Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 883–887, February 1994
Biochemistry

Chimeric restriction endonuclease

(*Flavobacterium okeanokoites/Escherichia coli*/hybrid restriction endonuclease/protein engineering/recognition and cleavage domains)

YANG-GYUN KIM AND SRINIVASAN CHANDRASEGARAN*

Department of Environmental Health Sciences, School of Hygiene and Public Health, The Johns Hopkins University, 615 North Wolfe Street, Baltimore, MD 21205-2179

Communicated by Hamilton O. Smith, October 12, 1993 (received for review July 23, 1993)

ABSTRACT ***Fok* I restriction endonuclease recognizes the nonpalindromic pentadeoxyribonucleotide 5′-GGATG-3′·5′-CATCC-3′ in duplex DNA and cleaves 9 and 13 nt away from the recognition site. Recently, we reported the presence of two distinct and separable domains within this enzyme: one for the sequence-specific recognition of DNA (the DNA-binding domain) and the other for the endonuclease activity (the cleavage domain). Here, we report the construction of a chimeric restriction endonuclease by linking the *Drosophila* Ultrabithorax homeodomain to the cleavage domain (F_N) of *Fok* I restriction endonuclease. The hybrid enzyme, Ubx–F_N, was purified, and its cleavage properties were characterized. The hybrid enzyme shows the same DNA sequence-binding preference as that of Ubx; as expected, it cleaves the DNA away from the recognition site. On the 5′-TTAATGGTT-3′ strand the hybrid enzyme cleaves 3 nt away from the recognition site, whereas it cuts the complementary 5′-AACCATTAA-3′ strand 8, 9, or 10 nt away from the binding site. Similarly engineered hybrid enzymes could be valuable tools in physical mapping and sequencing of large eukaryotic genomes.**

We have undertaken a detailed study of the *Fok* I restriction-modification (R-M) system from *Flavobacterium okeanokoites*. *Fok* I is a member of the type IIS class (1). Unlike other type II restriction enzymes, which have the cleavage site within or adjacent to their recognition sites, the type IIS enzymes cleave double-stranded DNA at precise distances from their recognition site (2). The *Fok* I endonuclease recognizes the nonpalindromic pentadeoxyribonucleotide 5′-GGATG-3′·5′-CATCC-3′ in duplex DNA and cleaves 9 and 13 nt away from the recognition site (2). This implies the presence of two separate protein domains: one for the sequence-specific recognition of DNA and the other for the endonuclease activity. Once the DNA-binding domain is anchored at the recognition site, a signal is transmitted to the endonuclease domain, probably through allosteric interactions, and the cleavage occurs.

Our studies on proteolytic fragments of *Fok* I endonuclease have defined an N-terminal DNA-binding domain and a C-terminal domain with nonspecific DNA-cleavage activity (3). These results have been confirmed by the study of C-terminal deletion mutants of *Fok* I endonuclease (4). We have also shown that introduction of additional amino acid residues between the recognition and cleavage domains of *Fok* I can alter the spacing between the recognition site and the cleavage site of the DNA substrate (5). Our studies indicate that the two domains are connected by a linker region, which appears to be amenable for repositioning of the DNA-sequence recognition domain with respect to the catalytic domain.

The modular structure of *Fok* I endonuclease suggests that it may be feasible to construct hybrid endonucleases with distinct sequence-specificities by linking other DNA-binding proteins, especially those that recognize sequences >6 bp to the cleavage domain of *Fok* I endonuclease. Most naturally occurring restriction enzymes recognize sequences that are 4–6 bp long. Although these enzymes are very useful in manipulating recombinant DNA, they are not suitable for handling large DNA segments. In many instances, it is preferable to have fewer but longer DNA strands, especially during genome mapping. So far, only ≈10 restriction endonucleases with recognition sequences >6 bp (rare cutters) have been identified (New England Biolabs catalog). R-M systems appear to have a single biological function—namely, to protect cells from infection by phage that would otherwise destroy them. The phage genomes are usually small. It stands to reason, then, that bacteria select for R-M systems with small recognition sites (4–6 bp) because these sites occur more frequently in the phages. Several independent strategies have been devised in the attempt to develop a facile method for the production and manipulation of large DNA segments (6–13). These strategies include RecA-assisted restriction endonuclease (RARE) cleavage (6), site-specific cleavage mediated by triple-helix formation (7), site-specific cleavage of duplex DNA by chemically cross-linking λ repressor to staphylococcal nuclease (8), and Achilles' heal cleavage (11–13).

Our approach to generate the rare cutters (restriction enzymes that recognize sequences >6 bp) is rather straightforward. Because DNA-binding proteins that recognize sequences >6 bp exist in nature (e.g., zinc finger motifs, homeodomain motifs, POU domain motifs, and DNA-binding protein domains of λ, *lac* repressors, *cro*, etc.), we reasoned that the obvious and logical next step is to fuse these DNA-binding domains to the C-terminal cleavage domain of *Fok* I and screen for active endonucleases with altered substrate specificity.

We chose the *Ubx* homeodomain of *Drosophila* as a representative of the helix–turn–helix motif proteins for fusion with the C-terminal cleavage domain of *Fok* I. The homeodomain is a 61-codon region of similarity in the sequences of several *Drosophila* genes that are important in embryonic development (14). These sequences are highly conserved in many other genes from higher and lower eukaryotes. The homeodomain is sufficient for sequence-specific DNA-binding activity, even without the flanking polypeptide sequence. The homeodomains contain a helix–turn–helix motif that contacts the DNA in a way similar to the prokaryotic repressors. The three-dimensional structures of the Antennapedia (15) and Engrailed (16, 17) homeodomains have been determined. They confirm the presence of a helix–turn–helix motif within these proteins. Furthermore,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: *fokIR*, gene coding for *Fok* I restriction endonuclease; F_N, cleavage domain of *Fok* I; *Ubx–F_N*, gene coding for the hybrid protein Ubx–F_N constructed by linking the *Drosophila* Ultrabithorax (Ubx) homeodomain (Ubx) to the cleavage domain (F_N) of *Fok* I; R-M, restriction-modification.

*To whom reprint requests should be addressed.

the DNA-sequence preferences of the homeodomains appear to be altered by amino acid changes in the recognition helix of the helix–turn–helix motifs. This result adds further support to the functional significance of the structural homology between the homeodomains. The Ultrabithorax homeodomain recognizes the 9-bp consensus DNA sites 5′-TTAAT(G/T)(G/A)CC-3′ (18, 19).

We describe here the construction of a chimeric restriction endonuclease by linking the Ubx homeodomain to the cleavage domain of *Fok* I. We also describe the purification and characterization of this hybrid enzyme.

MATERIALS AND METHODS

The complete nucleotide sequence of the *Fok* I R-M system has been published (20, 21). Experimental protocols for PCR have been described elsewhere (22). The procedures for cell growth and purification of proteins using His-bind resin (23) is as outlined in the Novagen pET system manual. The protocol for SDS/PAGE is as described by Laemmli (24).

Preparation of pUC13-Derived Substrates. pUC13-derived DNA substrates were prepared by blunt-end ligation of *Sma* I-cleaved pUC13 plasmid with 10-fold excess of a 31-bp insert (5′-GCGCTTAATGGTTTTTGCTTTTTGCTTTTTT-3′·5′-AAAAAAGCAAAAAGCAAAAACCATTAA-GCGC-3′) containing a known *Ubx* site, 5′-TTAATGGTT-3′. The inserted *Ubx* site was chosen on the basis of the original 7-bp consensus sequence 5′-TTAATGG-3′ (18). This sequence has been subsequently revised to a 9-bp consensus sequence 5′-TTAAT(G/T)(G/A)CC-3′ (19). Several clones were picked, and their plasmid DNAs were analyzed for the presence of 31-bp inserts. Clones containing pUC13-1, pUC13-2, or pUC13-3, each with one, two, and three inserts, respectively, were identified. Their DNA sequences were confirmed by Sanger's dideoxynucleotide chain-termination sequencing method (25).

Preparation of DNA Substrates with a Single *Ubx* Site. The polylinker region of pUC13-1, which has a single 31-bp insert was excised by using *Eco*RI/*Hin*dIII and gel-purified. Individual strands of this 82-bp substrate were radiolabeled by using [^{32}P]dATP or [^{32}P]dCTP and filling in the sticky ends of the fragment with Klenow enzyme. The products were purified from agarose gel by using DEAE membranes (NA 45 from Schleicher & Schuell). The DNA was eluted with high-salt elution buffer (TE buffer: 50 mM Tris·HCl, pH 7.6/1 mM EDTA/0.2 M NaCl/0.1% SDS), ethanol-precipitated, and resuspended in TE buffer.

Purification of Ubx-F_N Endonuclease. *Escherichia coli* BL21 (DE 3) [pET-15b *Ubx–F_N*] was grown in 1 liter of LB broth containing 100 μg of ampicillin per ml at 37°C to an OD_{600} unit of 0.5 and then induced for 2–3 hr with 1 mM isopropyl β-D-thiogalactoside. The cells were harvested by centrifugation and then resuspended in 100 ml of binding buffer (5 mM imidazole/0.5 M NaCl/20 mM Tris·HCl, pH 7.9). Cells were disrupted at maximum intensity on a Branson sonicator for 1 hr at 4°C. Accurate determination of the hybrid enzyme in the crude extract was not feasible because of the low yield and the contaminating nonspecific nucleases of the cell. Therefore, only a qualitative description of the purification procedure of the hybrid enzyme is outlined here. The sonicated cells were centrifuged at 15,000 × *g* for 1 hr at 4°C. The supernatant was filtered through a 0.45-μm filter and then loaded onto a His-bind column (2 ml). After being washed with 10 vol of binding buffer and 10 vol of wash buffer (60 mM imidazole/0.5 M NaCl/20 mM Tris·HCl, pH 7.9), the column was eluted with 5 ml of eluate buffer (1 M imidazole/0.5 M NaCl/20 mM Tris·HCl, pH 7.9). A PD-10 column (Pharmacia) was used to change the buffer of the eluted fractions to buffer A [10 mM Tris phosphate, pH 8.0/7 mM 2-mercaptoethanol/1 mM EDTA/10% (vol/vol) glycerol]/0.2 M NaCl. This solution was then loaded onto a phosphocellulose column (Whatman) (4 ml). The column was washed with 16 ml of buffer A/0.4 M NaCl and eluted with 12 ml of buffer A/1 M NaCl. This eluate was then dialyzed against 500 ml of buffer A/0.2 M NaCl overnight at 4°C and loaded onto a SP-Sepharose (Pharmacia) column (3 ml). After being washed with 15 ml of buffer A/0.4 M NaCl, the column was eluted with a 40-ml total gradient of 0.4–1 M NaCl in buffer A. Fractions were analyzed by electrophoresis on 0.1% SDS/15% polyacrylamide gels. Proteins were stained with Coomassie blue (24). Identity of the hybrid protein was confirmed by probing the immunoblot with rabbit antiserum raised against *Fok* I endonuclease. The restriction endonuclease activity of the fractions was assayed by using pUC13 as substrate. The hybrid enzyme Ubx–F_N was purified to near-homogeneity with this procedure. In a control experiment, the crude extract of a 1-liter culture of *E. coli* BL21 (DE 3) [pET-15b] cells, which does not contain the *Ubx–F_N* insert, was subjected to exactly identical purification procedure as described above. None of the fractions from the SP-Sepharose column exhibited any sequence-specific endonuclease activity or nonspecific nuclease activity. This result suggests that the nonspecific nucleases of the cell do not copurify with the hybrid enzyme.

RESULTS AND DISCUSSION

Construction of the Clone Producing the Hybrid Enzyme Ubx–F_N Using PCR. The homeodomain of Ubx, a 61-amino acid protein sequence encoded by the homeobox of *Ubx* (14), is a sequence-specific DNA-binding domain with a structure related to helix–turn–helix motifs found in bacterial DNA-binding proteins (15–17). The Ubx homeodomain recognizes the 9-bp consensus DNA sites 5′-TTAAT(G/T)(G/A)CC-3′ (18, 19). We have used the PCR technique to link the Ubx homeodomain to the F_N of *Fok* I and to express the Ubx–F_N enzyme in *Escherichia coli*. A schematic representation of the engineered Ubx–F_N hybrid protein is shown in Fig. 1. The oligonucleotide primers used to construct the hybrid gene is shown in Fig. 2*A*. The *Ubx* homeodomain template was provided by Philip Beachy (Department of Molecular Biology and Genetics, Johns Hopkins School of Medicine). Construction of the clone expressing the hybrid protein was done as

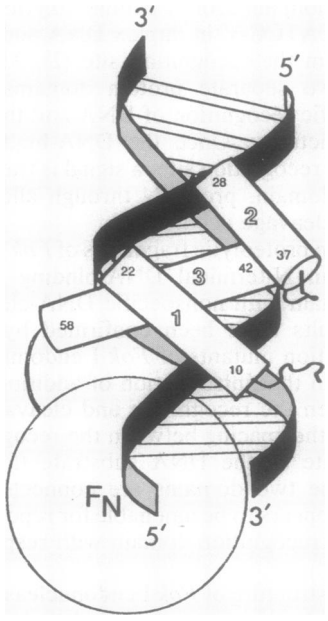

FIG. 1. Diagram showing orientation of the Ubx homeodomain with respect to the *Fok* I nuclease domain (F_N) in relation to the DNA substrate. The crystal structure of an Engrailed homeodomain–DNA complex was reported by Kissinger *et al.* (16).

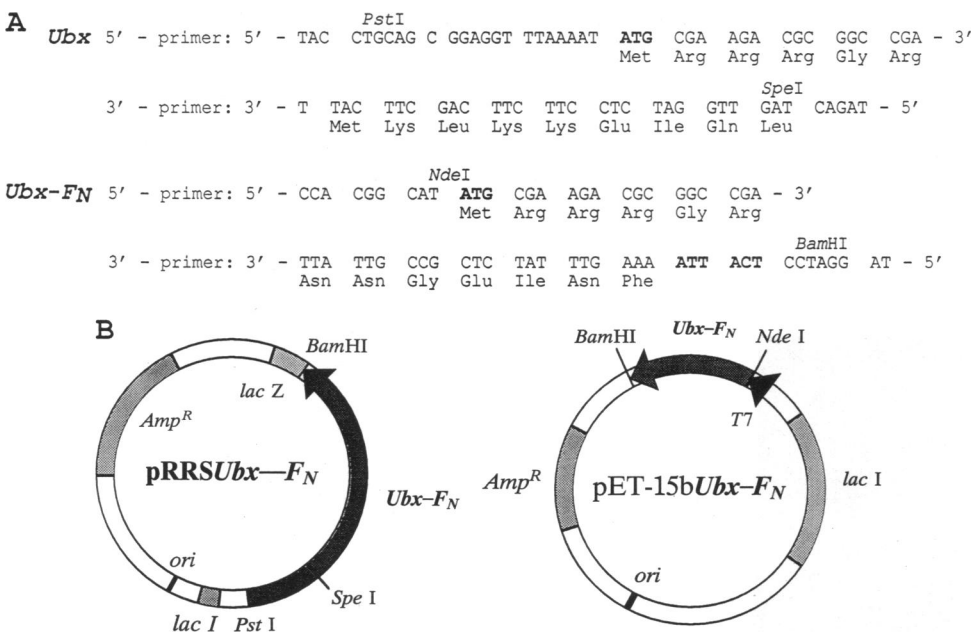

FIG. 2. Construction of expression vector of hybrid enzyme, Ubx–F_N. (*A*) Sequences of 5′ and 3′ primers used to construct the hybrid gene, *Ubx–F_N*. The *Ubx* primers are flanked by *Pst* I and *Spe* I sites. The *Ubx–F_N* primers are flanked by *Nde* I and *Bam*HI sites. Start and stop codons are shown in boldface letters. (*B*) Structure of plasmids pRRS*Ubx-F_N* and pET-15b*Ubx–F_N*. The PCR-modified *Ubx* homeobox was ligated into the *Pst* I/*Spe* I-cleaved vector pRR*SfokIR* containing a single *Spe* I site (5) to generate pRRS*Ubx–F_N*. The PCR-generated fragment using *Ubx–F_N* primers was inserted at the *Bam*HI/*Nde* I sites of pET-15b (26) to form pET-15b*Ubx–F_N*.

follows: (*i*) the PCR-generated *Ubx* homeobox was digested with *Pst* I/*Spe* I and gel-purified. This fragment was then ligated into the *Pst* I/*Spe* I-cleaved vector pRR*SfokIR* containing a single *Spe* I site (5) to replace the DNA segment coding for the *Fok* I DNA-binding domain and hence, form the *Ubx–F_N* hybrid gene (Fig. 2*B*). This construct links the 61-amino acid Ubx homeodomain to the C-terminal 196-amino acids of *Fok* I that constitute the *Fok* I cleavage domain. After transfection of competent RR1 cells with the ligation mix, several clones were identified by restriction analysis, and their DNA sequences were confirmed by the dideoxynucleotide chain-termination method of Sanger *et al.* (25). (*ii*) The hybrid gene was PCR-amplified using the *Ubx–F_N* primers (shown in Fig. 2*A*). The PCR-generated DNA was digested with *Nde* I/*Bam*HI and gel-purified. This fragment was then ligated into the *Nde* I/*Bam*HI-cleaved pET-15b vector (26), which contains a T7 promoter for expression of the hybrid gene. This construct will tag the hybrid protein with six consecutive histidine residues at the N terminus. These residues serve as the affinity tag for purification of this protein by metal-chelation chromatography (23) with Novagen's His-bind resin. This histidine tag can be subsequently removed by thrombin (Novagen pET system manual). Competent BL21 (DE 3) cells (26) were transformed with the ligation mix, and several clones containing the recombinant DNA (Fig. 2*B*) were identified by restriction analysis. These clones appear to tolerate the hybrid protein, even though they do not contain a modification gene to protect them from chromosomal cleavage. This result could be attributed to the tight control exerted over the expression of the hybrid enzyme by the pET system, as well as to the efficient repair of the DNA damage by the DNA ligase within these cells. The host, BL21 (DE 3), contains a chromosomal copy of T7 RNA polymerase gene under *lacUV5* control, and the expression is induced by addition of isopropyl β-D-thiogalactoside. After induction with 1 mM isopropyl β-D-thiogalactoside, the hybrid enzyme was purified to near-homogeneity using His-bind resin, phosphocellulose, and S-Sepharose chromatography. The SDS/PAGE profile of the purified hybrid enzyme is shown in Fig. 3. Its size is ≈32 kDa, which agrees well with that predicted for the fusion protein. Identity of the hybrid protein was further confirmed by probing the immunoblot with rabbit antiserum raised against *Fok* I endonuclease (data not shown).

Analysis of the DNA-Sequence Preference of the Ubx–F_N Hybrid Enzyme. The linearized pUC13 and its derivatives used as substrates to characterize Ubx–F_N are shown in Fig. 4*A*. The inserted *Ubx* site was chosen based on the original 7-bp consensus sequence 5′-TTAATGG-3′ (18); this has been subsequently revised to a 9-bp consensus sequence, 5′-TTAAT(G/T)(G/A)CC-3′ (19). To determine whether there were any *Ubx* sites within pUC13, the *Sca* I-linearized plasmid was digested with Ubx–F_N in the presence of tRNA (Fig. 4*B*, lane 3). In these reactions the molar ratio of DNA was in large excess (≈10-fold) compared with the protein. The reaction conditions were optimized to give a single double-stranded cleavage per substrate molecule. Reduction of $MgCl_2$ concentration to 2 mM was critical for lowering the nonspecific nuclease activity of the hybrid enzyme Ubx–F_N. Although tRNA was routinely added to the reaction mix-

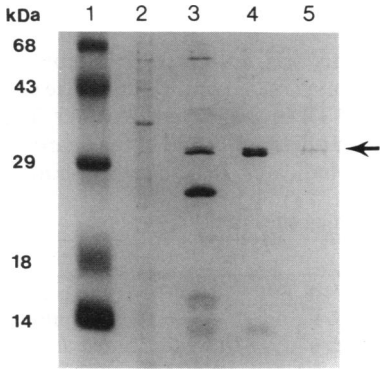

FIG. 3. SDS/PAGE profiles at each step in the purification of Ubx–F_N hybrid enzyme (arrow). Lanes: 1, protein standards; 2, crude extract from induced cells; 3, His-bind resin column; 4, phosphocellulose column; and 5, SP-Sepharose column.

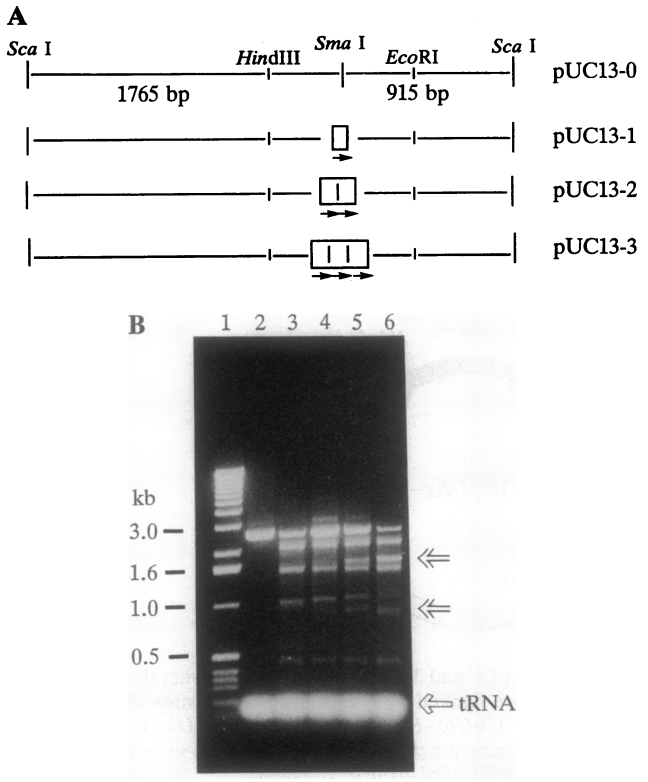

FIG. 4. Characterization of Ubx–F_N hybrid protein using the linearized pUC13 DNA substrates containing *Ubx* site(s). (*A*) pUC13-derived DNA substrates. □, 31-bp insert containing the *Ubx* site 5′-TTAATGGTT-3′. The number of tandem repeats of the 31-bp insert in these substrates are shown as pUC13-*X*, where *X* varies from 0 to 1, 2, and 3. Orientation of the *Ubx* site(s) are indicated by arrows. (*B*) DNA substrates (1 μg) were partially digested with the hybrid enzyme (2 ng) in buffer containing 20 mM Tris·HCl (pH 7.6), 75 mM KCl, 1 mM dithiothreitol, bovine serum albumin at 50 μg/ml, 10% glycerol, tRNA at 100 mg/ml, and 2 mM $MgCl_2$ at 31°C for 5 hr. Products were analyzed by 1% agarose gel electrophoresis. The substrate was present in large excess compared with the Ubx–F_N hybrid protein (molar ratio was ≈10:1). The reaction condition was optimized to yield a single double-stranded cleavage per substrate molecule. The nonspecific nuclease activity of the hybrid enzyme increases with increasing $MgCl_2$ concentration. Addition of tRNA reduces the nonspecific nuclease activity in the crude extract, and tRNA was routinely added to reaction mixtures. It appears to have no effect on cleavage properties of the hybrid enzyme (data not shown). The two fragments, ≈1.8 kb and ≈0.9 kb, respectively, resulting from binding of the hybrid enzyme at the inserted *Ubx* site of pUC13 and cleaving near this site, are indicated by arrows. Lanes: 1, 1-kb ladder; 2, pUC13-0; 3, pUC13-0 digested with Ubx–F_N; 4, pUC13-1 digested with Ubx–F_N; 5, pUC13-2 digested with Ubx–F_N; and 6, pUC13-3 digested with Ubx–F_N.

tures, it apparently has no effect on the properties of the hybrid enzyme (data not shown). Under these conditions, the linearized pUC13 DNA is cleaved into four fragments. The appearance of four distinct bands in the agarose-gel electrophoretic profile indicates that Ubx–F_N binds DNA in a

Table 1. Ubx-binding sites in pUC13

Sequence	Remarks
5′-TTAATGTCA-3′	Putative *Ubx* site present in pUC13
5′-TTAATGAAT-3′	Putative *Ubx* site present in pUC13
5′-TTAATGGTT-3′	*Ubx* site inserted at the *Sma* I site of pUC13*

*The inserted *Ubx* site was chosen on the basis of the original 7-bp consensus sequence 5′-TTAATGG-3′ (18); this has been subsequently revised to a 9-bp consensus sequence 5′-TTAAT(G/T)(G/A)CC-3′ (19).

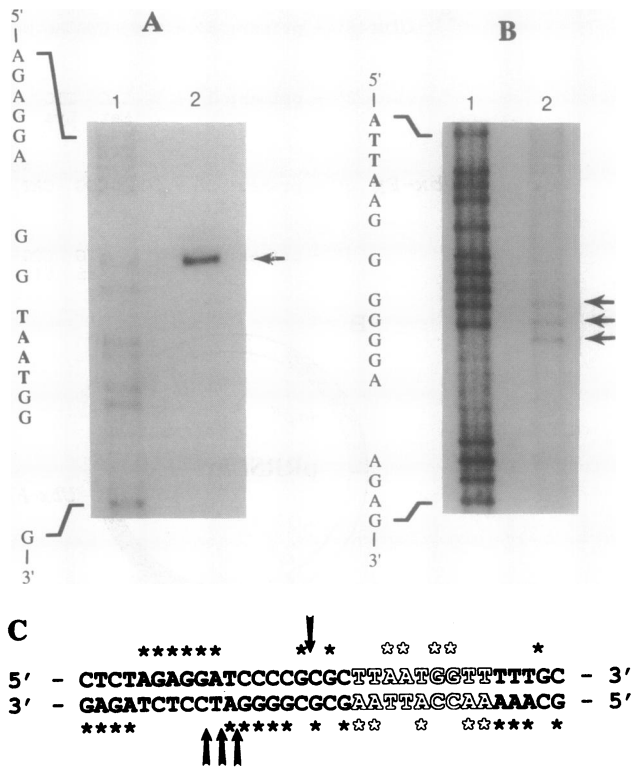

FIG. 5. Analysis of the distance of cleavage from the recognition site by Ubx–F_N. Cleavage products of the ^{32}P-labeled 82-bp DNA substrate containing a single *Ubx* site by Ubx–F_N along with (G + A) Maxam–Gilbert sequencing reactions (27) were separated by electrophoresis on a 9% polyacrylamide gel containing 6 M urea; the gel was dried and exposed to an x-ray film for 6 hr. (*A*) Cleavage product(s) from substrate containing ^{32}P-label on the 5′-TTAATGGTT-3′ strand. Lanes: 1, (G + A) sequencing reaction; 2, Ubx–F_N. (*B*) Substrate containing ^{32}P-label on the complementary strand 5′-AACCATTAA-3′. Lanes: 1, (G + A) sequencing reaction; 2, Ubx–F_N. (*C*) Map of cleavage site(s) of Ubx–F_N based on the DNA substrate containing a single *Ubx* site. The recognition site is shown by outlined letters. Site(s) of cleavage are indicated by arrows. Purine residues are indicated ⋆ for easy comparison with the gel purine tracts.

sequence-specific manner and there are two strong binding sites within the linearized pUC13 for the hybrid protein, because the four fragments can be divided into two sets; the total size of these fragments in each set adds up to that of the linearized pUC13 DNA substrate. There are also two weaker bands visible in the agarose-gel profile, indicating another weak binding site for Ubx–F_N within pUC13.

The pUC13 derivatives were constructed by inserting a 31-bp DNA fragment containing a known *Ubx* recognition sequence 5′-TTAATGGTT-3′ at the *Sma* I site of pUC13. Cleavage at the inserted *Ubx* site of the linearized pUC13 derivatives should yield ≈1.8-kb and ≈0.9-kb fragments as products. The agarose-gel electrophoretic profile of the partial digests of the substrates containing one, two, and three inserts, respectively, by Ubx–F_N is shown in Fig. 4*B* (lanes 4–6). The two additional fragments (≈1.8 kb and ≈0.9 kb, respectively) could be explained as resulting from the binding of the hybrid protein at the inserted *Ubx* site of pUC13 and cleaving near this site. As expected, the intensity of these bands increase with the number of 31-bp inserts in pUC13. The two putative strong Ubx-binding sites in pUC13 and the inserted *Ubx* site are shown in Table 1. All these sites have 5′-TAAT-3′ as their core sequence; and these preferred sites are consistent with those reported for the Ubx homeodomain. The affinity of Ubx homeodomain for these sites is modulated by the nucleotide bases surrounding the core site. It appears that the hybrid protein does turn over because most DNA

substrates in the reaction are cleaved. The cleavage is more specific at higher temperatures.

Analysis of Cleavage Distance from the Recognition Site by the Hybrid Enzyme. To determine the distance of cleavage from the recognition site by Ubx–F_N, the cleavage products of the ^{32}P-labeled DNA substrates containing a single *Ubx* site were analyzed by PAGE (Fig. 5). The digestion products were analyzed alongside the Maxam–Gilbert (G + A) sequencing reactions of the substrates (27). As expected, the cut sites are shifted away from the recognition site. On the 5′-TTAATGGTT-3′ strand Ubx–F_N cuts the DNA 3 nt away from the recognition site (Fig. 5*A*), whereas on the 5′-AACCATTAA-3′ strand it cuts 8, 9, or 10 nt away from the recognition site. Analysis of the cut sites of Ubx–F_N based on the cleavage of the DNA substrate containing a single *Ubx* site is summarized in Fig. 5*C*. The cleavage occurs 5′ to the 5′-TTAATGGTT-3′ sequence and is consistent with the way the Ubx–F_N hybrid protein was engineered (Fig. 1).

Chimeric Restriction Endonucleases. Our work on *Fok* I endonuclease (a type IIS enzyme) revealed the presence of two distinct and separable domains within this enzyme: one for the sequence-specific recognition and the other for the endonuclease activity. The modular structure of *Fok* I suggested that it may be feasible to construct hybrid endonucleases with discrete sequence-specificities by linking other DNA-binding proteins to the cleavage domain of *Fok* I endonuclease. Therefore, our approach has been simple and direct. We have successfully engineered a chimeric restriction endonuclease by linking the Ubx homeodomain to the cleavage domain of *Fok* I. The hybrid protein, Ubx–F_N, also shows nonspecific nucleolytic activity, especially at high $MgCl_2$ concentrations. There appears to be no suppression of the catalytic domain by the recognition domain when not bound to DNA sites, as occurs in native *Fok* I enzyme. The nonspecific nuclease activity of the hybrid enzyme is, therefore, an inherent property of our present construct. However, we can control the nonspecific nuclease activity of the hybrid enzyme by reducing $MgCl_2$ concentration. The Ubx homeodomain of the hybrid enzyme targets the protein to the appropriate site on the DNA. At low levels of $MgCl_2$, we can coax the cleavage domain to cut near the binding site, while greatly reducing its nonspecific nuclease activity concomitantly. The hybrid enzyme, thus, still does not behave exactly like the naturally occurring restriction endonucleases. The allosteric coupling effect of the DNA-recognition domain and the cleavage domain of *Fok* I need to be studied in more detail. One approach would be to include more linker region of *Fok* I when hybrids are constructed. Alternatively, nuclease domain of the hybrid protein might be controlled by varying salt concentration, pH, and glycerol and by altering the metal-ion requirements of the enzyme.

It is possible that the type IIS endonucleases evolved by random fusions of the DNA-binding domains to nonspecific endonucleases. Over time, these proteins were probably further refined into sequence-specific type IIS endonucleases by acquiring the allosteric interaction between the recognition domain and the catalytic domain. Future experiments are necessary to refine the hybrid endonuclease by random mutagenesis into an enzyme that communicates between the DNA-binding domain and the catalytic domain. Such refinement should remove the nonspecific nuclease activity inherent in our present construct that is probably deleterious to the host cell.

Our initial success in engineering a chimeric restriction endonuclease is quite encouraging. This is an important step toward engineering "artificial" restriction enzymes with tailor-made sequence specificity. This approach offers us an opportunity to use the repertoire of all naturally occurring DNA-binding proteins to create additional restriction enzymes, the sole determinant of the specificity of such enzymes being the DNA-binding protein. These engineered hybrid endonucleases would greatly facilitate the manipulation and mapping of genomic DNA. Furthermore, these hybrids will provide valuable information about protein structure and protein design.

The recognition by a protein of a specific sequence of bases along a strand of double-helical DNA is a key element in many biological phenomena. Perhaps the most important example is in genetic regulation, especially during embryonic development. To fully understand genetic regulation, one must not only identify the regulatory proteins but also understand their affinity for one or more sequences from a very large number present over the whole genome. By engineering hybrid endonucleases of these transcription factors one could not only identify the sequences to which they bind but also their affinity for the various substrate sites. Ultimately, one might be able to target specific genes for destruction within cells by using this approach.

We thank Professors Hamilton O. Smith and Brown Murr for encouragement and helpful discussions. We also thank Drs. Philip Beachy, Stephen Ekker, Lin Li, and J.-F. Tomb for helpful suggestions and Pat Sauers for typing the manuscript. We are grateful to the Environmental Health Sciences Center Core Facility (supported by Grant ES 03819) for synthesis of oligonucleotides. This work was supported by National Institutes of Health Grant GM 42140. S.C. has a Faculty Research Award (FRA 65569) from the American Cancer Society.

1. Sugisaki, H. & Kanazawa, S. (1981) *Gene* **16,** 73–78.
2. Szybalski, W., Kim, S. C., Hasan, N. & Podhajska, A. J. (1991) *Gene* **100,** 13–26.
3. Li, L., Wu, L. P. & Chandrasegaran, S. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4275–4279.
4. Li, L., Wu, L. P., Clarke, R. & Chandrasegaran, S. (1993) *Gene* **133,** 79–84.
5. Li, L. & Chandrasegaran, S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 2764–2768.
6. Ferrin, L. J. & Camerini-Otero, R. D. (1991) *Science* **254,** 1494–1497.
7. Strobel, S. A., Doucette-Stamm, L. A., Riba, L., Housman, D. E. & Dervan, P. B. (1991) *Science* **254,** 1639–1642.
8. Pei, D. & Schultz, P. G. (1990) *J. Am. Chem. Soc.* **112,** 4579–4580.
9. Shin, J. A., Ebright, R. H. & Dervan, P. B. (1991) *Nucleic Acids Res.* **19,** 5233–5236.
10. Ebright, R. H., Ebright, Y. W., Pendergrast, P. S. & Gunasekera, A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2882–2886.
11. Koob, M. (1992) *Methods Enzymol.* **216,** 321–329.
12. Koob, M., Grimes, E. & Szybalski, W. (1988) *Science* **241,** 1084–1086.
13. Kur, J., Koob, M., Burkiewicz, A. & Szybalski, W. (1992) *Gene* **110,** 1–7.
14. Hayashi, S. & Scott, M. P. (1992) *Cell* **63,** 883–894.
15. Qian, Y. Q., Billeter, M., Otting, M., Muller, M., Gehring, W. J. & Wuthrich, K. (1989) *Cell* **59,** 573–580.
16. Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B. & Pabo, C. O. (1990) *Cell* **63,** 579–590.
17. Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. & Pabo, C. O. (1991) *Cell* **67,** 517–528.
18. Ekker, S. C., Young, K. E., Von Kessler, D. P. & Beachy, P. A. (1991) *EMBO J.* **10,** 1179–1186.
19. Ekker, S. C., von Kessler, D. P. & Beady, P. A. (1992) *EMBO J.* **11,** 4059–4072.
20. Kita, K., Kotani, H., Sugisaki, H. & Takanami, M. (1989) *J. Biol. Chem.* **264,** 5751–5756.
21. Looney, M. C., Moran, L. S., Jack, W. E., Feehery, G. R., Benner, J. S., Slatko, B. E. & Wilson, G. G. (1989) *Gene* **80,** 193–208.
22. Skoglund, C. M., Smith, H. O. & Chandrasegaran, S. (1990) *Gene* **88,** 1–5.
23. Hochuli, E., Döbeli, H. & Schacher, A. (1987) *J. Chromatogr.* **411,** 177–184.
24. Laemmli, U. K. (1970) *Nature (London)* **222,** 680–685.
25. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
26. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60–89.
27. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.