674–681 *Nucleic Acids Research, 1999, Vol. 27, No. 2*

A detailed study of the substrate specificity of a chimeric restriction enzyme

Jeff Smith, Jeremy M. Berg and Srinivasan Chandrasegaran[1,*]

Department of Biophysics and Biophysical Chemistry, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205-2179, USA and [1]Department of Environmental Health Sciences, The Johns Hopkins University School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205-2179, USA

Received July 27, 1998; Revised and Accepted November 16, 1998

ABSTRACT

Recently, the crystal structure of the designed zinc finger protein, ΔQNK, bound to a preferred DNA sequence was reported. We have converted ΔQNK into a novel site-specific endonuclease by linking it to the *Fok*I cleavage domain (F_N). The substrate specificity and DNA cleavage properties of the resulting chimeric restriction enzyme (ΔQNK-F_N) were investigated, and the binding affinities of ΔQNK and ΔQNK-F_N for various DNA substrates were determined. Substrates that are bound by ΔQNK with high affinity are the same as those that are cleaved efficiently by ΔQNK-F_N. Substrates bound by ΔQNK with lower affinity are cleaved with very low efficiency or not at all by ΔQNK-F_N. The binding of ΔQNK-F_N to each substrate was ~2-fold weaker than that for ΔQNK. Thus, the fusion of the *Fok*I cleavage domain to the zinc finger motif does not change the DNA sequence specificity of the zinc finger protein and does not change its binding affinity significantly.

INTRODUCTION

Cys_2-His_2 zinc fingers can be designed to bind targeted DNA sequences (1–17). Zinc finger domains by themselves bind DNA and additional functions can be added by fusing other functional domains. Designed zinc fingers are being used to construct hybrid proteins that act as transcription activators, transcription inhibitors and restriction enzymes (18–26). Our aim is to use designed zinc fingers to generate hybrid restriction enzymes with targeted specificity. Chimeric restriction enzymes have been created by fusing the *Fok*I endonuclease domain (F_N) to DNA binding domains of other proteins (18–22). *Fok*I is a type IIs restriction enzyme with separate N- and C-terminal domains that correspond to the DNA-binding and endonuclease domains, respectively (27–33). The endonuclease domain by itself has non-specific DNA cleavage activity; however, when attached to another DNA-binding domain the sequence specificity of the hybrid appears to be determined by the DNA-binding domain.

Recently, one such chimeric nuclease Zif-QQR-F_N was shown to stimulate homologous recombination through targeted cleavage *in vivo* in *Xenopus* oocytes. This approach could be used to induce targeted genetic manipulation in a variety of organisms. However, one important issue needs to be addressed for the utility of chimeric nucleases as tools for gene targeting: what is the specificity of the zinc finger recognition *in vivo* and *in vitro* once it is fused to a nuclease domain? It is important to determine whether the variants of the canonical target (the degenerate sites) are cleaved or not *in vivo* as well as *in vitro*. What is the minimum binding affinity of the zinc finger protein for the substrate that is necessary for sequence specific cleavage when it is converted into a chimeric nuclease?

To date, five different designed zinc fingers have been fused to F_N to generate chimeric restriction enzymes (18–20). The cleavage sites for many of these hybrid enzymes were identified from λ-phage DNA. For four of the hybrid enzymes, the identified cleavage site was a preferred degenerate site that had been predicted from the length encoded multiplex binding assay (LEMB). In this assay, the identity of a base at a specific position in the binding site is encoded in the length of the DNA fragment (8). Measuring the DNA band intensities on a denaturing gel provides a quantitative assessment of preferred bases in the recognition site. However, the fifth hybrid enzyme's cleavage site was not predicted by LEMB suggesting that either (i) construction of a fusion protein has altered the specificity of the zinc finger, (ii) the predictions based on the LEMB assay were incorrect or (iii) factors other than DNA binding influences DNA cleavage efficiency.

In order to test these possibilities, we constructed a zinc finger-F_N fusion using the well characterized zinc finger ΔQNK and compared the specificities of the fusion to that of the original zinc finger. We have investigated in detail the *in vitro* specificity of a zinc finger-F_N fusion towards alternative binding sites. These studies provide further characterization of *in vitro* cleavage preferences by zinc finger-F_N fusion proteins.

MATERIALS AND METHODS

Construction of the clone producing the hybrid enzyme ΔQNK-F_N using PCR

Experimental procedures for PCR are described elsewhere (27–29). The ΔQNK zinc finger was PCR amplified using oligodeoxyribonucleotide d(GACTAGTCCCTTCTTATTCTGGT) (start primer) and d(GGACTAGTCCCTTCTTATTCTGATG)

*To whom correspondence should be addressed. Tel: +1 410 614 2289; Fax: +1 410 955 0617; Email: chandra@welchlink.welch.jhu.edu

(stop primer) and was digested with *Nde*I/*Spe*I and then ligated into *Nde*I/*Spe*I-cleaved pET-15b: *Ubx-F*$_N$ vector, which contains the sequence encoding the *Fok*I nuclease (F_N) domain. This construct replaces the *Ubx* homeodomain with the gene coding for the three zinc fingers QNK QDK RHR abbreviated ΔQNK. The ligation mixture was used to transform competent RR1 (pACYC184:*lig*) cells. The glycine linker $(Gly_4Ser)_3$ was inserted between the zinc finger motif and the F_N domain using procedures described elsewhere (32). The ΔQNK fusion constructs were confirmed by dideoxynucleotide sequencing (34). The pET-15b:*Zif-F*$_N$ plasmids were then transformed into BL21 (DE3) cells which also carry the pACYC184:*lig* plasmid.

Purification of ΔQNK-F_N and ΔQNK

ΔQNK-F_N protein purification: 4.2 l of BL21 (DE3)[pACYC184:*lig*, pET-15b:*Zif-F*$_N$] cells were grown in LB containing 100 μg/ml of ampicillin, 20 μg/ml tetracycline and 100 μM $ZnCl_2$ at 37°C. When A_{600nm} reached 0.4, the growth temperature was shifted to 22°C. Cells were induced at A_{600nm} = 0.5 with 0.7 mM of IPTG for 4.5 h at 22°C. Cells were harvested by centrifugation, resuspended in binding buffer [20 mM Tris–HCl, pH 7.9, 5 mM imidazole, 100 μM $ZnCl_2$, 10% (v/v) glycerol, 500 mM NaCl and 0.1 mM PMSF] and disrupted by 10 min sonication on ice. After centrifugation at 4°C for 40 min, the crude extract was passed through a 0.45 μm filter and applied to a His-bind affinity column. Purification of protein using His-bind (pET System Manual, 6th Edition) was done as outlined in Novagen pET system manual. The column was washed with 10 column vol of binding buffer, 6 vol of binding buffer with 60 mM imidazole and 4 vol of binding buffer with 100 mM imidazole. As determined by western blots with rabbit polyclonal antibody raised against *Fok*I, ΔQNK-F_N began to elute in the last wash and was fully eluted in 4 column vol of binding buffer with 400 mM imidazole. Fractions containing ΔQNK-F_N were combined and diluted 10-fold in TDZ buffer (20 mM Tris–HCl, pH 7.9, 3 mM DTT, 100 μM $ZnCl_2$ and 0.1 mM PMSF). The diluted solution was loaded over a heparin–sepharose column, washed with 10 vol of TDZ with 100 mM NaCl and 2 vol of TDZ with 600 mM NaCl. ΔQNK-F_N was eluted in 2 vol of TDZ with 1 M NaCl. The elution was concentrated and dialyzed on a 25 kDa molecular weight cut off colloidion membrane and then loaded on an S-100 HR gel filtration column. The column was equilibrated against dialysis buffer [20 mM Tris–HCl, pH 7.9, 100 mM $ZnCl_2$, 10% (v/v) glycerol, 500 mM NaCl, 3 mM DDT and 0.1 mM PMSF]. Following gel filtration, pure fractions according to silver stained SDS–PAGE were combined and the fusion protein was stored in 50% glycerol at –70°C. ΔQNK was partially purified using the boiling method as described (15) and purification was completed using a heparin–sepharose column as is described above for ΔQNK-F_N.

Generation of synthetic substrates

Based on the length encoded multiplex binding studies of ΔQNK zinc fingers, the following binding sites were synthesized: d(TTTGGGGCGGAATTT), d(TTTGGGGCAGAATTT), d(ATAGGGTCGGAA), d(ATAGGGTCAGAA), d(ATAGGGGTGGAAATA) and d(ATAGGGATCGAAATA). The oligos were phosphorylated with T4 polynucleotide kinase and annealed to their corresponding complementary oligonucleotides. The 15mer oligo duplexes were used in a 3 to 1 molar excess over pUC18 DNA in Pharmacia Biotech pUC18 *Sma*I/BAP + ligase Ready-To-Go kit. The recombinant DNA was precipitated according to the Ready-To-Go kit procedure and was electroporated into DH5α cells. The transformed cells were grown on LB plates containing 75 μg/ml ampicillin, 40 mg/ml X-Gal and 0.12 mM IPTG. White colonies were picked, plasmids were isolated using Qiagen plasmid purification kit and the number of inserted binding sites was determined by dideoxynucleotide sequencing (34).

ΔQNK-F_N activity assays

The conditions for sequence-specific cleavage by ΔQNK-F_N are as follows: ~20 nM protein and ~2 nM pUC18 with inserted synthetic binding site were incubated in 20 mM Tris–HCl (pH 8.5), 75 mM NaCl, 100 μM $ZnCl_2$, 1 mM DTT, 5% (v/v) glycerol, 100 mg/ml tRNA, 3.75 ng/ml double stranded 15mer oligodeoxyribonucleotide containing the d(GGGTCAGAA) binding site and 50 mg/ml bovine serum albumin with a total volume of 19 μl for 30 min at 22°C. After incubation, 10 mM $MgCl_2$ was added raising the total reaction volume to 20 μl and the mixture was incubated at 22°C for 1 h. The reaction was stopped by heating at 70°C for 10 min to denature the endonuclease domain. Proteinase K (65 mg/ml) was added and the mixture was incubated at 37°C for 30 min to digest the protein and then at 70°C for 10 min to inactivate the proteinase K. The reaction products were analyzed by 1% agarose gel electrophoresis. The 15mer containing d(GGGTCAGAA) was added to the reaction to prevent the ΔQNK-F_N from binding to DNA via only two of its three zinc fingers.

Measuring initial rates

The pUC18 plasmids containing the synthetic substrates were linearized with *Sca*I, dephosphorylated with calf intestinal phosphorylase and radiolabeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The labeled pUC18 substrates were gel purified on a 2% low melting point gel, phenol/chloroform extracted and ethanol precipitated. A 100 μl cleavage reaction was set up for each substrate using the conditions and concentrations described above with the following exceptions: 13 nM ΔQNK-F_N per reaction instead of 20 nM, 1 nM labeled substrate was used, and there was no 30 min preincubation with ΔQNK-F_N before adding 10 mM $MgCl_2$. The reaction was timed from the addition of ΔQNK-F_N. A 10 μl fraction was removed at each time point, heat inactivated at 70°C and digested with Proteinase K. The Proteinase K was heat inactivated at 70°C and the samples were run out on a 1% agarose gel. The gel was dried and then imaged using a Molecular Dynamics Storm 860 PhosphorImager. The two bands resulting from cleavage were quantified using ImageQuant software and summed to determine the product concentration. The data was plotted using Cricket graph software.

Gel-shift experiments and K_D determination

The DNA probes were prepared by isolating the 50 bp oligonucleotide containing the single inserted synthetic substrate from pUC18 using *Bam*HI or *Pst*I and *Eco*RI. The ends were labeled with [α-^{32}P]ATP using Klenow fragment and then filled in completely by a subsequent incubation with 2 mM dNTPs. Unincorporated [α-^{32}P]ATP was removed using a Qiagen

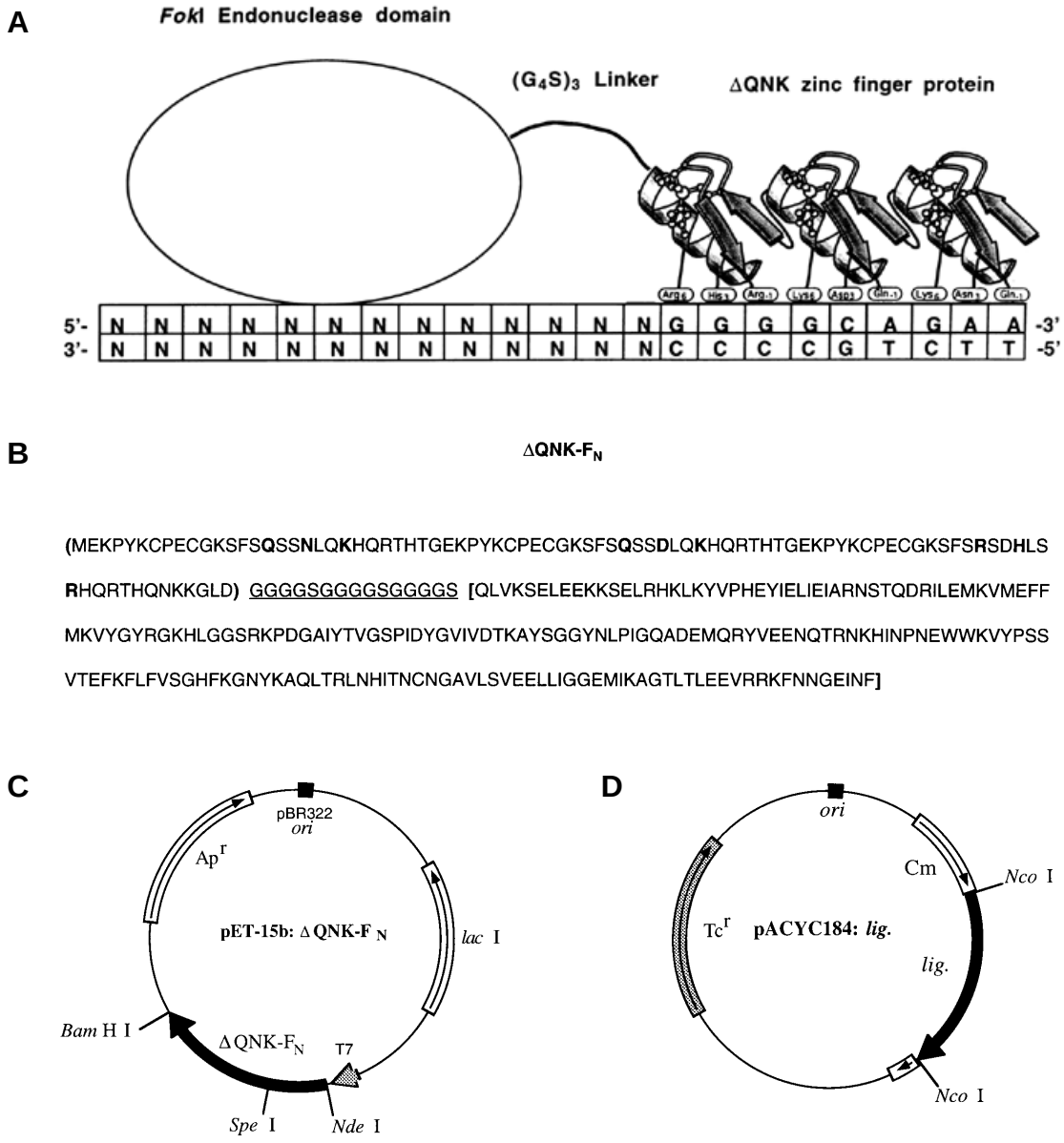

Figure 1. Plasmid and protein constructs. (**A**) Schematic representation of the chimeric restriction enzyme ΔQNK-F_N aligned with its cognate DNA site. (**B**) Amino acid sequence of the chimeric restriction enzyme, ΔQNK-F_N. The zinc finger protein ΔQNK is shown in parenthesis. Specificity determining residues in ΔQNK are shown in boldface (36). The $(G_4S)_3$ linker is underlined and the *Fok*I nuclease domain is in brackets. (**C**) pET-15b:Δ*QNK-F*$_N$; (**D**) pACYC184:*lig.*

QIAquick Nucleotide Removal Kit. The labeled probe was purified on a 10% native acrylamide gel (30:1 acrylamide:bisacrylamide). The DNA was removed from the gel slice by soaking overnight in 0.5 M ammonium acetate and with ethanol precipitation. All six probes were resuspended in distilled water to final concentrations of ~4500 c.p.m./µl. Various concentrations of ΔQNK and ΔQNK-F_N were incubated with 1 µl of each of the six different radioactive probes in 20 µl gel shift buffer containing 25 mM Tris–HCl (pH 8.0), 100 mM NaCl, 2 mM DTT, 100 µM $ZnCl_2$, 10% glycerol, 50 µg/ml BSA and 4 µg/ml polydeoxyinosinic-deoxycytidylic acid [poly(dI-dC)] at 4°C for 30 min. The final concentration of probe was ~6 nM. The reaction mixtures were separated on 10% acrylamide gels (30:1 acrylamide:bisacrylamide) in 50 mM Tris (pH 8.0) and 400 mM glycine buffer at 100 V for 3 h (19,35). The concentration of ΔQNK was determined by absorbance at 280 nm using an extinction coefficient of 4200 M/cm. The concentration of ΔQNK-F_N was determined using a Bradford assay and a BSA standard curve. The dissociation constant of ΔQNK and ΔQNK-F_N for each of the six substrates was estimated with a gel-shift assay. We determined the binding affinities using five different protein concentrations centered around the estimated dissociation constant. We ran three independent ΔQNK gel shift experiments for all six substrates and three independent ΔQNK-F_N gel shift experiments only for

probes containing GCG, GCA and TCA. One gel shift using ΔQNK-F_N was run for probes containing TCG, GTG and ATC. The shifted bands were visualized by autoradiography. These bands were quantified using a densitometer and ImageQuant software. The K_D values were determined from the slope of Scatchard plots prepared on Cricket Graph software.

RESULTS

Construction of overproducing clone of ΔQNK-F_N

Kim and Berg recently solved the crystal structure of ΔQNK bound to a preferred DNA binding site (36). We used PCR to link the zinc-finger protein, ΔQNK, to the cleavage domain of *Fok*I endonuclease (18) (Fig. 1A). The construct links the consensus framework based zinc finger through a glycine linker to the C-terminal 196 amino acids of *Fok*I, which constitutes the *Fok*I cleavage domain (F_N) (Fig. 1B). Since there is no methylase available to protect the host genome from cleavage by the hybrid endonuclease, we cloned the hybrid gene into a tightly controlled expression system to minimize leaky expression (Fig. 1C). Furthermore, protection was provided by increasing the level of DNA ligase within the cells by placing the *Escherichia coli lig* gene on a compatible plasmid, pACYC184, downstream of the chloramphenicol promoter (Fig. 1D).

Purification of ΔQNK-F_N

The ΔQNK-F_N was expressed in BL21(DE3) cells (37). BL21 (DE3) cells contain a chromosomal copy of the T7 RNA polymerase gene driven by the isopropyl β-D-thiogalactoside (IPTG) inducible *lacUV5* promoter. After induction of the hybrid protein with 0.7 mM IPTG for 4 h at 22°C, the hybrid protein was purified to homogeneity using His· Bind resin (38), a heparin–sepharose column and gel-filtration chromatography. The SDS–PAGE profiles of the ~38 kDa purified hybrid enzyme are shown in Figure 2A. We confirmed the identity of the hybrid protein by immunoblotting with rabbit antiserum raised against *Fok*I endonuclease (Fig. 2B).

DNA substrate specificity of the ΔQNK-F_N hybrid restriction enzyme

The substrate specificity of ΔQNK was previously studied using the LEMB (8). The base preferences at each position of the three zinc fingers are shown in Figure 3A. The first and last zinc fingers have a high specificity for d(GAA) and d(GGG), respectively. The middle zinc finger of ΔQNK appears to have the following degeneracy: d[(G/T)C(G/A)] with G preferred over T at the first position and G slightly preferred over A at the third position. We synthesized a series of substrates, each of the form d(GGG NNN GAA), in which the region expected to be contacted by the middle finger (designated NNN) was varied. We will henceforth refer to these DNA substrates according to identities of these three bases. Four sites GCG, GCA, TCG and TCA were synthesized. To test the limits of recognition we also synthesized GTG and ATC. The GTG substrate has a preferred G at the first and third positions, but the most preferred C at the second position was replaced with the least preferred base, T. The GTG substrate would reveal the importance of the middle base and the effect of a single unfavorable base in the recognition site. The ATC substrate has the least preferred base at every position and was therefore

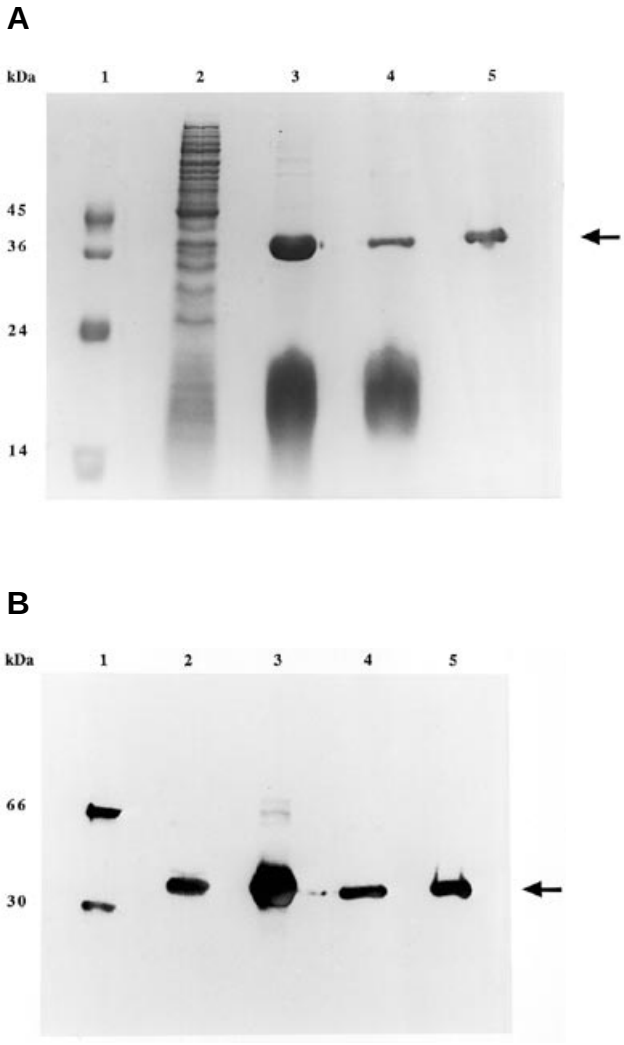

Figure 2. Purification of ΔQNK-F_N. (**A**) A Coomassie Brilliant Blue stained SDS–PAGE gel showing fractions from each step in the purification of the ΔQNK-F_N hybrid enzyme. Lane 1, protein standards; lane 2, supernatant from induced cells after sonication; lane 3, His·Bind column; lane 4, heparin–sepharose column; lane 5, gel-filtration column. (**B**) A western blot showing fractions from each step in the purification of the ΔQNK-F_N. The blot was probed with rabbit polyclonal antibody against *Fok*I. Lane identity is same as above. Arrows indicate the ΔQNK-F_N. Purification protocol is described in Materials and Methods.

predicted to be a very poor substrate for the hybrid restriction enzyme. These synthetic substrates were inserted into pUC18 as described in Materials and Methods. When used in a cleavage assay, each pUC18 construct was first linearized at the *Sca*I site so that cleavage at the inserted substrate by ΔQNK-F_N would produce two DNA fragments of ~1.7 and 1.0 kb (Fig. 3B, arrows). The pUC18 plasmid alone has no natural ΔQNK-F_N binding sites (Fig. 3B, lane 2); and all pUC18 derivatives are not cleaved in the absence of ΔQNK-F_N (Fig. 3B, lanes 3, 5, 7, 9, 11 and 13). The almost complete disappearance of the substrate band and strong appearance of 1.7 and 1.0 kb bands indicate a strong specific cleavage at GCG (Fig. 3B, lane 4). Substrates GCA and TCA were also cleaved, but more weakly. The TCA site was different from both GCG and GCA in that TCA was not strong enough to

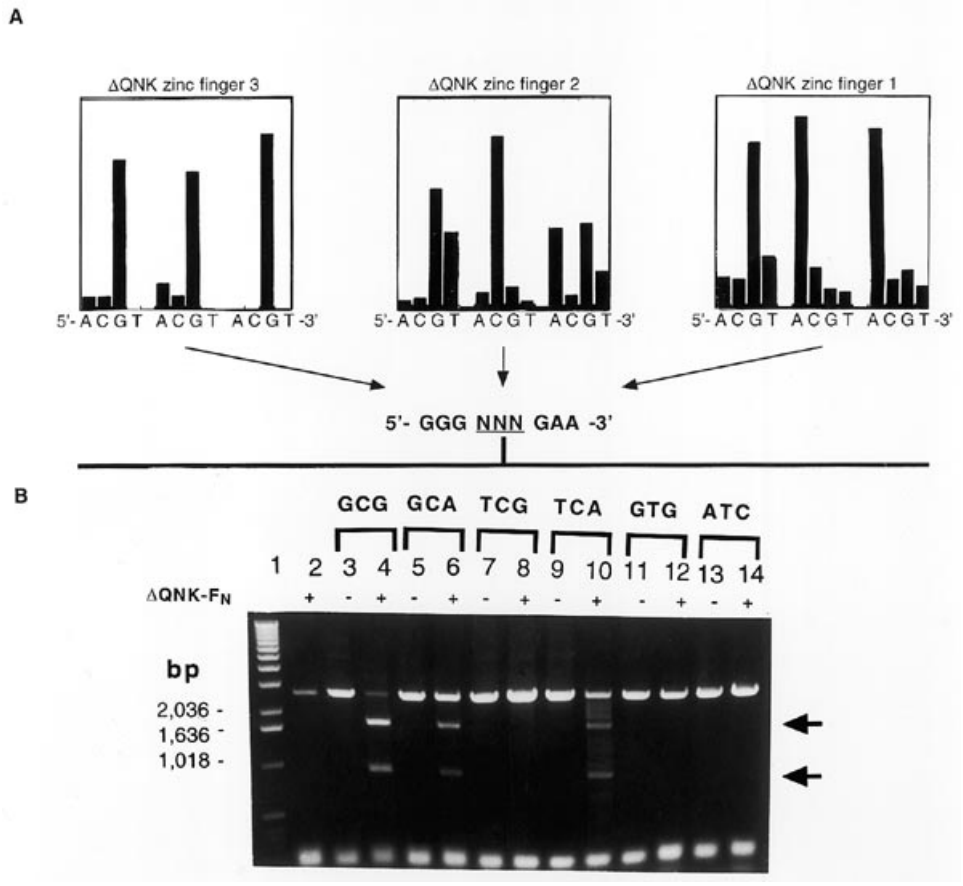

Figure 3. ΔQNK-F_N cleavage of pUC18 derivatives with a single synthetic DNA insert. (**A**) The predicted binding preferences of ΔQNK. The bar height (*y* axis) represents the fraction of sites selected with the base identified under the bar (*x* axis) at the indicated position in the binding site. The preference for fingers 1 and 2 were determined by LEMBs and were calculated from 13 selected sequences for finger 3 (36). (**B**) The pUC18 derivatives containing ΔQNK-F_N binding sites were linearized at the *Sca*I site. Cleavage by ΔQNK-F_N at the insert will result in 1700 and 1000 bp fragments. Lane 1, kb ladder; lane 2, pUC18; lanes 3–14, pUC18 with the insert: 5′-GGG NNN GAA-3′ where NNN represents the middle triplet that was varied as indicated above the brackets. Lanes 2, 4, 6, 8, 10, 12 and 14 contain ΔQNK-F_N. Lanes 3, 5, 7, 9, 11 and 13 do not. Arrows indicate cleavage fragments. tRNA appears at the bottom of the gel. Cleavage conditions are described in Materials and Methods.

Table 1. Substrate specificity and affinity for ΔQNK and ΔQNK-F_N

Substrate	ΔQNK zinc finger K_D (nM)	ΔQNK-F_N fusion K_D (nM)	ΔQNK-F_N fusion cleavage activity	Initial rates of cleavage by ΔQNK-F_N (nM/min)
d(GGG **GCG** GAA)	16[a]	25[a]	+++	0.25
d(GGG **GCA** GAA)	24[a]	54[a]	++	0.04
d(GGG **TCA** GAA)	142[a]	>158.4	+	< 0.001
d(GGG **TCG** GAA)	990[a]	ND	-	ND
d(GGG **GTG** GAA)	7500[b]	ND	-	ND
d(GGG **ATC** GAA)	25000[b]	ND	-	ND

ND, not determined.
[a]Error ≤17%.
[b]Error >17%.

limit cleavage to specific sites as evidenced by higher background cleavage. This order of cleavage preference was further supported by determination of the initial rates of cleavage of ΔQNK-F_N for GCG, GCA and TCA (Fig. 4 and Table 1). The GCG substrate was cleaved 6-fold faster than GCA which, in turn, cleaved 60-fold faster than TCA. This order of cleavage preference matches with that expected from the LEMB results. Also as predicted, GTG and ATC did not direct specific cleavage (Fig. 3B, lanes 12 and 14). However, contrary to expectations, TCG did not direct specific cleavage (Fig. 3, lane 8). The

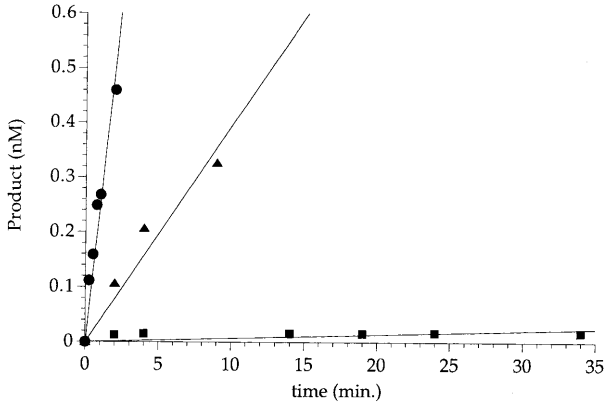

Figure 4. Cleavage profiles of various substrates by ΔQNK-F_N. Samples were collected from ΔQNK-F_N cleavage reactions of GCG(●), GCA(▲), and TCA(■) at indicated times. Reaction conditions are described in the Materials and Methods.

discrepancy between the predicted and observed ability of TCG to direct specific cleavage led us to further test whether ΔQNK-F_N had an altered DNA binding specificity from that of ΔQNK.

Comparing substrate specificity of ΔQNK and ΔQNK-F_N using gel-shift assays

If TCG failed to direct specific cleavage because the binding specificity of the ΔQNK-F_N fusion had changed, then ΔQNK and ΔQNK-F_N should have different affinities for the substrates as measured by gel-shift assays. For example, ΔQNK would shift TCG more strongly than TCA as predicted from LEMB, and ΔQNK-F_N would shift TCG less strongly than TCA based on cleavage results. However, if ΔQNK and ΔQNK-F_N had the same specificity for TCG then they would shift TCG identically, relative to the other five substrates. A 50 bp DNA containing a single inserted recognition site was used as the gel shift probe (Fig. 5). From the gel-shift assays, we found that the orders of specificity of ΔQNK and ΔQNK-F_N were identical, namely GCG > GCA > TCA > TCG > GTG, ATC. We conclude that the specificity of the ΔQNK zinc finger was not changed by the addition of the *Fok*I endonuclease domain. The optimal binding site GCG was shifted reproducibly higher by the zinc finger protein compared to other degenerate binding sites (Fig. 5). This is likely to be due to bending or kinking of the DNA within the complex containing the optimal substrate.

K_D values of ΔQNK and ΔQNK-F_N

Our conclusions based on the qualitative gel-shift results were confirmed by the K_D values of ΔQNK and ΔQNK-F_N determined from three independent gel shift experiments (Table 1). Although the gel-shifts were run in the absence of Mg^{2+}, ΔQNK-F_N is able to maintain partial activity by using the Zn^{2+} which was present for zinc finger folding. The ΔQNK-F_N gel-shifts result in multiple bands because the probe was being degraded by this partial nuclease activity. We determined the K_D values for ΔQNK-F_N by summing the band intensities which are bracketed in Figure 4. ΔQNK and ΔQNK-F_N have the highest affinity for GCG with K_D = 16 and 25 nM, respectively. The order of K_D values is the same for ΔQNK and ΔQNK-F_N: TCA > GCA > GCG. The percent error in the K_D measurements for ΔQNK are ≤17% for GCG, GCA, TCA and TCG. The error is >17% for GTG and ATC because of less well-defined gel-shift patterns. GCG and GCA were the only substrates for which a K_D for ΔQNK-F_N could be determined.

The effects of multiple recognition sites on cleavage

A tandem repeat of low affinity sites may increase the chance of cleavage by the hybrid enzyme. To test this possibility, we selected pUC18 clones with four tandem repeats of synthetic sites with various orientations (Table 2). The pUC18 derivatives were linearized at the *Sca*I site and digested with ΔQNK-F_N (Fig. 6). Plasmid pUC18 with multiple inserts were not cleaved in the absence of ΔQNK-F_N (Fig. 6, lanes 2, 4, 6 and 8). As expected, ΔQNK-F_N cleaves the substrates containing a mixture of GCG and GCA sites strongly (Fig. 6, lane 3). ΔQNK-F_N demonstrates

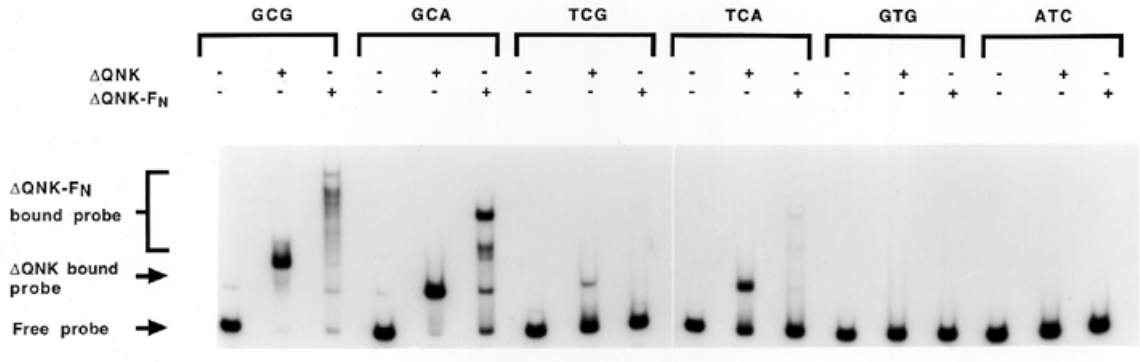

Figure 5. Gel-shift assays comparing ΔQNK zinc finger and ΔQNK hybrid enzyme affinity for various substrates. The DNA probe is a 50mer containing a single binding site: d(GGG NNN GAA) where NNN represents the middle triplet. This was varied as indicated above the brackets. Lane description is identical for both gels. Lane 1, DNA probe; lane 2, +ΔQNK zinc finger [831 nM]; lane 3, +ΔQNK-F_N [158 nM]. ΔQNK repeatedly shifted the most preferred substrate d(GGG GCG GAA) higher than the other substrates. This may be attributed to the bending of the DNA by ΔQNK-F_N when it binds to its optimal substrate. Gel-shift conditions are described in Materials and Methods. K_D values for ΔQNK-F_N were determined by summing the intensities of bands bracketed by ΔQNK-F_N bound probe. This excluded the lowest shifted band because this band is believed to be due to the contaminating ΔQNK resulting from the degradation of ΔQNK-F_N. In addition, ΔQNK-F_N appears to be active in presence of Zn^{2+} resulting in multiple bands for the complex.

Table 2. Specificity of ΔQNK-F_N for substrates containing multiple binding sites

Middle triplet	Synthetic DNA substrates[a]	Activity
GCA / GCG	5'-**TTCTGCCCC**AAAAA**TTCTGCCCC**AAAAA**TTCCGCCCC**AAATTGGGGCAATTTAAA**TTCTGCCCC**-3'	++
TCA	5'-ATA**GGGTCAGAA**ATAATA**GGGTCAGAA**ATTAT**TTCTGACCC**TATATA**GGGTCAGAA**ATA-3'	+
GTG	5'-TAT**TTCACCCC**TATATA**GGGGTGGAA**ATATA**GGGGTGGAA**ATAATA**GGGGTGGAA**ATA-3'	-
ATC	5'-TAT**TTCGATCCC**TATTAT**TTCGATCCC**TATATA**GGGATCGAA**ATAATA**GGGATCGAA**ATA-3'	-

[a]Arrows indicate orientation of binding site.

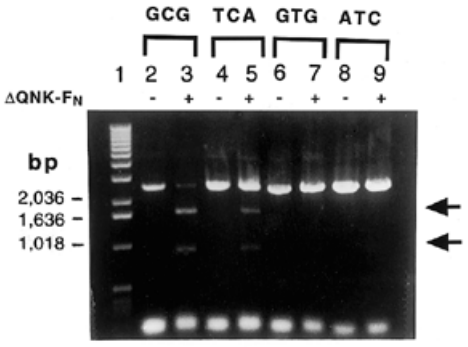

Figure 6. ΔQNK-F_N cleavage of pUC18 with multiple synthetic DNA inserts. Each substrate contains four tandem inserts with the sequence of the middle triplet indicated above the brackets. Refer to Table 2 for the complete sequence of each substrate. Lane 1, kb ladder; lanes 2, 4, 6 and 8, –ΔQNK-F_N; lanes 3, 5, 7 and 9, +ΔQNK-F_N. Arrows indicate the cleaved fragments. tRNA appears at the bottom of the gel.

a weaker cleavage of the substrate containing multiple TCA sites (Fig. 6, lane 5). The smear that occurs with the TCA single insert in lane 10 of Figure 3 is not apparent here however. This suggests that multiple sites do increase sequence specific interactions and may be used to reduce non-specific cleavage background. Substrates with multiple GTG and ATC sites are not cleaved by ΔQNK-F_N (Fig. 6, lanes 7 and 9). This suggests that multiple tandem sites do not increase interactions with unfavorable substrates enough so that they can be cleaved. pUC18 containing three tandem TCG recognition sites was also treated with ΔQNK-F_N and did not result in any specific cleavage (data not shown). Although multiple sites appear to decrease non-specific cleavage background, they do not alter the cleavage specificity of the hybrid enzyme.

DISCUSSION

Kim *et al.* constructed the first two zinc finger-F_N hybrid restriction enzymes, CP-QDR-F_N and Sp1-QNR-F_N (18). Their preferred cleavage sites in λ-DNA were identified and compared to the specificity predicted from the multiplex binding assays. CP-QDR-F_N cleaved a site in λ that was predicted by LEMB where as Sp1-QNR-F_N cleaved a site that was not predicted by LEMB. Two possibilities were given to explain the apparent new specificity of Sp1-QNR-F_N: (i) the addition of a functional domain, in this case a cleavage moiety, alters the sequence specificity of the zinc finger or (ii) the predictions based on the LEMB assay are incorrect. We reasoned that by making a fusion with a well-characterized zinc finger, we would be able to shed some light on the apparent discrepancy in the substrate specificity. Therefore, we engineered a chimeric restriction enzyme using the zinc finger ΔQNK whose crystal structure with a cognate DNA site has been determined (36). The substrate specificity of this hybrid restriction enzyme was investigated in detail. We have observed a similar discrepancy in the substrate cleavage specificity of the hybrid enzyme when compared to the zinc finger specificity predicted by LEMB assays. Our gel shift experiments and K_D measurements demonstrate that the specificity of the zinc finger and the hybrid enzyme is the same for all four degenerate sites as well as two unfavorable sites. The specificity of the zinc fingers is not changed by the fusion of the *Fok*I endonuclease domain to the C-terminus of the zinc finger. The discrepancy appears to be due to an assumption in the analysis of data using the LEMB. This assay is based on the assumption that the specificity at one position within the bindings site is independent of the other bases. Our binding and cleavage results with TCG substrate, compared with GCG and GCA and TCA reveal that this assumption is invalid for this protein.

Since fusion of the cleavage moiety, F_N, to a zinc finger does not change the specificity, the capacity to engineer proteins with designed DNA specificity using zinc fingers can be directly extended to engineering hybrid restriction enzymes with the same sequence specificity. The K_D measurements also have strong implications towards future Zif-F_N constructs: a comparison of the K_D values for ΔQNK and the cleavage activities of ΔQNK-F_N in Table 1 indicates the minimum binding affinity required for the substrate (~140 nM) for sequence specific cleavage. Using zinc fingers that have only one distinct site with a $K_D \leq 140$ nM, one can engineer hybrid restriction enzymes that cleave the one distinct site and not at variant sites. These new insights into *in vitro* cleavage specificity of the zinc finger fusion protein may be

useful in designing better chimeric nucleases to induce targeted genetic manipulation in a variety of organisms.

Finally, randomly oriented tandem recognition sites appear to decrease non-specific cleavage without changing the specificity of the zinc fingers. The multiple site results do not exclude the possibility that multiple sites with specific orientation and separation may alter the cleavage specificity. Such sites may aid the dimerization of the *Fok*I nuclease domain. A recent report from Bitinaite *et al*. (39) indicates that dimerization of the *Fok*I nuclease domain may stimulate an increased cleavage activity. Digestion experiments using substrates with a wide range of site orientations and spacings must be done in order to determine if changes in specificity occur. Our current fusion proteins leave out domains of *Fok*I that maintain an allosteric control over the endonuclease domain (33). These fusion proteins often produce low background smears of degraded DNA because of the uncontrolled nuclease activity of the endonuclease domain. The reduction of this DNA degradation background by multiple sites will be useful in generating cleaner digests until new fusions that limit the nuclease activity can be engineered.

ACKNOWLEDGEMENTS

We thank Professor Hamilton O. Smith for helpful suggestions and Kay Castleberry for typing the manuscript. We also thank Mr Gregory J. Gatto for helping with the preparation of the figures. This work was supported by grants from the National Institutes of Health (GM53923) and National Science Foundation (MCB9415861). We also thank Environmental Health Sciences Center Core Facility (supported by Grant ES03819) for synthesis of oligonucleotides. S.C and J.M.B. are members of the scientific advisory board of Sangamo BioSciences, Inc.

REFERENCES

1 Berg,J.M. (1997) *Nature Biotechnol*., **15**, 323.
2 Corbi,N., Perez,M., Malone,R. and Passananti,C. (1997) *FEBS Lett*., **417**, 71–74.
3 Greisman,H. and Pabo,C. (1997) *Science*, **275**, 657–661.
4 Rebar,E., Greisman,H. and Pabo,C. (1996) *Methods Enzymol*., **276**, 129.
5 Berg,J. (1995) *Acc. Chem. Res*., **28**, 14–19.
6 Choo,Y. and Klug,A. (1995) *Curr. Opin. Biotechnol*., **6**, 431.
7 Wu,H., Yang,W.-P. and Barbas,C.,III (1995) *Proc. Natl Acad. Sci. USA*, **92**, 344–348.
8 Desjarlais,J. and Berg,J. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 11099.
9 Choo,Y. and Klug,A. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 11163.
10 Jamieson,A., Kim,S.-H. and Well,J. (1994) *Biochemistry*, **33**, 5689.
11 Rhodes,D. and Klug,A. (1993) *Sci. Am*., **268**, 56–65.
12 Fairall,L., Schwabe,J.W.R., Chapman,L., Finch,J.T. and Rhodes,D. (1993) *Nature*, **366**, 483–487.
13 Pavletich,N.P. and Pabo,C.O. (1993) *Science*, **261**, 1701.
14 Desjarlais,J.R. and Berg,J.M. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 2256–2260.
15 Desjarlais,J. and Berg,J. (1992) *Proteins*, **12**, 101–104.
16 Pavietich,N.P. and Pabo,C.O. (1991) *Science*, **252**, 809–817.
17 Berg,J.M. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 99–102.
18 Kim,Y.-G., Cha,J. and Chandrasegaran,S. (1996*) Proc. Natl Acad. Sci. USA*, **93**, 1156–1160.
19 Huang,B., Schaeffer,C.J., Li,Q. and Tsai,M.-D. (1996*) J. Prot. Chem*., **15**, 481–489.
20 Kim,Y.-G., Shi,Y., Berg,J.M. and Chandrasegaran,S. (1997) *Gene*, **203**, 43–49.
21 Kim,Y.-G. and Chandrasegaran,S. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 883–887.
22 Kim,Y.-G., Smith,J., Durgesha,M. and Chandrasegaran,S. (1998) *Biol. Chem*., **379**, 489–495.
23 Pomerantz,J., Sharp,P. and Pabo,C. (1995) *Science*, **267**, 93–96.
24 Pomerantz,J., Pabo,C. and Sharp,P. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 9752–9756.
25 Kim,J.-S., Kim,J., Cepek,K., Sharp,P. and Pabo,C. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 3616–3362.
26 Choo,Y., Sanchez-Garcia,I. and Klug,A. (1994) *Nature*, **372**, 642.
27 Sugisaki,H. and Kamazawa,S. (1981) *Gene*, **16**, 73–78.
28 Szybalski,W., Kim,S.C., Hasan,N. and Podhajska,A.J. (1991) *Gene*, **100**, 13–26.
29 Li,L., Wu,L.P. and Chandrasegaran,S. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 4275–4279.
30 Li,L., Wu,L.P., Clark,R. and Chandrasegaran,S. (1993) *Gene*, **133**, 79–84.
31 Li,L. and Chandrasegaran,S. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 2764–2768.
32 Kim,Y.-G., Li,L. and Chandrasegaran,S. (1994) *J. Biol. Chem*., **269**, 31978–31982.
33 Wah,D.A., Hirsch,J.A., Dormer,L.F., Schildkraut,I. and Aggarwal,A.K. (1997) *Nature*, **388**, 97–100.
34 Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
35 Shi,Y. and Berg,J. (1995) *Science*, **268**, 282–284.
36 Kim,C. and Berg,J.M. (1996) *Nature Struct. Biol*. **3**, 940–946.
37 Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol*., **185**, 60–89.
38 Hochuli,E., Dobeli,H. and Schacher,A. (1987) *J. Chromatogr*., **411**, 177–184.
39 Bitinaite,J., Wah,D., Aggarwal,A. and Schildkraut,I. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 10570–10575.