Engineering a nicking endonuclease N.*Alw*I by domain swapping

Yan Xu, Keith D. Lunnen, and Huimin Kong*

New England Biolabs, 32 Tozertypeoad, Beverly, MA 01915

Edited by Brian W. Matthews, University of Oregon, Eugene, OR, and approved September 19, 2001 (received for review May 2, 2001)

Changing enzymatic function through genetic engineering still presents a challenge to molecular biologists. Here we present an example in which changing the oligomerization state of an enzyme changes its function. Type IIs restriction endonucleases such as *Alw*I usually fold into two separate domains: a DNA-binding domain and a catalytic/dimerization domain. We have swapped the putative dimerization domain of *Alw*I with a nonfunctional dimerization domain from a nicking enzyme, N.*Bst*NBI. The resulting chimeric enzyme, N.*Alw*I, no longer forms a dimer. Interestingly, the monomeric N.*Alw*I still recognizes the same sequence as *Alw*I but only cleaves the DNA strand containing the sequence 5′-GGATC-3′ (top strand). In contrast, the wild-type *Alw*I exists as a dimer in solution and cleaves two DNA strands; the top strand is cleaved by an enzyme binding to that sequence, and its complementary bottom strand is cleaved by the second enzyme dimerized with the first enzyme. N.*Alw*I is unable to form a dimer and therefore nicks DNA as a monomer. In addition, the engineered nicking enzyme is at least as active as the wild-type *Alw*I and is thus a useful enzyme. To our knowledge, this is the first report of creating a nicking enzyme by domain swapping.

Restriction endonucleases are DNA cleavage enzymes often found in bacteria (1). Together with DNA methyltransferases, restriction enzymes offer bacteria an effective defense system against bacteriophage and other foreign DNAs. Based on the cofactor requirement of the purified enzymes, restriction enzymes fall into three major classes: type I, type II, and type III. Type II restriction enzymes, which require only Mg^{2+}, are typically homodimers that bind to palindromic DNA sequences and cleave both DNA strands symmetrically. Type IIs restriction enzymes also require Mg^{2+} but recognize asymmetric DNA sequences and cleave both DNA strands outside of their recognition sequences at fixed distances. The defined sequence-specific cleavage activity of type II and type IIs restriction enzymes has proved to be a powerful tool for molecular biologists.

Both type II and type IIs restriction endonucleases cleave both DNA strands within a very short distance and therefore break the DNA molecules at their cleavage sites. There are proteins in the literature that break only one DNA strand and therefore introduce a nick into the DNA molecule. Most nicking endonucleases are found in protein complexes involved in the initiation of DNA replication (2) or in DNA repair (3). Because of the complexity of the recognition sequence and requirement for other proteins, such enzymes are not very useful in manipulating DNA. Two restriction endonuclease-like nicking enzymes, N.*Bst*SEI (4) and N.*Bst*NBI (5), have been reported. They were isolated independently from different strains of *Bacillus stearothermophilus*. Both of them recognize the 5′-GAGTC-3′ sequence and nick 4 bp downstream of that sequence on the same DNA strand specifically. Nicking endonuclease can introduce nicks and gaps at specific sites on DNA, and they have been used to prepare DNA substrates for *in vitro* mismatch repair (6). Nicking endonucleases also have potential applications in strand displacement amplification. Strand displacement amplification is a DNA amplification technology that can achieve 10^6-fold amplification in 30 min without thermocycling (7). In a strand displacement amplification reaction, the template DNA first is nicked by a restriction enzyme, and then the nick is extended by a DNA polymerase, displacing the parental DNA strand (7).

Thus far there are more than 3,000 restriction enzymes identified, among which only ≈200 different specificities have been demonstrated (8). Efforts to engineer novel endonucleases have had little success. *Fok*I is a type IIs restriction enzyme that exhibits a bipartite nature; it has an N-terminal DNA recognition domain and a C-terminal DNA cleavage domain (9, 10). The modular nature of *Fok*I led to the invention of several enzymes with new specificities (11). Fusion of the Ubx homeodomain to the *Fok*I cleavage domain yielded an enzyme that cleaves on both sides of the Ubx recognition site (12). Similar approaches have been used to engineer enzymes that can cleave near Z-DNA (13), DNA-RNA hybrids (14), and Gal4 recognition sites (15). However, three major drawbacks are associated with such chimeric enzymes. First, the chimeric enzymes cleave at multiple sites on both sides of the recognition sequence rather than at a unique location; therefore, the specificity is much relaxed. Second, the enzymatic cleavage activity of the chimeric enzymes is very low. Third, the selectivity (difference in activity at the correct site and at a site with one mismatch) is much reduced compared with that of a restriction enzyme.

Converting the type II restriction endonuclease *Eco*RV into a nicking endonuclease has been reported by combining a subunit with an inactive catalytic center with a subunit with a defect in the DNA-binding site (16). Although the engineered *Eco*RV heterodimers introduce nicks into DNA specifically within the *Eco*RV recognition sequence, they nick both DNA strands without discrimination because of the palindromic nature of its recognition sequence (16). Thus the heterodimeric *Eco*RV is not a strand-specific nicking endonuclease; its sequential cleavages can lead to double-stranded breakage (16).

Although more than 200 sequence specificities can be cleaved by more than 3,000 characterized restriction enzymes (8), only one DNA sequence, 5′-GAGTC-3′, can be nicked specifically by two naturally existing nicking endonucleases, N.*Bst*SEI (4) and N.*Bst*NBI (5). More nicking endonucleases with new sequence specificities are desired. For example, N.*Bst*NBI or N.*Bst*SEI could not be used in a strand displacement amplification reaction if there is an internal 5′-GAGTC-3′ sequence within the target DNA, because the template DNA strand would be broken by both enzymes. Our previous study showed that the nicking endonuclease N.*Bst*NBI is a naturally mutated type IIs endonuclease with diminished double-stranded cleavage activity (17). Thus far only two strand-specific nicking endonucleases have been reported, indicating that the natural occurrence of such nicking enzymes may be quite limited, although the lack of simple detection methods may be another factor. Here we describe an approach to generating strand-specific nicking endonuclease from existing type IIs restriction enzymes. Type IIs

This paper was submitted directly (Track II) to the PNAS office.

*To whom reprint requests should be addressed. E-mail kong@neb.com.

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.

www.pnas.org/cgi/doi/10.1073/pnas.241215698

enzymes recognize nonpalindromic sequences and thus cleave two DNA strands asymmetrically. By using a domain-swapping strategy we have obtained a nicking endonuclease, N.*Alw*I. It derives from *Alw*I, a type IIs endonuclease that recognizes 5′-GGATC-3′ and cleaves four and five bases outside the recognition sequence. N.*Alw*I recognizes the same sequence but just nicks the top strand specifically. Biochemical studies provide a new insight into understanding the cleavage reactions catalyzed by type IIs restriction endonucleases.

Materials and Methods

Reagents. *Taq* DNA polymerase was purchased from Perkin–Elmer. [^{33}P]dATP was from NEN. Chromatographic columns including Heparin Hyper D, Source Q, Source S, Hyper TSK, and Sephacryl-100/300 were purchased from Amersham Pharmacia. All other reagents including restriction enzymes, *Escherichia coli* strains, and plasmids were obtained from New England Biolabs.

Generation of N.*Alw*I by a Two-Step PCR Approach. The 5′ end of the *Alw*I endonuclease gene was amplified from plasmid pLT7K-*Alw*I, which contains the entire *Alw*I endonuclease gene, by using primers P1 (5′-TTACCGGTAAGGAGGTGATCTAATGAG-CACGTGGCTTCTTGGAA-3′) and P2 (5′-TTCACCAA-GAACAATAAAGTCTTCATACTCAAAGATCACATCAG-3′). The 3′ end of the N.*Bst*NBI endonuclease gene was amplified from plasmid pNbstnbi, which contains the complete N.*Bst*NBI endonuclease gene, by using primers P3 (5′-GACTTTATTGT-TCTTGGTGAA-3′) and P4 (5′-TTCTCGAGTTAAAACCT-TACCTCCTTGTCAACAA-3′). The two PCR products then were mixed at an equal molar ratio and amplified by using primers P1 and P4 (Fig. 3*A*, Swapping Construct 2). The 1,710-bp PCR product was digested with *Age*I and *Xho*I and cloned into plasmid pLT7K (18) to obtain plasmid pAB2.

PCR-Mediated Mutagenesis and Activity Analysis. By using the PCR-mediated mutagenesis approach (19), four *Alw*I mutants were obtained. *Alw*I-D420A, *Alw*I-E433A, *Alw*I-L502A-L504A, and *Alw*I-L504A-F507A each have one or two point mutations in which individual residues were replaced with alanine. Each of the four mutant plasmids was transformed into *E. coli* strain ER2566 (New England Biolabs) for expression. After induction in the presence of 0.3 mM isopropyl β-D-thiogalactoside for 3 h, cells were harvested and sonicated in a buffer containing 20 mM Tris, pH 7.4, 50 mM NaCl, 0.1 mM EDTA, and 10 mM β-mercaptoethanol. Various amounts of crude protein extract were incubated with 1 μg of lambda DNA (dam$^-$) in a 50-μl reaction for 20 min, mixed with 1× DNA loading dye supplemented with 1% SDS, and separated on a 1% agarose gel.

Expression and Purification of N.*Alw*I and *Alw*I. The *E. coli* strain ER2566 containing pAB2 was grown at 37°C until the OD_{600} reached 0.8. The expression of N.*Alw*I protein was induced in the presence of 0.3 mM isopropyl β-D-thiogalactoside at 30°C for 4 h. After induction, cells were spun down, sonicated in a sonication buffer (10 mM Tris, pH 7.4, 50 mM NaCl, 0.1 mM EDTA, and 10 mM β-mercaptoethanol), and centrifuged at 15,000 × *g* for 20 min. The supernatant fraction was passed through Heparin Hyper D, Source Q, and Heparin TSK. The loading buffer for chromatography contained 20 mM Tris, pH 7.4, 0.1 mM EDTA, 50 mM NaCl, and 1 mM DTT. The NaCl gradient ranged from 50 mM to 1 M. The efficiency of protein purification was monitored enzymatically. ER2566 containing pLT7K-*Alw*I was grown and induced in the same way. *Alw*I was purified by using four chromatographic columns: Heparin Hyper D, Source Q, Source S, and Hyper TSK.

Construction of Plasmids pAC0 and pAC1 and pAC2. Plasmid pACYC184 was digested with *Sfc*I and *Ava*I, filled in with Klenow enzyme, and self-ligated to give plasmid pAC1 (containing one *Alw*I recognition sequence, 5′-GGATC-3′, and three 5′-GAGTC-3′ sequences). Plasmid pACYC184 was digested with *Sfc*I and *Bsa*BI, filled in with Klenow enzyme, and self-ligated to give plasmid pAC0 (containing no 5′-GGATC-3′ sequence but two 5′-GAGTC-3′ sequences). Plasmid pACYC184 was digested with *Sfc*I and *Bgl*I (partial digestion), filled in with Klenow enzyme, and self-ligated to give plasmid pAC2 (containing two 5′-GGATC-3′ sites).

Construction of Oligonucleotide-Derived DNA Duplex. Four oligonucleotides, alw1f (5′-AATTCCGAAGGATGAGAGTCAGG-GGATCATGATTTCCGAACTATCACGAAT-3′), alw1r (5′-ATTCGTGATAGTTCGGAAATCATGATCCCCTGACTCT-CATCCTTCGGAATT-3′), alw0f (5′-AATTCCGAAGGAT-GAGAGTCAGGGGATTATGATTTCCGAACTATCAC-GAAT-3′), and alw0r (5′-ATTCGTGATAGTTCGGAAAT-CATAATCCCCTGACTCTCATCCTTCGGAATT-3′) were synthesized either with nonlabeled ends or with fluorescein labeling at both ends. Fluorescein-labeled oligonucleotides were annealed with their complementary nonlabeled oligonucleotides, i.e., alw1f paired with alw1r and alw0f paired with alw0r.

Mapping the Cleavage Site of N.*Alw*I. Manual sequencing was done by using the T7 Sequenase kit V2.0 (Amersham Pharmacia). pAC1 (4 μg) was used in the following six parallel reactions in the presence of the primer alw15f (5′-CACGGGGCCTGCCAC-CATA-3′). Four reactions were sequencing reactions with the addition of one of the four dideoxynucleotides (20). Two reactions were extended fully with the addition of an extension mix, heat-inactivated at 70°C for 20 min, and then digested with either N.*Alw*I (1.4 ng) or *Alw*I (1 unit; ref. 21). To map the cleavage site on the complementary strand, six more reactions were carried out in the presence of the primer alw16r (5′-AACGGT-TAGCGCTTCGTTA-3′). The reaction products were separated on an 8% Quick Point sequencing gel (Invitrogen) and detected by autoradiography.

Determination of the Molecular Mass of *Alw*I and N.*Alw*I by Gel Filtration. Purified *Alw*I and N.*Alw*I were used in the gel filtration assay. The loading buffer contained 10 mM Tris, pH 7.4, 500 mM NaCl, 0.1 mM EDTA, and 10 mM β-mercaptoethanol. HiPrep 26/60 Sephacryl-100 and Sephacryl-300 (Amersham Pharmacia) were used to determine the molecular mass of N.*Alw*I and *Alw*I in the loading buffer. Protein standards (Amersham Pharmacia) including catalase (232 kDa) aldolase, BSA (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), ribonuclease A (13.7 kDa), and plus blue Dextran 2000 (2,000 kDa) were used to generate standard curves.

Results

***Alw*I and N.*Bst*NBI Share Sequence Similarity and Catalytic Residues.** The deduced polypeptide sequences of *Alw*I and N.*Bst*NBI share 28.7% sequence identity and 41.5% similarity (Fig. 1). Site-directed mutagenesis on N.*Bst*NBI had shown that Asp-456 and Glu-469 are essential for the cleavage activity of N.*Bst*NBI and are likely to be the catalytic residues (17). The corresponding aspartate and glutamate in *Alw*I are Asp-420 and Glu-433, respectively. Substitution of alanine for Asp-420 completely knocked out *Alw*I cleavage activity; changing Glu-433 to alanine reduced the cleavage activity over 100-fold (Fig. 2). The results indicate that Asp-420 and Glu-433 in *Alw*I are critical for the cleavage activity and suggest that the catalytic residues are conserved between *Alw*I and N.*Bst*NBI.

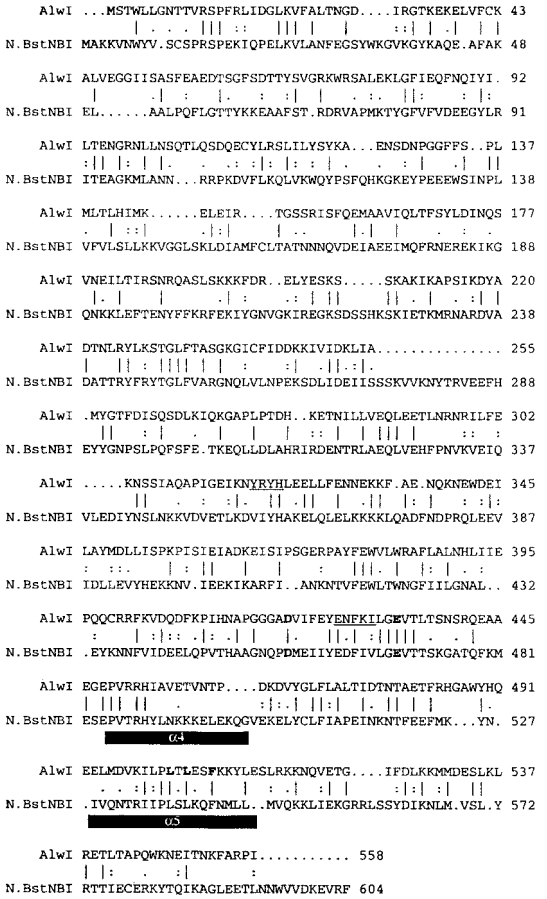

Fig. 1. Sequence comparison of *Alw*I and N.*Bst*NBI. The GCG GAP program (gap weight = 8; length weight = 2) was used to compare the protein sequences of *Alw*I (top) and N.*Bst*NBI (bottom). The residues mutated into alanines are shown in bold face. The two domain-swapping points are shown immediately after the underlined residues in *Alw*I. Putative dimerization regions in *Alw*I, which correspond to α4 and α5 helixes in *Fok*I, are indicated by black bars.

Converting *Alw*I into a Strand-Specific Nicking Enzyme. *Fok*I crystallization studies showed that *Fok*I dimerization is mediated primarily through two α-helixes, α4 and α5, located less than 10 amino acid residues away from its catalytic site PD—DTK (22). Studies on N.*Bst*NBI and the related type IIs endonucleases *Mly*I and *Ple*I suggested that dimerization is essential also for double-stranded DNA cleavage of *Mly*I and *Ple*I (17). Based on these results, we reasoned that if the dimerization function of a type IIs restriction enzyme were eliminated by mutation, it might become a nicking enzyme. Three amino acid residues were chosen to be mutated because they are located in the region

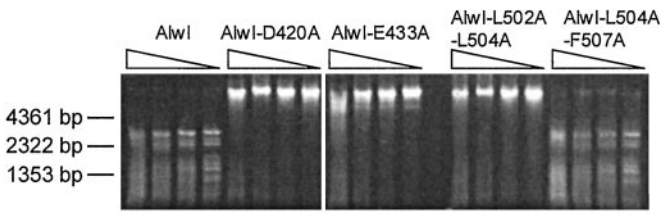

Fig. 2. Activity analysis of *Alw*I mutants. The *Alw*I mutants were expressed in *E. coli* strain ER2566. Crude cell extracts were incubated with 1 μg of lambda dam$^-$ DNA for 20 min, separated on a 1% agarose gel, and stained with ethidium bromide. Serial 2-fold dilutions were analyzed for each mutant.

corresponding to the α5-helix of *Fok*I, which is essential for dimerization (Fig. 1). The two corresponding α-helixes in *Alw*I were determined based on secondary structure analysis and sequence alignment with *Fok*I. To inactivate the dimerization function of *Alw*I, two constructs with double mutations (L504A-F507A and L502A-L504A) were made in the putative dimerization domain. One construct, *Alw*I-L504A-F507A, showed cleavage activity comparable to that of the wild-type *Alw*I, and the other double mutation, L502A-L504A, completely abolished the *Alw*I activity (Fig. 2). In addition, no nicking activity was observed based on the digestion result of lambda DNA (Fig. 2). Thus, these mutations in the putative dimerization domain failed to generate a nicking *Alw*I endonuclease.

Previous studies suggested that the nonfunctional dimerization domain of N.*Bst*NBI may have caused the great reduction in its double-strand cleavage activity and therefore resulted in a naturally existing nicking enzyme (17). Taking into account the similarity between *Alw*I and N.*Bst*NBI at the primary sequence level (Fig. 1) and at the level of secondary structure (data not shown), a domain-swapping strategy seemed feasible to engineer an *Alw*I-derived nicking enzyme by swapping the cleavage and/or dimerization domains of *Alw*I with the corresponding domains from N.*Bst*NBI. This chimeric enzyme might combine the substrate specificity of *Alw*I and the cleavage characteristic of the nicking enzyme N.*Bst*NBI.

One construct was made by fusing the putative DNA recognition domain of *Alw*I (located in the N-terminal half, residues 1–320) to the C-terminal region of N.*Bst*NBI (including both the putative catalytic and dimerization domains, residues 361–604; Fig. 3*A*). However, this chimeric protein (Fig. 3*A*, Swapping Construct 1) exhibited no detectable cleavage activity (data not shown). Swapping at this junction might have disrupted protein folding and resulted in a completely inactive enzyme. The other chimeric enzyme, as shown in Fig. 3*A*, Swapping Construct 2, was constructed by fusing a longer N-terminal part of *Alw*I (residues 1–430) including both its putative DNA-binding domain and catalytic center to the putative dimerization domain of N.*Bst*NBI at the extreme C-terminal end (residues 467–604). This chimeric endonuclease, containing the dimerization domain from N.*Bst*NBI, showed DNA nicking activity and was named N.*Alw*I.

Characterization of the DNA Cleavage Activity of N.*Alw*I. N.*Alw*I was expressed and purified to ≈80% homogeneity (see *Materials and Methods*). On SDS-PAGE gels, it migrated as a 65-kDa protein (Fig. 3*B*), which is consistent with the expected molecular mass.

The nicking activity of N.*Alw*I was examined first by using plasmid DNAs. The supercoiled form of an undigested plasmid can be converted into a nicked closed circular form, when one strand is cleaved by a nicking endonuclease, or into a linear form, when both strands are cleaved in proximity by a restriction enzyme. To test whether the nicking activity of N.*Alw*I is sequence-dependent, two plasmids were constructed: pAC1 with one *Alw*I recognition sequence and pAC0 with no *Alw*I recognition sequence. The plasmids pAC0 and pAC1 contain two and three N.*Bst*NBI recognition sites, respectively. As expected, the supercoiled pAC0 (Fig. 4*A*, lane 1) was not digested by *Alw*I (lane 2) and N.*Alw*I (lane 3) but was nicked by N.*Bst*NBI (lane 4). When plasmid pAC1 (lane 5) was incubated with the wild-type *Alw*I, the supercoiled form was converted into the linear form (lane 6). In contrast, when pAC1 was incubated with the engineered N.*Alw*I, the supercoiled form was converted into a nicked closed circular form (lane 7), which migrated at the same position as the relaxed form produced by N.*Bst*NBI (lane 8). These results demonstrate that the nicking activity of N.*Alw*I depends on the presence of the sequence 5′-GGATC-3′.

To test whether the nicking activity of N.*Alw*I is strand-specific, double-stranded oligodeoxyribonucleotides containing

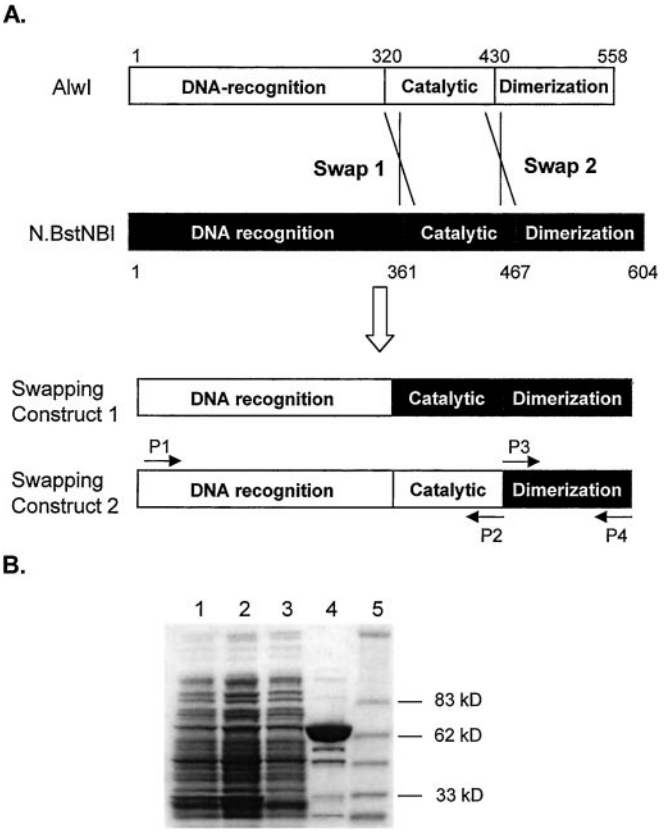

Fig. 3. Production of N.*Alw*I. (*A*) Schematic diagram of domain swapping in engineering N.*Alw*I. (*B*) Purification of N.*Alw*I. After induction, ER2566 cells were spun down and sonicated. The crude cell extract (lane 1) was centrifuged at 15,000 × *g* for 20 min to separate the supernatant fraction (lane 2) from the insoluble pellet (lane 3). The supernatant fraction was passed through Heparin Hyper D, Source Q, and Heparin TSK, giving the purified N.*Alw*I (lane 4).

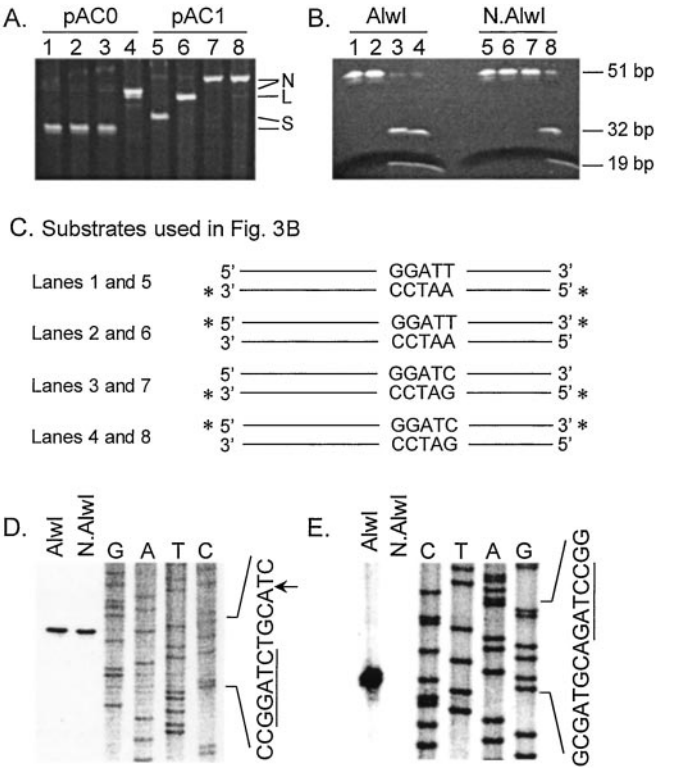

Fig. 4. Characterization of DNA cleavage activity of N.*Alw*I. (*A*) N.*Alw*I cleavage activity on plasmid DNA. Plasmid DNA (≈200 ng) was incubated in a 15-μl digestion reaction for 20 min in the presence of *Alw*I (2 units), N.*Bst*NBI (10 units), or N.*Alw*I (3 ng). Reactions were fractionated on a 1% agarose gel and stained with ethidium bromide. Lane 1, supercoiled plasmid pAC0; lane 2, pAC0 digested with *Alw*I; lane 3, pAC0 digested with N.*Alw*I; lane 4, pAC0 digested with N.*Bst*NBI; lane 5, supercoiled plasmid pAC1; lane 6, pAC1 digested with *Alw*I; lane 7, pAC1 digested with N.*Alw*I; lane 8, pAC1 digested with N.*Bst*NBI. S, supercoiled form of DNA; L, linearized form of DNA; N, nicked form of DNA. (*B*) N.*Alw*I cleavage activity on double-stranded oligodeoxyribonucleotides. Duplex substrates were prepared by annealing a fluorescein-labeled oligonucleotide with its nonlabeled complementary sequence (see *Materials and Methods*). Double-stranded duplex (≈50 pmol) was incubated with 100 ng of *Alw*I or 1 μg of N.*Alw*I in a 10:l reaction at 37°C for 1 h. Reactions were stopped by adding equal volume of a stop dye (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol FF). Approximately 1.4 μl of the reaction were separated on an 8% Quick Point sequencing gel (Invitrogen) and DNA bands were detected under a UV light. Detailed descriptions of lanes 1–8 are shown in *C*. (*C*) Schematic diagram of DNA substrates used in *B*. The pentamer (bases 24–28) sequence in the top strands indicates a wild-type recognition sequence (5′-GGATC-3′) or altered sequence (5′-GGATT-3′). Asterisks (*) indicate the fluorescein labels. (*D*) Mapping cleavage sites by modified sequencing reactions. pAC1 (4 μg), containing a single 5′-GGATC-3′ sequence, was used in manual sequencing reactions primed by either *Alw*15f (see *Materials and Methods*). Two additional reactions of each primer were extended fully without the addition of dideoxynucleotides and then digested with either *Alw*I (1 unit) or N.*Alw*I (1.4 ng). Reactions were fractionated on an 8% polyacrylamide gel and detected by autoradiography. (*E*) Identical to *D* except that the primer alw16r was used.

one or no recognition sequence were digested with N.*Alw*I and *Alw*I (Fig. 4*B*). The double-stranded DNA substrates used in the digestion reactions were prepared by annealing fluorescein-labeled oligonucleotides with their nonlabeled complementary oligonucleotides (Fig. 4*C*). Neither *Alw*I nor N.*Alw*I showed endonuclease activity on nonspecific substrates containing one base substitution in the *Alw*I recognition sequence 5′-GGATC-3′ (Fig. 4*B*, lanes 1, 2, 5, and 6). DNA cleavage activity was observed in the specific substrates. As expected, the wild-type *Alw*I cleaved the labeled bottom strand into 33- and 18-bp fragments (Fig. 4*B*, lane 3) and the labeled top strand into 32- and 19-bp fragments (Fig. 4*B*, lane 4). However, N.*Alw*I only cleaved the labeled top strand containing the sequence 5′-GGATC-3′ into the expected 32- and 19-bp fragments (lane 8) but did not digest the labeled bottom strand (lane 7). These results provide further evidence that N.*Alw*I is a sequence-dependent (requiring 5′-GGATC-3′) and strand-specific nicking endonuclease.

To map the cleavage site of N.*Alw*I precisely, the N.*Alw*I cleavage products were compared with sequencing ladders in a modified sequence reaction (see *Materials and Methods*). As illustrated in Fig. 4*D*, both *Alw*I and N.*Alw*I cut the top strand; the cleavage site was located four bases downstream of the recognition site 5′-GGATC-3′ in both cases. The bottom strand was cleaved by wild-type *Alw*I but was not cut by N.*Alw*I (Fig. 4*E*). Thus, the nicking site of N.*Alw*I is located on the top strand, four bases away from the recognition sequence (5′-GGATC-NNNN-3′).

Analysis of the Mechanism of N.*Alw*I Nicking Activity. Our rationale for engineering nicking endonucleases by using domain swapping was based on the following observations. First, dimerization is responsible for the double-stranded cleavage by type IIs enzymes (22–24). Second, the dimerization function is impaired in the nicking enzyme N.*Bst*NBI (17). This implies that a chimeric nicking enzyme containing the dimerization domain of N.*Bst*NBI would not form a dimer and therefore would become a nicking enzyme. To test this hypothesis, gel filtration was used to compare the molecular mass of wild-type *Alw*I and the

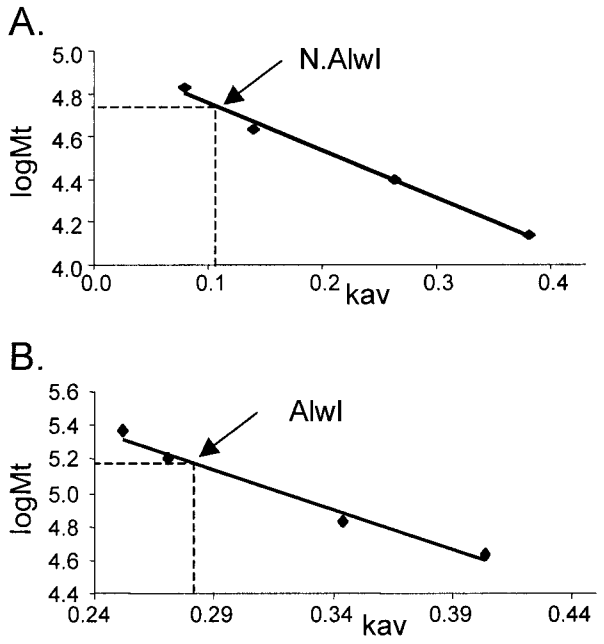

Fig. 5. Measurement of the molecular mass of *Alw*I and N.*Alw*I by gel filtration. (*A*) HiPrep 26/60 Sephacryl-100 (with a total bed volume $V_t = 318$ ml) was used to determine the molecular mass of N.*Alw*I under native conditions. The elution volume of blue Dextran 2000 (2,000 kDa) was used to determine the void volume (V_o). The elution volumes (V_e) of protein standards BSA (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa) were measured. The partition coefficient (K_{av}) is calculated as $(V_e - V_o)/(V_t - V_o)$. The standard curve was obtained by plotting K_{av} against the log of the molecular mass (logMr). The K_{av} of N.*Alw*I was obtained (0.11), and its molecular weight was determined by using the standard curve. (*B*) HiPrep 26/60 Sephacryl-300 was used to determine the molecular weight of *Alw*I following the same procedure as N.*Alw*I. Protein standards used were catalase (232 kDa), aldolase (158 kDa), BSA (67 kDa), and ovalbumin (43 kDa). The K_{av} of *Alw*I was 0.28.

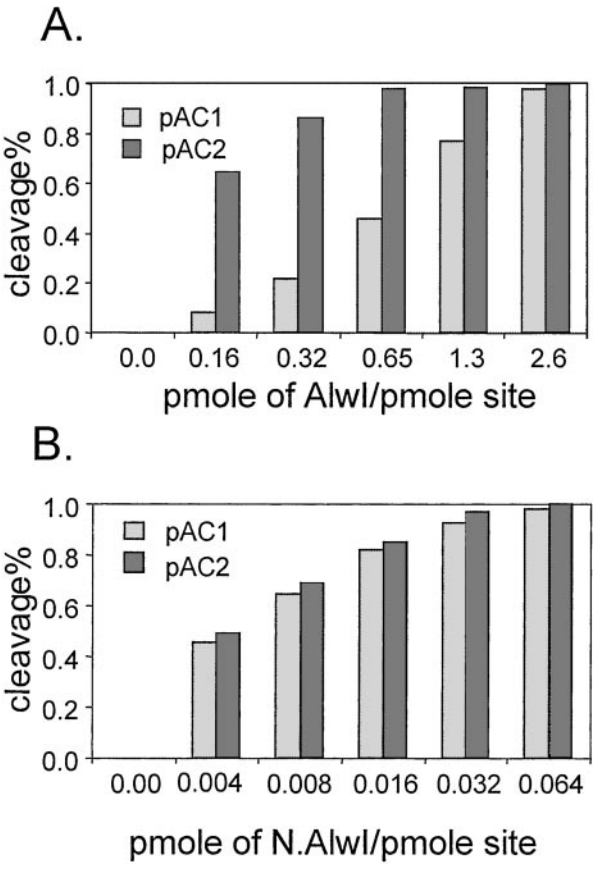

Fig. 6. Substrate preference of *Alw*I and N.*Alw*I. Equal amounts of pAC1 (125 ng) or pAC2 (125 ng) were incubated with various amounts of either *Alw*I or N.*Alw*I at 37°C for 1 h. A serial 2-fold dilution was carried out for both *Alw*I (starting with 20 ng) and N.*Alw*I (starting with 0.5 ng). Digestion reactions were fractionated on a 1% agarose gel. DNA bands were quantified by using IMAGEJ (29). The percentage of complete DNA cleavage was plotted against the enzyme amount (per pmol of recognition sites).

engineered N.*Alw*I under native conditions. Purified N.*Alw*I (Fig. 4*B*) was passed through a Sephacryl-100 column. Its molecular mass was calculated to be ≈54.6 kDa (Fig. 5*A*), suggesting that N.*Alw*I ($M_r = 65{,}448$) is a monomer in solution. On the other hand, *Alw*I (≈95% homogeneity) remained in the exclusion fraction after passing through Sephacryl-100 (data not shown), indicating an apparent molecular mass of more than 100 kDa under native conditions. On a Sephacryl-300 column, *Alw*I showed a molecular mass of 146 kDa (Fig. 5*B*), indicating that the wild-type *Alw*I ($M_r = 63{,}894$) exists as a dimer under native conditions.

To obtain more evidence that *Alw*I cleaves DNA as a dimer and N.*Alw*I nicks DNA as a monomer, plasmid pAC2 containing two recognition sites was constructed (see *Materials and Methods*). Restriction endonucleases that cleave DNA in a multimeric form prefer DNA substrates containing multiple recognition sites (25, 26). As shown in Fig. 6*A*, *Alw*I had a higher specific activity toward pAC2 than pAC1, suggesting it may cleave DNA as dimers. This property was more dramatic at lower enzyme concentration; the activity of *Alw*I on pAC2 is ≈7-fold higher than its activity on pAC1 at the enzyme/DNA ratio of 0.016 (Fig. 6*A*). In contrast, N.*Alw*I did not discriminate substrates containing one or two sites (Fig. 6*B*). This was also the case for both *Alw*I and N.*Alw*I when plasmid pUC19 (containing 10 *Alw*I sites) was used as substrate (data not shown).

As shown in Fig. 6*A*, 2.6 pmol of *Alw*I is required to achieve complete digestion of one pmol of GGATC sites. In contrast, 0.064 pmol of N.*Alw*I was sufficient to nick one pmol of 5′-GGATC-3′ sites (Fig. 6*B*). Considering the fact that one double strand cleavage takes two nicking events, N.*Alw*I has ≈20-fold higher specific activity than the wild-type *Alw*I.

Discussion

Restriction endonucleases have provided molecular biologists with a powerful tool to manipulate DNA. Although more than 200 sequence specificities can be cleaved by more than 3,000 characterized restriction enzymes (8), only one DNA sequence, 5′-GAGTC-3′, can be nicked specifically by two naturally existing nicking endonucleases, N.*Bst*SEI (4) and N.*Bst*NBI (5). More nicking endonucleases with new sequence specificities will be useful in DNA engineering.

We have demonstrated in this paper that it is possible to create nicking endonucleases with new sequence specificities by a protein-engineering approach. The engineering nicking enzyme N.*Alw*I is useful. It has a stringent selectivity of the sequence 5′-GGATC-3′. A single mutation in the recognition sequence such as changing 5′-GGATC-3′ to 5′-GGATT-3′ abolished N.*Alw*I cleavage activity (Fig. 4*B*). In addition, N.*Alw*I only cleaves the top strand because of the asymmetric nature of its recognition sequence. Furthermore, the engineered N.*Alw*I is a very active enzyme. Approximately 50% of 1 pmol of DNA could be cleaved by 0.004 pmol of N.*Alw*I in 1 h. This rate corresponds to a turnover rate of ≈120 cleavage events per hour ($1/0.004 \times 50\% = 120$; Fig. 6*B*). Indeed, N.*Alw*I showed ≈20-fold higher

activity than that of the wild-type *Alw*I toward plasmid pAC1 (Fig. 6). In summary, the engineered N.*Alw*I is an active, sequence-dependent, and strand-specific nicking endonuclease. These features distinguish N.*Alw*I from other previously engineered endonucleases and make possible its commercial application. Similar domain-swapping strategies may be applicable to engineering other type IIs restriction enzymes into nicking restriction endonucleases.

The cleavage mechanism of *Alw*I seems to be different from that of *Fok*I, which is the only well studied type IIs restriction enzyme. Mutagenesis and crystallization studies suggested that *Fok*I has a single catalytic center (10, 27). Although *Fok*I exists as a monomer in solution (28), biochemical analysis shows that *Fok*I dimerizes in the presence of divalent metal and cognate DNA, and the formation of dimeric enzyme/DNA complex is essential for activation of cleavage (23). The current working model for *Fok*I is that the dimerization interface is sequestered by the DNA-binding domain in solution (22, 24). In the presence of its recognition sequence and divalent cation, the monomeric *Fok*I binds to its recognition sequence and dimerizes with another DNA-bound *Fok*I monomer (23). The dimerization triggers a conformational change, which leads to double strand cleavage (23, 24). Because *Alw*I exists as a dimer in solution (Fig. 5*B*), it may interact with DNA in a manner different than that of *Fok*I.

*Alw*I probably binds to its DNA substrate as a dimer and interacts with two DNA recognition sites in a cooperative manner. This speculation is supported by results from a comparative DNA cleavage study using substrates containing either one or two recognition sites (Fig. 6). We found that additional sites on the DNA substrates stimulated the cleavage rates of *Alw*I (Fig. 6*A*), indicating interaction among the binding sites of *Alw*I. When the two sites are located in the same DNA molecule such as pAC2, the interaction is a bimolecular reaction involving a dimeric *Alw*I and one pAC2 molecule with two sites, which is more efficient than the trimolecular reaction involving one dimeric *Alw*I and two pAC1 molecules (Fig. 6*A*). Once the dimeric *Alw*I binds to two recognition sites, the strand containing the 5′-GGATC-3′ sequence is probably cleaved by the *Alw*I monomer, which binds to that site, and its complementary strand probably is cleaved by the second *Alw*I monomer, which dimerizes with the first monomer. Support for this proposal comes from the observation that when the dimerization domain of *Alw*I was swapped with the nonfunctional domain from N.*Bst*NBI, the resulting N.*Alw*I exists as a monomer in solution (Fig. 5*A*) and only cleaves the strand containing the 5′-GGATC-3′ sequence (Fig. 4). In addition, the monomeric N.*Alw*I shows the same activity toward substrates with either one or two recognition sites, suggesting that N.*Alw*I only interacts with one recognition site.

In conclusion, the oligomerization state of the type IIs endonuclease *Alw*I have been changed from dimer to monomer. This alters the *Alw*I-catalyzed DNA cleavage reaction; the monomeric N.*Alw*I cleaves one DNA strand, whereas the dimeric *Alw*I cleaves both DNA strands. Thus, we have shown that changing the oligomerization state of an enzyme changes its function.

We are very grateful to Geoffrey Wilson for sharing the unpublished sequence and construct of *Alw*I endonuclease. We thank Lauren Higgins for constructing pNbstnbi and Janos Posfai for assistance in sequence analysis. We thank Rick Morgan and Andy Gardner for help in mapping the cleavage sites of *Alw*I and N.*Alw*I. We thank Caroline Besnier, Lauren Higgins, Thomas Evans, and Francine Perler for helpful discussions. We thank Aneel Aggarwal for sharing unpublished *Fok*I results. We are very grateful to Richard Roberts, Elisabeth Raleigh, and Charles Cantor for critical reading of this manuscript. We thank Donald Comb for support. Part of this work was supported by National Institutes of Health Grant GM60057 (to H.K.).

1. Heitman, J. (1993) *Genetic Engineering* (Plenum, New York), Vol. 15, pp. 57–108.
2. Higashitani, A., Greenstein, D., Hirokawa, H., Asano, S. & Horiuchi, K. (1994) *J. Mol. Biol.* **237,** 388–400.
3. Malik, H. S. & Henikoff, S. (2000) *Trends Biochem. Sci.* **25,** 414–418.
4. Abdurashitov, M. A., Belichenko, O. A., Shevchenko, A. V. & Degtyarev, S. K. (1996) *Mol. Biol.* **30,** 1261–1267.
5. Morgan, R. D., Calvet, C., Demeter, M., Agra, R. & Kong, H. (2000) *Biol. Chem.* **381,** 1123–1125.
6. Wang, H. & Hays, J. B. (2000) *Mol. Biotechnol.* **15,** 97–104.
7. Walker, G. T., Little, M. C., Nadeau, J. G. & Shank, D. D. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 392–396.
8. Roberts, R. J. & Macelis, D. (2001) *Nucleic Acids Res.* **29,** 268–269.
9. Li, L., Wu, L. P. & Chandrasegaran, S. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4275–4279.
10. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I. & Aggarwal, A. K. (1997) *Nature (London)* **388,** 97–100.
11. Chandrasegaran, S. & Smith, J. (1999) *Biol. Chem.* **380,** 841–848.
12. Kim, Y. & Chandrasegaran, S. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 883–887.
13. Kim, Y. G., Kim, P. S., Herbert, A. & Rich, A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 12875–12879.
14. Kim, Y. G., Shi, Y., Berg, M. & Chandrasegaran, S. (1997) *Gene* **203,** 43–49.
15. Kim, Y. G., Smith, J., Durgesha, M. & Chandrasegaran, S. (1998) *Biol. Chem.* **379,** 489–495.
16. Stahl, F., Wende, W., Jeltsch, A. & Pingoud, A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 6175–6180.
17. Higgins, L. S., Besnier, C. & Kong, H. (2001) *Nucleic Acids Res.* **29,** 2492–2501.
18. Kong, H., Lin, L. F., Porter, N., Stickel, S., Byrd, D., Posfai, J. & Roberts, R. J. (2000) *Nucleic Acids Res.* **28,** 3216–3223.
19. Morrison, H. G. & Desrosiers, R. C. (1993) *BioTechniques* **14,** 454–457.
20. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
21. Brown, N. L., Hutchison, C. A. & Smith, M. (1980) *J. Mol. Biol.* **140,** 143–148.
22. Wah, D. A., Bitinaite, J., Schildkraut, I. & Aggarwal, A. K. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 10564–10569.
23. Vanamee, E. S., Santagata, S. & Aggarwal, A. K. (2001) *J. Mol. Biol.* **309,** 69–78.
24. Bitinaite, J., Wah, D. A., Aggarwal, A. K. & Schildkraut, I. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 10570–10575.
25. Nobbs, T. J., Halford, S. E. (1995) *J. Mol. Biol.* **252,** 399–411.
26. Kong, H. & Smith, C. L. (1998) *Biol. Chem.* **379,** 605–609.
27. Waugh, D. S. & Sauer, R. T. (1994) *J. Biol. Chem.* **269,** 12298–12303.
28. Kaczorowski, T., Skowron, P. & Podhajska, A. J. (1989) *Gene* **80,** 209–216.
29. Rasband, W. (2000) IMAGEJ (National Institutes of Health, Bethesda, MD), Version 1.20s.