced. On the bottom strand, the guanines at –6, –3 and +5 were strongly protected, while the adenine at –4 was weakly protected. Modification of the adenine at +9 was also strongly enhanced. These results are also summarized in Figure 10.

Methylation interference analysis was performed to assess the relative importance of the protected guanines and adenines (and/or others) for binding by I-*Cre*I; the results are shown in Figure 9B. On the top strand, the guanines at +1 and +3 were strongly required for binding, whereas those at –5 and –2 were required to some extent. Interestingly, the N-7 of the guanine at +7 was apparently not critical for binding, although it was strongly protected from methylation (Figs 9A and 10). Whether

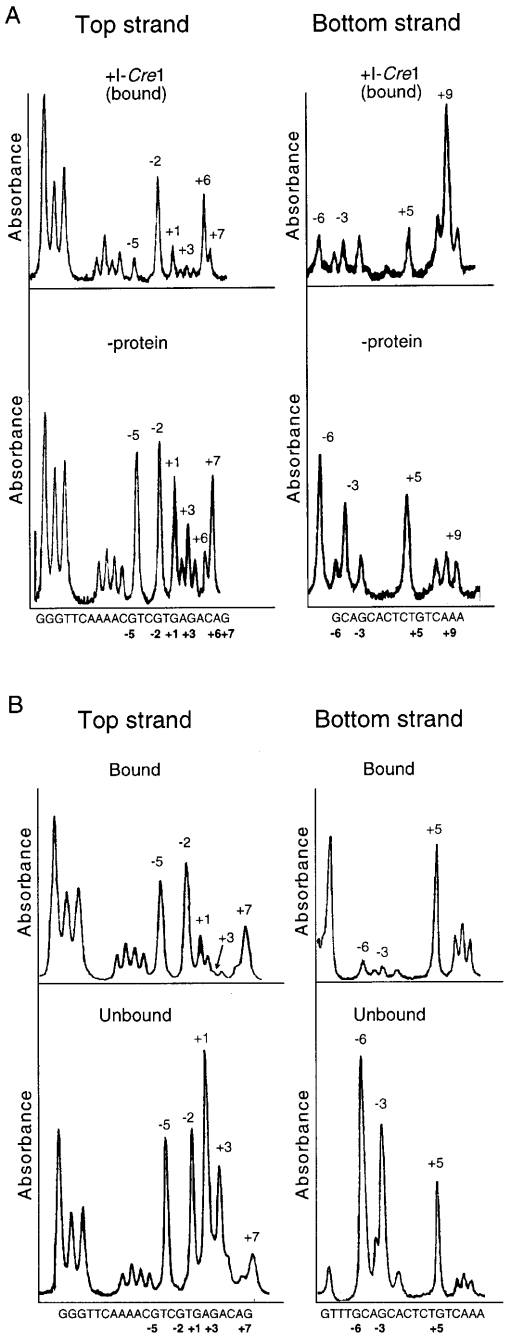

Figure 9. Methylation protection and interference analysis of I-*Cre*I binding to target DNA. The methylation reactions using singly-end-labeled (either top or bottom strand) target DNA were carried out as described in Materials and Methods. In some cases, bound and unbound DNA fractions were separated electrophoretically and recovered from a native gel prior to their cleavage and analysis on a denaturing gel. After autoradiography, the films were scanned, and the affected regions are shown. The DNA sequence is provided below each set of scans, and the positions of most (but not all) of the bases (Gs or As) that were either protected from methylation, or whose methylation interfered with binding, are indicated. (**A**) Methylation protection analysis. The lanes containing 'bound' DNA fractions for each strand were scanned (+I-*Cre*I), and compared to the corresponding region of the DNA methylated in the absence of protein (–protein). (**B**) Methylation interference analysis. The lanes containing bound (Bound) and unbound (Unbound) DNA fractions were scanned and presented for comparison.

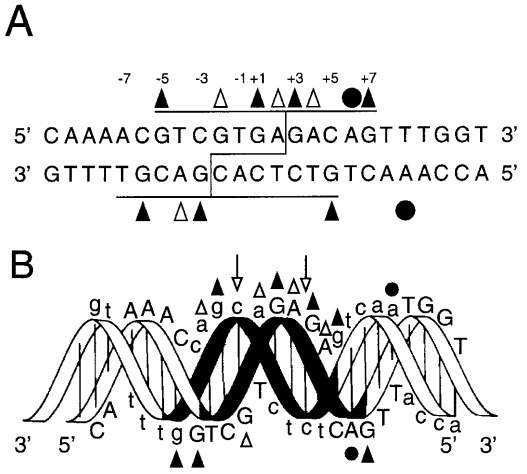

Figure 10. Summary of the hydroxyl radical and methylation protection analysis. The recognition sequence of I-*Cre*I is given, and the nucleotides are numbered from the center of the cleavage site. (**A**) Linear representation. The over- and underlined regions were protected from hydroxyl radical cleavage. The filled triangles indicate bases strongly protected from methylation, while the open triangles were weakly protected from methylation. The black dots indicate methylation enhancement. The vertical lines indicate the cleavage sites. (**B**) Helical representation. The bases on the top strand are in capital letters, and on the bottom strand in lower case letters. The regions of the backbone protected from hydroxyl radical cleavage are filled. The arrows indicate the cleavage sites, and the other designations are as in (A).

the adenines in the top strand were required for binding was difficult to assess because of low reactivity; however, the adenine at +4 may be at least partially necessary for binding. On the bottom strand, the guanines at –6 and –3 were strongly required, but the guanine at +5 was only weakly required for binding. The adenines at –4 and –1 on the bottom strand also appear to play a role in binding.

DISCUSSION

Intron-encoded endonucleases, such as I-*Cre*I, are critical for efficient intron homing, and therefore have played an important role in the distribution of introns (1–3). Homing endonucleases are also of interest because they recognize long stretches of DNA like transcription factors, yet they cut DNA like the well-known Type II restriction enzymes (28). Thus, they have the potential to teach us new things about DNA–protein interactions. I-*Cre*I, the subject of this study, is somewhat unusual among the LAGLI-DADG family of homing endonucleases, because it contains only a single copy of the signature motif, and its recognition site shows partial dyad symmetry (5).

To obtain more information on I-*Cre*I, we over-expressed the protein in *E.coli* using the inducible, stringently controlled T7 expression system (19). I-*Cre*I accumulated to 5–10% of the total cellular protein, and did not form an inclusion body, remaining in the soluble fraction. A simple, two step purification procedure was developed that gave I-*Cre*I preparations of >90% purity with 20% recovery.

The molecular mass of I-*Cre*I in solution was examined using size-exclusion chromatography and chemical cross-linking. The results obtained with the two techniques were congruent, and indicated that the protein is a homodimer. Previous work showed that I-*Sce*II is also a homodimer (29), but that I-*Sce*I and PI-*Sce*I

are monomers (30,31). Mutagenesis of these two-motif enzymes also showed that the LAGLI-DADG peptide is required for endonuclease activity; however, exactly how many copies of this motif are needed is not clear. While only P1 (i.e. motif #1) was needed by I-*Sce*II for endonuclease activity (11), both P1 and P2 (motif #2) were required by PI-*Sce*I (13). Since I-*Sce*II is a dimer (29), however, the P2 mutants presumably still had two copies of the wild-type P1 motif. Thus, those data suggest that two copies of LAGLI-DADG are needed for endonuclease activity (although not just any two copies will do, since an I-*Sce*II chimera containing two P2 motifs was inactive; 11). Our finding that I-*Cre*I is a dimer is consistent with that suggestion. In apparent contradiction, however, stands I-*Ceu*I, the only other single-LAGLI-DADG enzyme known besides I-*Cre*I (5). I-*Ceu*I is apparently a monomer in solution (14). It would be interesting to know if I-*Ceu*I remains monomeric throughout its DNA binding/reaction cycle.

The requirements and conditions for optimal I-*Cre*I activity were also investigated. Like other LAGLI-DADG endonucleases (4,14,29–31), Mg^{2+} (or Mn^{2+}) was absolutely required for DNA cleavage. Monovalent salts were not required, and in fact were mildly inhibitory, a characteristic shared by some (14,30,32), but not all, LAGLI-DADG endonucleases (4,31). I-*Cre*I activity was optimal at alkaline pH, which is also the case for other LAGLI-DADG enzymes (4,14,29–31). Unlike I-*Ceu*I (14), I-*Sce*I (30) and I-*Dmo*I (4), however, the efficiency of cleavage by I-*Cre*I decreased significantly above pH 9.5. Finally, I-*Cre*I exhibited a high temperature optimum (50–70°C), retaining full activity even at 70°C. This was an unexpected result for a DNA endonuclease that does not originate from a thermophilic organism. Only I-*Dmo*I, which comes from the hyperthermophilic archaeon, *Desulfurococcus mobilis*, has been shown to be fully active at 70°C (4). Interestingly, I-*Ceu*I was reported to be more active at 50°C than at lower temperatures, although higher temperatures were apparently not tested (14). The adaptive value, if any, of the high temperature optima for I-*Cre*I and I-*Ceu*I is not clear. However, based on these analyses, we can conclude that I-*Cre*I produced by *E.coli* would be active under the conditions typically found in *C.reinhardtii* chloroplasts, although its activity might not be optimal. To what extent the enzyme produced in *C.reinhardtii* chloroplasts might differ (due to chloroplast-specific post-translational modifications) from the enzyme produced in *E.coli* is not known, since only the latter has been characterized.

Given some of the aforementioned characteristics of I-*Cre*I, it was surprising that the purified enzyme proved to be unstable, even at –70°C. The rate of loss of I-*Cre*I activity at the standard reaction temperature (37°C) was in the order of 50% in 18 min. Although this contrasts with the extreme instability of I-*Sce*I (30), which lost 50% of its activity in 1 min at 37°C, the instability of I-*Cre*I is much greater than I-*Ceu*I, which required 3 h to lose 50% of its activity (14). We also observed that I-*Cre*I could be stabilized by adding DNA, and that non-substrate DNAs were as effective as the substrate. Interestingly, I-*Sce*I was also found to be stabilized by DNA; however, only DNAs that contained the recognition site, or a close variant, were effective (30). The fact that I-*Cre*I was stabilized by DNAs that lacked the recognition site suggests that the protein interacts with DNA non-specifically as well as specifically.

A particularly interesting characteristic of I-*Cre*I concerns the fact that efficient release of the cleavage products from the enzyme required a strong protein denaturant (SDS). This result indicates that the enzyme forms very stable complexes with the DNA products. Two other LAGLI-DADG enzymes, PI-*Sce*I and I-*Sce*I, have been shown to bind very tightly to the 3′ exon product (13,32); however, I-*Cre*I seems to hold on tightly to both products, suggesting that the interaction of I-*Cre*I with its target is different from the yeast enzymes. This suggestion is supported by the footprinting data (discussed below). Not surprisingly, the DNA cleaving activity of I-*Cre*I is inefficient, in terms of enzyme turnover, as is the case for I-*Sce*I (32). *In vivo*, however, a poor turnover rate for I-*Cre*I may not be much of a problem, since the amount of the putative substrate, intronless 23S DNA, should be very small. Moreover, it is possible that, *in vivo*, the formation of a stable complex of I-*Cre*I with its DNA products could facilitate the subsequent recombination events (7,8).

The mobility shift assay demonstrated that I-*Cre*I binds to its target DNA in the absence of Mg^{2+}. This allowed us to study DNA binding in the absence of cleavage. Hydroxyl radical protection analysis indicated that I-*Cre*I binds strongly to the DNA backbone for a continuous stretch of 12 nt on each strand, centered slightly asymmetrically (1 bp shift) around the cleavage site. Hydroxyl radical interference analysis confirmed that the strongly protected regions are required for binding. These stretches of DNA on each strand are also out of register with each other, with a 2 nt shift in the 3′ direction; a result which is characteristic of minor groove interactions (33). Thus, it can be stated that the overall span of target DNA required for tight binding by I-*Cre*I is at least 14 bp. This number is smaller by several nucleotides than the recognition sequence determined by digesting sequence ladders of the target (9,15). This difference might be due in part to ‘end effects’ on the cleavage activity of I-*Cre*I, since the sequence ladder approach involves digesting a population of nested molecules, including those with the recognition sequence at the very end of the DNA. Finally, we note that these footprinting data are consistent with the observation that I-*Cre*I holds on tightly to both DNA products.

The pattern of hydroxyl radical protection obtained with I-*Cre*I is quite different from the hydroxyl radical protection pattern of PI-*Sce*I, which binds strongly mainly to the 3′ exon (13). PI-*Sce*I is the only other LAGLI-DADG endonuclease for which there are high-resolution, hydroxyl radical protection patterns. However, DNase I footprinting of I-*Dmo*I has been performed, and the results indicated strong binding to both exons (34).

The methylation protection analysis demonstrated that I-*Cre*I protected guanines throughout the region of DNA defined by hydroxyl radical footprinting (Fig. 10). This result indicates that I-*Cre*I binds strongly in the major groove, and for more than one turn of the helix. Hypermethylation of certain adenines in and near the right half of the protected region was also observed, suggesting that some distortion of the DNA structure accompanies I-*Cre*I binding.

The methylation interference analysis revealed that only some of the protected purines were strongly required for binding (Fig. 9B), suggesting that I-*Cre*I might tolerate, at the least, some single-site changes in the target sequence. An extensive analysis of the effects of mutagenizing the target site on cleavage by I-*Cre*I has not been published. However, Durrenberger and Rochaix (15) did digest several variants of the target site that occur naturally in different rRNAs. Our results are generally consistent with theirs, except for the fact that I-*Cre*I was able to cleave a *Paramecium* mitochondrial rRNA gene, which has a T-A pair at the –3 position instead of C-G (15). The methylation interference data indicate

that the G at –3 on the bottom strand should be strongly required for binding (Fig. 9B). These comparisons are complicated, however, by the different conditions used for cleavage (e.g. inclusion of Mg^{2+}) versus binding (no Mg^{2+}). It will be useful to carry out quantitative binding assays with target-site mutants compared to wild-type DNA.

In summary, we have reported the over-expression, purification and initial characterization of the homing endonuclease, I-*Cre*I. We also characterized its binding to target DNA using high resolution footprinting and interference analyses, making it the first single-LAGLI-DADG endonuclease to be studied in this way. Finally, these results provided the basis for further work on this unusual site-specific endonuclease, including X-ray crystallographic studies (35), which were published after the revision of this manuscript.

ACKNOWLEDGEMENTS

We thank Dr William Jack (New England BioLabs) for the expression vector, Dr Jung-Keun Suh for helpful advice and Dr Jon Robertus for use of lab facilities. This work was supported by grants from the R.A.Welch Foundation (F-1164), the USDA (NRICGP 92-37301-7682), and the Texas Advanced Technology Program (ATP 003658-326). D.L.H. would also like to thank A.S.H., A.J.H. and J.P.H. for their patience and support during manuscript preparation.

REFERENCES

1 Dujon, B. (1989) *Gene* **82**, 91–114.
2 Belfort, M. (1990) *Annu. Rev. Genet.* **24**, 363–385.
3 Lambowitz, A.M. and Belfort, M. (1993) *Annu. Rev. Biochem.* **62**, 587–622.
4 Dalgaard, J.Z., Garrett, R.A. and Belfort, M. (1994) *J. Biol. Chem.* **269**, 28885–28892.
5 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 ed. Cold Spring Harbor Laboratory Press, New York, pp. 111–143.
6 Mueller, J.E., Clyman, J., Huang, Y.-J., Parker, M.M. and Belfort, M. (1996) *Genes Dev.* **10**, 351–364.
7 Durrenberger, F., Thompson, A.J., Herrin, D.L. and Rochaix, J.-D. (1996) *Nucleic Acids Res.* **24**, 3323–3331.
8 Durrenberger, F. and Rochaix, J.-D. (1991) *EMBO J.* **10**, 3495–3501.
9 Thompson, A.J., Yuan, X., Kudlicki, W. and Herrin, D.L. (1992) *Gene* **119**, 247–251.
10 Michel, F., Jacquier, A. and Dujon, B. (1982) *Biochimie* **64**, 867–881.
11 Henke, R.M., Butow, R.A. and Perlman, P.S. (1995) *EMBO J.* **14**, 5094–5099.
12 Hodges, R.A., Perler, F.B., Noren, C.J. and Jack, W.E. (1992) *Nucleic Acids Res.* **20**, 6153–6157.
13 Gimble, F.S. and Stephens, B.W. (1995) *J. Biol. Chem.* **270**, 5849–5856.
14 Marshall, P., Davis, T.B. and Lemieux, C. (1994) *Eur. J. Biochem.* **220**, 855–859.
15 Durrenberger, F. and Rochaix, J.-D. (1993) *Mol. Gen. Genet.* **236**, 409–414.
16 Pognonec, P., Kato, H., Sumimoto, H., Kretzschmar, M. and Roeder, R.G. (1991) *Nucleic Acids Res.* **19**, 6650.
17 Rochaix, J.-D., Rahire, M. and Michel, F. (1985) *Nucleic Acids Res.* **13**, 975–984.
18 Thompson, A.J. and Herrin, D.L. (1991) *Nucleic Acids Res.* **19**, 6611–6618.
19 Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) *Methods Enzymol.* **185**, 60–89.
20 Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248–254.
21 Laemmli, U.K. (1970) *Nature* **227**, 680–685.
22 Herrin, D.L., Plumley, F.G., Ikeuchi, M., Michaels, A.S. and Schmidt, G.W. (1987) *Arch. Biochem. Biophys.* **254**, 397–408.
23 Davies, G.E. and Stark, G.R. (1970) *Proc. Natl. Acad. Sci. USA* **66**, 651–656.
24 Tullius, T.D. and Dombroski, B.A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5469–5473.
25 Bencini, D.A., O'Donovan, G.A. and Wild, J.R. (1984) *BioTechniques* **2**, 4–5.
26 Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
27 Wang, J. (1995) Master's thesis, The University of Texas at Austin.
28 Roberts, R.J. and Halford, S.E. (1993) In Linn, S.M., Lloyd, R.S. and Roberts, R.J. (eds), *Nucleases*, 2nd ed. Cold Spring Harbor Laboratory Press, New York, pp. 35–88.
29 Wernette, C.M., Saldahna, R., Perlman, P.S. and Butow, R.A. (1990) *J. Biol. Chem.* **265** 18976–18982.
30 Monteilhet, C., Perrin, A., Thierry, A., Colleaux, L. and Dujon, B. (1990) *Nucleic Acids Res.* **18**, 1407–1413.
31 Gimble, F.S. and Thorner, J. (1993) *J. Biol. Chem.* **268**, 21844–21853.
32 Perrin, A., Buckle, M. and Dujon, B. (1993) *EMBO J.* **12**, 2939–2947.
33 Oakley, M.G. and Dervan, P.B. (1990) *Science* **248**, 847–850.
34 Aagaard, C., Awayez, M.J. and Garrett, R.A. (1997) *Nucleic Acids Res.* **25**, 1523–1530.
35 Heath, P.J., Stephens, K.M., Monnat, R.J. and Stoddard, B.L. *Nature Struct. Biol.* **4**, 468–476.