J|A|C|S
ARTICLES

Published on Web 11/10/2009

Redesign of High-Affinity Nonspecific Nucleases with Altered Sequence Preference

Yi-Ting Wang,[†,‡,§] Jon D. Wright,[||] Lyudmila G. Doudeva,[†] Hua-Ci Jhang,[†] Carmay Lim,[*,||,‡,§] and Hanna S. Yuan[*,†,‡,#]

Institute of Molecular Biology, Institute of Biomedical Sciences, and Institute of Chemical Biology and Molecular Biophysics, Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan, R.O.C, Institute of Bioinformatics and Structural Biology and Department of Chemistry, National Tsing Hua University, HsinChu, Taiwan, R.O.C., and Institute of Biochemistry and Molecular Biology, National Taiwan University, Taipei, Taiwan, R.O.C

Received August 24, 2009; E-mail: hanna@sinica.edu.tw; carmay@gate.sinica.edu.tw

Abstract: It is of crucial importance to elucidate the underlying principles that govern the binding affinity and selectivity between proteins and DNA. Here we use the nuclease domain of Colicin E7 (nColE7) as a model system to generate redesigned nucleases with improved DNA-binding affinities. ColE7 is a bacterial toxin, bearing a nonspecific endonuclease domain with a preference for hydrolyzing DNA phosphodiester bonds at the 3′O-side after thymine and adenine; i.e., it prefers Thy and Ade at the −1 site. Using systematic computational screening, six nColE7 mutants were predicted to bind DNA with high affinity. Five of the redesigned single-point mutants were constructed and purified, and four mutants had a 3- to 5-fold higher DNA binding affinity than wild-type nColE7 as measured by fluorescence kinetic assays. Moreover, three of the designed mutants, D493N, D493Q, and D493R, digested DNA with an increased preference for guanine at +3 sites compared to the wild-type enzyme, as shown by DNA footprint assays. X-ray structure determination of the ColE7 mutant D493Q−DNA complex in conjunction with structural and free energy decomposition analyses provides a physical basis for the improved protein−DNA interactions: Replacing D493 at the protein−DNA interface with an amino acid residue that can maintain the native hydrogen bonds removes the unfavorable electrostatic repulsion between the negatively charged carboxylate and DNA phosphate groups. These results show that computational screening combined with biochemical, structural, and free energy analyses provide a useful means for generating redesigned nucleases with a higher DNA-binding affinity and altered sequence preferences in DNA cleavage.

Introduction

DNA-binding proteins are attractive targets for protein redesign because they play a central role in many biological events, from DNA replication, transcription, recombination and repair to chromosome assembly and apoptotic DNA degradation. Previous studies on protein−DNA complexes have elucidated some general trends in protein−DNA recognition: (1) sequence-specific DNA-binding proteins often interact with DNA bases using polar and charged amino acid side chains, whereas nonspecific DNA-binding proteins tend to contact DNA at backbone phosphates;[1,2] (2) a greater specificity is usually achieved via interactions in DNA major grooves rather than in minor grooves.[3] Attempts have been made to predict the DNA-binding site given the 1D protein sequence[4] or 3D structure[5,6] or to engineer a new DNA-binding protein with different sequence specificities.[7] However, no simple generic code for protein−DNA recognition has been found due to complicated interaction patterns involving not only hydrogen-bonding and nonbonded contacts but also conformational changes in the protein and DNA upon binding.[8]

Protein engineering of a new DNA-binding protein is usually achieved by indirect protein evolution strategies[9] or by computational protein design algorithms.[7] Indirect protein evolution strategies have been used mainly to alter the affinity and specificity of a DNA-binding protein on the basis of selected mutations or random screening methods such as phage display and two-hybrid screening. These studies have made variant site-specific endonucleases that target novel DNA sites.[10−12] On the other hand, computational protein design could create various novel DNA-

[†] Institute of Molecular Biology, Academia Sinica.
[‡] Taiwan International Graduate Program, Academia Sinica.
[§] National Tsing Hua University.
[||] Institute of Biomedical Sciences, Academia Sinica.
[#] National Taiwan University.

(1) Mandel-Gutfreund, Y.; Schueler, O.; Margalit, H. *J. Mol. Biol.* **1995**, *253*, 370–82.
(2) Hsia, K.-C.; Li, C.-L.; Yuan, H. S. *Curr. Opin. Struct. Biol.* **2005**, *15*, 126–134.
(3) Luscombe, N. M.; Thornton, J. M. *J. Mol. Biol.* **2002**, *320*, 991–1009.
(4) Sarai, A.; Kono, H. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 379–98.
(5) Chen, Y. C.; Wu, C. Y.; Lim, C. *Proteins: Struct., Funct., Bioinf.* **2007**, *67*, 671–680.
(6) Chen, Y. C.; Lim, C. *Nucleic Acids Res.* **2008**, *36*, 7078–7087.
(7) Lippow, S. M.; Tidor, B. *Curr. Opin. Biotechnol.* **2007**, *18*, 305–11.
(8) Pabo, C. O.; Nekludova, L. *J. Mol. Biol.* **2000**, *301*, 597–624.
(9) Arnold, F. H.; Wintrode, P. L.; Miyazaki, K.; Gershenson, A. *Trends Biochem. Sci.* **2001**, *26*, 100–6.
(10) Rebar, E. J.; Greisman, H. A.; Pabo, C. O. *Methods Enzymol.* **1996**, *267*, 129–49.
(11) Seligman, L. M.; Chisholm, K. M.; Chevalier, B. S.; Chadsey, M. S.; Edwards, S. T.; Savage, J. H.; Veillet, A. L. *Nucleic Acids Res.* **2002**, *30*, 3870–9.

binding proteins by tethering two unrelated domains together.[13−18] For example, fusion of the homing endonuclease domains of I-DmoI and I-CreI generated a highly specific engineered endonuclease,[14] while tethering site-specific zinc fingers to a nonspecific nuclease domain of restriction enzyme FokI created artificial site-specific nucleases.[13,15] Moreover, a mutated homing endonuclease I-MsoI, redesigned using an atomic-level force field, was shown to cleave DNA with altered specificity.[19]

These previous works focused on *site-specific* DNA-binding proteins and endonucleases that bind DNA with *high* affinity and specificity. *Nonspecific* DNA-binding proteins are less predictable as they are more flexible and bind to various DNA sequences with *low* affinity, posing a challenging problem for protein redesign. Herein, we present a systematic computational method to generate redesigned nonspecific nucleases with improved DNA-binding affinity. As a test system, we use Colicin E7 (ColE7), an *Escherichia coli* toxin that contains a nonspecific nuclease domain capable of cleaving phosphodiester linkages at many positions along the DNA with a preference for making nicks after thymine and adenine.[20] The crystal structures of the nuclease domain of ColE7 (nColE7), in the presence and absence of DNA,[21−23] show that nColE7 interacts mainly with the DNA backbone, with residues interacting in both minor and major grooves to clamp the phosphates (see Figure 1).

Based on the crystal structure of nColE7 complexed with an 8-bp DNA, systematic computational screening was used to identify candidate residues for protein engineering. Among the six candidates suggested by computational screening, five mutants were constructed and purified. Four of the five redesigned nColE7 mutants bound DNA with higher affinity, and three of the mutants cleaved DNA with altered sequence preference, as compared to the wild-type enzyme. To provide a structural basis for the observed higher DNA-binding affinity of the mutants, the crystal structure of one of the mutants, D493Q in complex with an 18-bp DNA, was determined. Comparative structural and free energy decomposition analyses further identified the physical basis for the mutant's high DNA-binding affinity. This study thus provides a systematic way to generate redesigned nonspecific endonucleases with improved DNA-binding affinity, while the structural and free energy analyses provide insight into the interactions between a nonspecific DNA-binding protein and DNA.

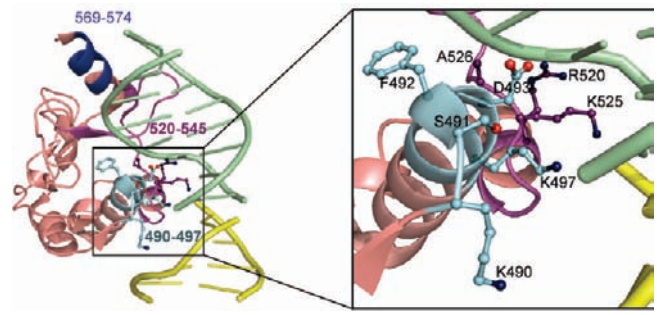

Figure 1. Suggested mutated residues in nColE7 based on computational screening method. Three groups of residues in the interface of nColE7−DNA were identified, residues 490−497(cyan), 520−545(purple), and 569−574 (blue). The interface residues that do not make any internal hydrogen bonds with nColE7 backbone atoms and the residues that were located in the DNA minor groove were disregarded. Eight residues were selected (shown in stick model) for further mutational analyses.

Results

Computational Redesign of nColE7 Mutants with Improved DNA-Binding Activity. To identify candidate residues for redesigning nucleases with better DNA-binding affinities and altered sequence cleavage preferences, the following computational strategy was employed. Using the crystal structure of nColE7 bound to an 8-bp DNA (PDB entry 1PT3, chain A), the solvent accessible surface area (SASA) of each residue in the presence and absence of DNA was calculated by MolMol.[24] 21 residues had a significant change in SASA and were analyzed by Whatif[25] for hydrogen bonds within the protein and at the protein−DNA interface. We used two criteria to select residues for in silico mutation: (1) those residues whose backbones do not make hydrogen bonds with other protein atoms were disregarded because their side chains might be too flexible, and (2) those residues that are located in the DNA minor groove were also disregarded because site-specific interactions tend to form at major grooves (see Introduction).[3] After screening out these residues, 8 residues (K490, S491, F492, D493, K497, R520, K525, A526) located in the DNA major groove were mutated in turn to arginine, asparagine, and glutamine, the three most frequently found residues at protein−DNA interfaces,[26,27] using a backbone-dependent side chain rotamer library implemented in Whatif,[25] and their hydrogen-bonding interactions were computed from these initial structures (see Figure 1). Of these mutants, D493R/Q/N, K497R, and A526R/Q/N (shaded in Table 1) appeared capable of making the same number of or more hydrogen bonds to the DNA as compared to those of wild-type nColE7 and were selected for further study.

In the second stage, to verify that the D493R/Q/N, K497R, and A526R/Q/N mutants could indeed form relatively stable internal and protein−DNA hydrogen bonds, the DNA-bound wild-type and mutated nColE7 were subjected to molecular dynamics (MD) simulation at a mean temperature of 300 K at pH 7 using the CHARMM34[28] program for 1.3 ns (see Methods). The simulation protocol adopted could maintain the

(12) Rosen, L. E.; Morrison, H. A.; Masri, S.; Brown, M. J.; Springstubb, B.; Sussman, D.; Stoddard, B. L.; Seligman, L. M. *Nucleic Acids Res.* **2006**, *34*, 4791–800.
(13) Smith, J.; Bibikova, M.; Whitby, F. G.; Reddy, A. R.; Chandrasegaran, S.; Carroll, D. *Nucleic Acids Res.* **2000**, *28*, 3361–9.
(14) Chevalier, B. S.; Kortemme, T.; Chadsey, M. S.; Baker, D.; Monnat, R. J.; Stoddard, B. L. *Mol. Cell* **2002**, *10*, 895–905.
(15) Urnov, F. D.; Miller, J. C.; Lee, Y.-L.; Beausejour, C. M.; Rock, J. M.; Augustus, S.; Jamieson, A. C.; Porteus, M. H.; Gregory, P. D.; Holmes, M. C. *Nature* **2005**, *435*, 646–651.
(16) Chan, S. H.; Bao, Y.; Ciszak, E.; Laget, S.; Xu, S. Y. *Nucleic Acids Res.* **2007**, *35*, 6238–48.
(17) Zhang, P. H.; Bao, Y. M.; Higgins, L.; Xu, S. Y. *Protein Eng., Des. Sel.* **2007**, *20*, 497–504.
(18) Lippow, S. M.; Aha, P. M.; Parker, M. H.; Blake, W. J.; Baynes, B. M.; Lipovsek, D. *Nucleic Acids Res.* **2009**, *37*, 3061–3073.
(19) Ashworth, J.; Havranek, J. J.; Duarte, C. M.; Sussman, D.; Monnat, R. J., Jr.; Stoddard, B. L.; Baker, D. *Nature* **2006**, *441*, 656–9.
(20) Wang, Y.-T.; Yang, W.-J.; Li, C.-L.; Doudeva, L. G.; Yuan, H. S. *Nucleic Acids Res.* **2007**, *35*, 584–594.
(21) Cheng, Y. S.; Hsia, K. C.; Doudeva, L. G.; Chak, K. F.; Yuan, H. S. *J. Mol. Biol.* **2002**, *324*, 227–36.
(22) Hsia, K. C.; Chak, K. F.; Liang, P. H.; Cheng, Y. S.; Ku, W. Y.; Yuan, H. S. *Structure* **2004**, *12*, 205–14.
(23) Doudeva, L. G.; Huang, H.; Hsia, K.-C.; Shi, Z.; Li, C.-L.; Shen, Y.; Yuan, H. S. *Protein Sci.* **2006**, *15*, 269–280.
(24) Koradi, R.; Billeter, M.; Wuthrich, K. *J. Mol. Graph.* **1996**, *14*, 51–55.
(25) Vriend, G. *J. Mol. Graph.* **1990**, *8*, 52–6, 29.
(26) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860–74.
(27) Luscombe, N. M.; Thornton, J. M. *J. Mol. Biol.* **2002**, *320*, 991–1009.
(28) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

Table 1. Predicted Hydrogen Bonds and Free Energy Changes in nColE7 Mutant−DNA Complexes

Suggested mutated residues	Internal Hydrogen Bond(s)[a]	Protein-DNA Hydrogen Bond(s)[b]	ΔΔG[c] (kcal/mol)
D493 (wild type)	✦	+	0
D493R	✦	++	−17.2 ± 0.3
D493Q	✦+	+	−7.2 ± 0.2
D493N	✦	+	−13.4 ± 0.2
K497 (wild type)	✦✦		0
K497R	✦✦	++	−14.4 ± 0.2
K497Q			
K497N			
A526 (wild type)	✦	✦	0
A526R	✦	✦+	−13.1 ± 0.2
A526Q	✦	✦	0.1 ± 0.2
A526N	✦+	✦+	−3.6 ± 0.2

[a]Each ✦ or + denotes a hydrogen bond within nColE7. '✦' denotes an interaction involving the protein backbone, whereas '+' denotes an interaction involving the protein side chain.

[b]Each ✦ or + denotes a hydrogen bond between nColE7 and DNA. '✦' denotes an interaction involving the protein backbone, whereas '+' denotes an interaction involving the side chain.

[c]ΔΔG is the free energy of mutant nColE7 binding to the 8-mer DNA, 5'-GCGATCGC-3', relative to that of the wild-type protein. The crystal structure of nColE7-DNA complex (PDB entry 1PT3, chain A) was used as the starting structure in the MD simulations (see text). The DNA-binding free energy of K497Q/N nColE7 was not computed, as it did not form any internal or protein-DNA hydrogen bonds, and was thus not considered to be a good candidate for redesigning a nuclease with better DNA-binding affinity.

overall protein structure, as evidenced by the reasonable root-mean-square deviation (rmsd) of the protein backbone atoms (1.8 ± 0.1 Å) and the DNA backbone atoms (1.5 ± 0.1 Å) in the wild-type complex simulation from those in the initial structure. It also retained all but one of the hydrogen bonds to the DNA found in the X-ray structure (see Supporting Table S1). The hydrogen bonds made by each mutant residue within nColE7 and between nColE7 and DNA were recalculated using the 1000 saved configurations from the last 100 ps of each trajectory (see Methods). As D493R/Q/N, K497R, and A526R/Q/N had one or more hydrogen bonds within nColE7 and at the nColE7−DNA interface (shaded in Table 1), they appear to be the most promising mutants.

To verify that the seven mutants could bind to DNA with high affinity, the free energy of each mutant binding to GCGATCGC (the DNA sequence in the wild-type complex X-ray structure, PDB entry 1PT3) relative to that of the wild-type protein was computed, as described in the Methods. The relative free energies in Table 1 indicate that, except for A526Q, the other nColE7 mutants (D493R, D493Q, D493N, K497R, A526R, and A526N) could apparently bind the 8-mer DNA more tightly than wild-type nColE7 (negative $\Delta\Delta G$). The computed free energy differences should *not* be considered as quantitative, since they were estimated using gas-phase free energies combined with an implicit solvent model, but they indicate whether the mutant protein can bind to the DNA sequence in the 1PT3 X-ray complex structure better than the wild-type nColE7.

Improved DNA Binding Affinity and Hydrolytic Activity. Since the computationally redesigned mutants were derived from nColE7 bound to an 8-mer DNA, 5′-GCGATCGC-3′, it was not known if they can bind to other DNA sequences with higher affinity. To test the computational redesign predictions, seven single-point nColE7 mutants were successfully constructed, expressed, and purified from *E. coli*. The five redesigned mutants, D493R, D493Q, D493N, K497R, and A526N, were chosen, as they were predicted to bind the 8-mer DNA tighter than the wild-type protein. The A526Q and K490R were chosen as negative controls, since the affinity of A526Q for the 8-mer DNA was predicted to be similar to wild-type nColE7, while K490R did not make any hydrogen bonding or van der Waals (vdW) contacts with the 8-mer DNA (see above). The DNA binding affinity and endonuclease activity of these mutants were measured by fluorescence resonance energy transfer (FRET), an assayed method developed previously, using a 16-bp double-stranded DNA substrate labeled respectively with the FAM (6-carboxyl-fluorescein) fluorophore and a TAMRA (6-carboxyl-tetramethylrhodamine) quencher at the 5′ end.[20] Note that the sequence of the 16-mer DNA used in the fluorescence kinetic assays differs from that of the 8-mer DNA used in the free energy calculations. The cleavage of the fluorophore-labeled DNA substrate by nColE7 resulted in an increase of fluorescence emission intensity, which was monitored by a

Table 2. DNA Binding Affinity, Catalytic Efficiency, and Cleavage Preference of nColE7 Proteins Measured by FRET and Footprint Assays

	Fluorescence Kinetic Assay (FRET)[a]			Cleavage preference[b] (%)							
				−1 site				+3 site			
	k_{cat} (x10^{-3}s^{-1})	K_m (nM)	k_{cat}/K_m (M^{-1}s^{-1})	T	A	G	C	T	A	G	C
Wild type	1.8±0.2	46.7±5	38,543.9 (100%)	42.1	29.5	12.2	17.1	25.1	25.8	26.3	22.8
D493N[c]	2.0±0.1	17.3±2	115,606.9 (299.9%)	**38.5**	**25.8**	**16.9**	**21.5**	**24.5**	**23.0**	**32.6**	**19.9**
D493Q	1.6±0.5	9.7±3	164,948.5 (427.9%)	**33.8**	**25.1**	**20.7**	**20.4**	**23.0**	**23.1**	**33.5**	**20.4**
D493R	1.7±0.5	9.6±2	177,083.3 (459.43%)	**33.6**	**25.2**	**20.6**	**20.6**	**22.5**	**22.7**	**34.6**	**20.2**
K497R	2.8±0.5	18.9±5	148,148.1 (384.4%)	42.1	26.8	11.1	20.0	25.3	25.6	26.2	22.9
A526N	1.8±0.1	43.2±4	41,666.7 (108.1%)	41.7	29.7	11.9	16.7	23.2	26.9	26.0	23.9
A526Q	1.7±0.2	46.9±5	36,247.3 (94.0%)	42.5	28.9	11.2	17.4	24.7	25.3	26.4	23.6
K490R	2.0±0.3	41.4±2	48,309.2 (125.3%)	41.1	27.5	13.6	17.8	24.8	25.6	26.5	23.1

[a] DNA substrate for the FRET assay: a 16-bp dsDNA with a sequence of 5′-CCACAGGTAGCGACAG-3′.[b] Cleavage preference was derived from the footprint assays of nColE7 in digesting of a 260-bp DNA substrate.[c] The shaded mutants, D493N, D493Q, D493R, K497R, and A526N, are computational redesigned mutants predicted to bind DNA tighter than the wild-type nColE7. The three mutants, D493N, D493Q, and D493R, which cleaved DNA with different sequence preferences at the −1 and +3 sites, are displayed in bold.

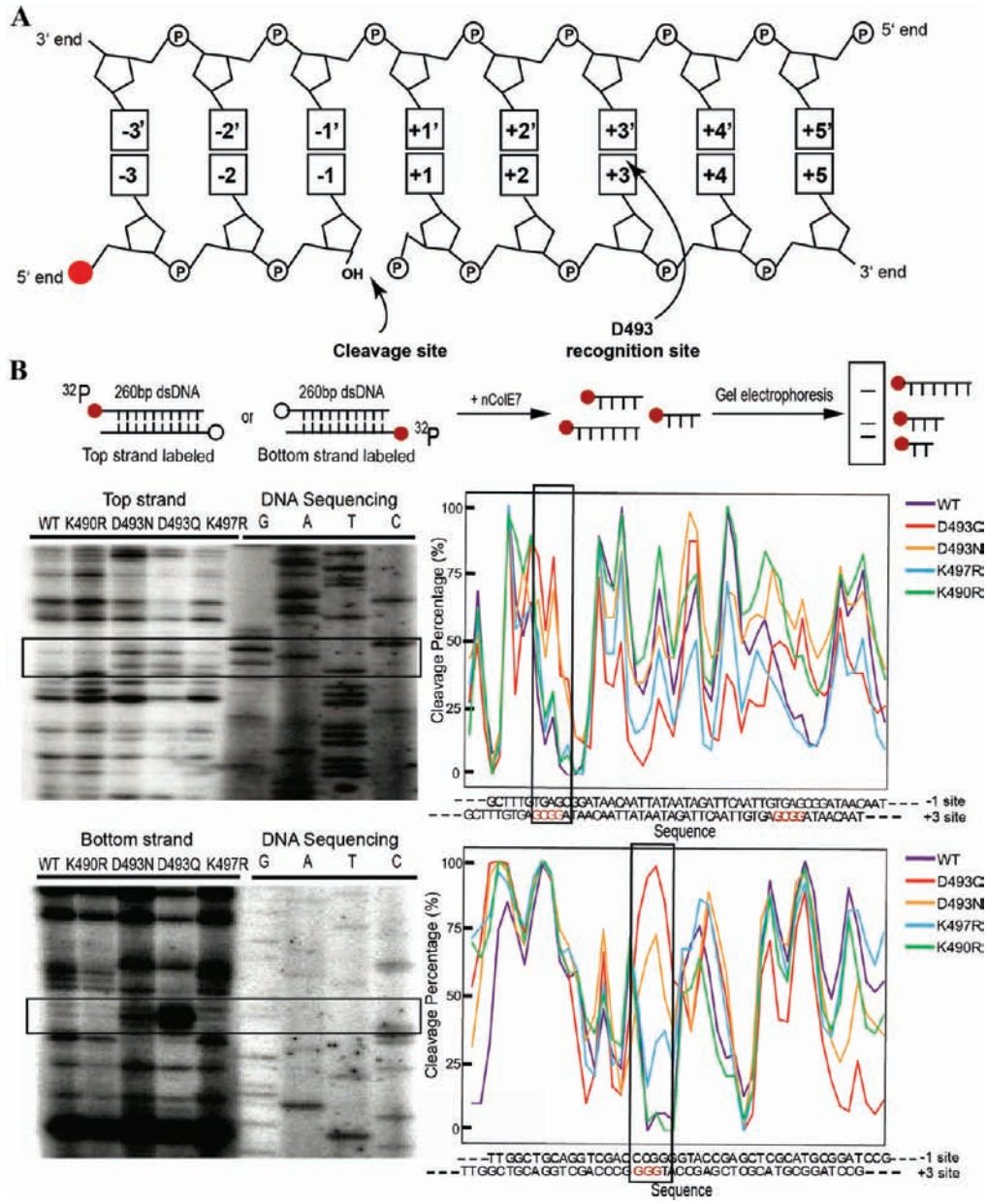

Figure 2. Different DNA cleavage preferences for nColE7 mutants measured by footprint assays. (A) Schematic diagram of the double-stranded DNA cleaved by nColE7. The DNA is numbered from a negative to a positive order from the 5′- to 3′-end with respect to the cleavage site (between the −1 and +1 site) in the cleaved strand, whereas the complementary strand is numbered with the same site number but with an apostrophe. (B) The designed mutants were incubated with a P^{32}-labeled 260-bp dsDNA, and the resultant digested DNA fragments were resolved by gel electrophoresis. Only partial footprint cleavage patterns from K490R, D493N, D493Q, and K497R are representatively shown here. The cleavage preference percentages are normalized and shown at the right panel of the footprint gels. The preference percentages are shown in the Y-axis, and the corresponding sequences at the −1 and +3 sites are shown along the X-axis. The regions marked in boxes clearly show that wild-type nColE7 does not prefer GC-rich regions at the +3 site; in contrast, D493Q and D493N prefer to cleave these regions.

spectrofluorometer. For a nonspecific nuclease, it is not feasible to measure the binding affinity and cleavage rate at a specific DNA site; therefore, the kinetic data obtained reflect the average result of nColE7 bound and cleaved at various sites.

The resulting steady-state kinetics data, k_{cat} and K_m, of the wild-type nColE7 and the redesigned mutants (shaded) are summarized in Table 2. The catalytic activities of all the mutants were in the range $(1.6-2.8) \times 10^{-3}$ s^{-1}, close to k_{cat} of the wild-type enzyme (1.8×10^{-3} s^{-1}), indicating that the mutation did not disrupt protein folding and enzyme activity. The D493N, D493Q, D493R, and K497R mutants, which were predicted to bind the 8-mer DNA more tightly than the wild-type protein (see Table 1), indeed manifested a significantly increased DNA-binding affinity (2.7-, 4.8-, 4.9-, and 2.5-fold increase, respectively), leading up to a 3- to 4-fold increase of overall enzyme activity (k_{cat}/K_m). On the other hand, the two negative controls, K490R (41.4 nM) and A526Q (46.9 nM), had a similar DNA binding affinity to that of the wild-type enzyme (46.7 nM). In summary, four of the five redesigned mutants tested had significantly higher DNA binding affinity and hydrolytic activity, suggesting that this screening method is a useful means of generating redesigned nonspecific DNA-binding proteins with high DNA-binding affinity.

D493Q and D493N Show Different Sequence Preferences in Cleavage. The nColE7 hydrolyzes DNA without sequence specificity but with sequence preference. It prefers to cleave DNA at the 3′O-side after Thy and Ade (−1 site in Figure 2A) with a cleavage preference order of Thy (41.2%) > Ade (29.5%) > Cyt (17.1%) > Gua (12.2%).[20] Since the redesigned mutants had improved DNA-binding affinity and the mutated residues were located at the DNA major groove, their DNA cleavage sequence preference might differ from that of the wild-type nColE7. To determine their DNA cleavage sequence preference, the designed mutants and negative controls were incubated with a P^{32}-labeled 260-bp dsDNA, and the resultant digested DNA fragments were resolved by gel electrophoresis (Figure 2B). The footprint gels of wild-type and mutated nColE7 revealed that K490R, K497R, A526Q, and A526N cleaved DNA with patterns similar to the wild-type enzyme pattern (the cleavage patterns of K490R and K497R are representatively shown in Figure 2B), whereas D493N, D493Q, and D493R cleaved DNA with patterns different from the wild-type enzyme pattern (Figure 2B). This result shows that, among the four higher-affinity mutants, three (D493N, D493Q, D493R) had an altered sequence cleavage preference, whereas one (K497R) had a cleavage preference almost identical to that of the wild-type enzyme.

The quantitative cleavage preference for the mutant at each site was normalized and plotted with respect to the DNA sequence (at the −1 and +3 sites) in Figure 2B. The DNA is cleaved between the −1 and +1 site, and the nucleotides are numbered from a negative to a positive order from the 5′- to 3′-end in the cleaved strand, whereas the complementary strand is numbered with the same site number but distinguished by an apostrophe (see Figure 2A). As the wild-type nColE7 prefers a Thy (41.2%) or Ade (29.5%) at the −1 site, but has no obvious preference for any DNA sequence at the +3 site, the cleavage preferences after Gua at the −1 and +3 sites are 12.2% and 26.3%, respectively. However, D493N, D493Q, and D493R had an increased cleavage preference at the region rich in Gua at the +3 site (marked in box in Figure 2B). As a result, the preferences were increased to 16.9% at the −1 site and 32.6% at the +3 site for D493N, to 20.7% at the −1 site and 33.5% at the +3 site for D493Q, and to 20.6% at the −1 site and 34.6% at the +3 site for D493R (see Table 2, displayed in bold). These results showed that three of the five designed mutants tested, D493Q, D493N, and D493R, not only bound DNA with higher affinity but also had a different sequence preference in cleavage as compared to the wild-type enzyme.

Crystal Structure of D493Q nColE7−DNA Complex Suggests a Structural Basis for the Mutant's Improved DNA Binding Affinity. To verify the accuracy of the predicted design and to elucidate the structural basis of the higher DNA binding activity of D493N and D493Q, these two mutants were screened for cocrystallization conditions with DNA. The D493Q mutant was successfully cocrystallized with an 18-bp dsDNA with a sequence of 5′-GGAATTCGAT▼CGAATTCC-3′ (▼ indicates cleavage site). Without the presence of the Zn^{2+} cofactor during crystallization, the D493Q mutant can only bind but cannot cleave the DNA substrate. The D493Q−DNA complex crystallized in a monoclinic $P2_1$ unit cell with two complexes per asymmetric unit. The structure was solved by molecular replacement using the H545Q nColE7−DNA (18-bp) structure (PDB entry: 2IVH)[20] as the searching model. The final model was refined to an *R*-factor of 20.1% for 10 155 reflections and an *R*-free of 26.4% for 1166 reflections up to a resolution of 2.9 Å. All the data collection and refinement statistics are listed in Table 3.

Table 3. Diffraction and Refinement Statistics for nColE7(D493Q)−DNA(18-bp) Complex

Data collection and processing	
Space group	$P2_1$
Cell dimensions (Å)	a = 60.6, b = 49.4, c = 92.4
(deg)	β= 102.64
Resolution (Å)	2.9
Observed reflections	37,972
Unique reflections	11,588
Completeness– all data (%)	94.9 (50.0– 2.9 Å)
Completeness– last shell (%)	75.5 (2.8– 2.9 Å)
R_{sym}[a]– all data (%)	6.9
R_{sym}[a]– last shell (%)	14.9
$I/\sigma(I)$, all data	26.1 (40.0– 2.8 Å)
$I/\sigma(I)$, last shell	4.0 (2.9– 2.8 Å)
Refinement	
Resolution range (Å)	50.0– 2.9
Reflections (work/test)	10,155/ 1,166
R–factor/ R–free (%)[b]	20.1/ 26.4
Non–hydrogen atoms	
Protein	2,128
DNA	1,464
Solvent molecules	71
Model quality	
r.m.s. deviations in	
bond lengths (Å)	0.0106
bond angles (deg)	1.43
Average B–factor (Å^2)	48.3
Protein atoms	45.5
DNA atoms	53.4
Solvent atoms	27.0

[a] $R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/I_i(hkl)$. [b] *R*-factor $= \Sigma_{(hkl)}||F_o(hkl)| - |F_c(hkl)||/\Sigma_{hkl}F_o(h)$.

The overall structure of the D493Q−DNA complex is shown in Figure 3 with one D493Q bound to one 18-bp DNA (i.e., one-half of the asymmetric unit). The omit map shows a well-defined electron density on the mutational site of Q493 (Figure 3B). D493Q was bound to DNA mainly at the minor groove with some residues interacting at the neighboring major groove. A disordered loop (residues 550 to 555) in the H545Q nColE7−DNA complex (PDB entry: 2IVH) became ordered with a visible electron density in the mutant complex. Therefore, the atomic model in D493Q was completely built from residues 445 to 576. The rmsd between H545Q nColE7 and the D493Q mutant was only 0.53 Å over 115 C_α atoms, demonstrating that the mutation did not disturb the overall protein structure.

A detailed examination of the protein−DNA interactions showed that most hydrogen bonds and vdW interactions identified in the H545Q−DNA complex were retained in the D493Q−DNA complex. D493Q was bound to the 18-bp DNA at the same site as the H545Q nColE7, even though the two complexes were crystallized in different crystal forms. The schematic diagram of the protein−DNA interactions identified in the D493Q−DNA complex is displayed in Figure 4A. The scissile phosphate and neighboring phosphate backbones made interactions with the protein atoms identical to the ones found in the H545Q−DNA complex,[20] except that additional interactions were identified between the ordered loop residues (R447, I549, and S550) and phosphate (P9) atoms.

The wild-type D493 side chain in the H545Q−DNA complex made a hydrogen bond with the Ade13 N6 atom (+3 site),

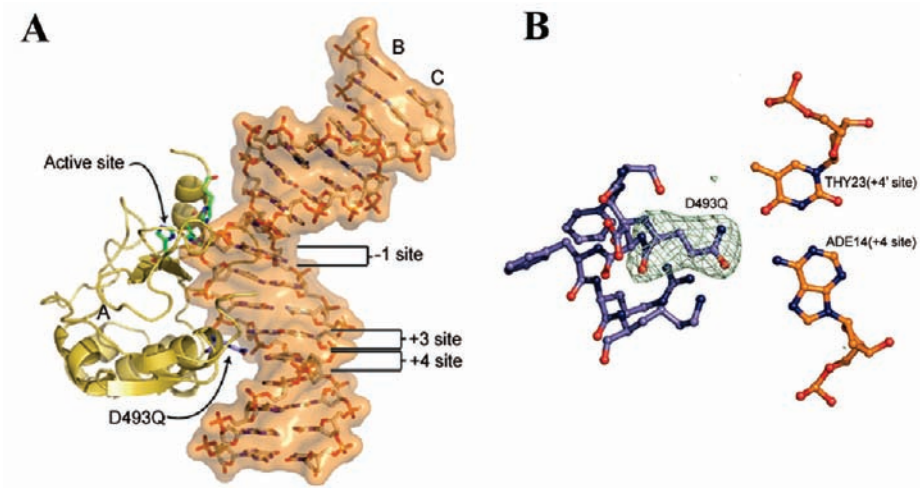

Figure 3. Crystal structure of nColE7 mutant D493Q in complex with DNA. (A) Overall crystal structure of nColE7 D493Q mutant (chain A) bound to an 18-mer DNA (chains B and C). The active site (at −1 site) and the D493 (site +3) or Q493 (site +4) recognition sites are marked on the DNA. (B) The (F_o−F_c) omit difference electron density maps surrounding Q493, contoured at 4σ.

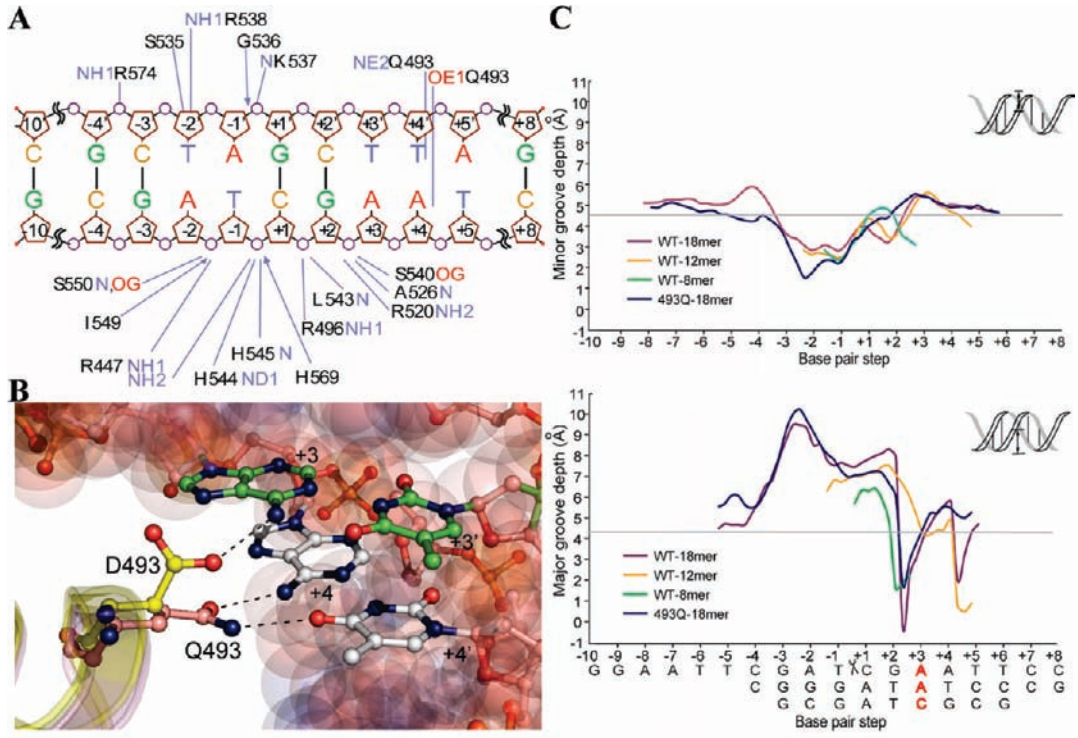

Figure 4. Protein−DNA interactions between D493Q nColE7 mutant and DNA. (A) Schematic diagram of the protein−DNA interactions identified in the crystal structure of D493Q−DNA complex. Hydrogen bonds are represented by solid lines, and nonbonded contacts are represented by arrows (<3.2 Å). (B) Superposition of the D493Q−DNA complex structure with the H545Q−DNA complex structure (PDB entry: 2IVH) shows that the D493 side chain in the H545Q−DNA complex (in yellow) made a hydrogen bond with Ade13 N6 (+3 site), whereas the Q493 side chain in the D493Q−DNA complex (in pink) changed its conformation and made two hydrogen bonds with Ade14 N6 (+4 site) and Thy23 O4 (+4′ site). (C) The major and minor groove depths of DNA bound to nColE7 in four structures are displayed, including wild-type and mutated nColE7 bound to an 8-bp DNA (1PT3), a 12-bp DNA (1ZNS), and an 18-bp DNA (2IVH and this study). The binding of nColE7 to different sequences of DNA induces a conserved DNA conformational change with a deep major groove from −2 to +2 sites (bases 6−12) followed by a dramatic change to a shallow major groove at the +3 site (base 13).

whereas the mutant Q493 side chain in the D493Q−DNA complex changed its conformation and made two hydrogen bonds with Ade14 N6 (+4 site) and Thy23 O4 (+4′ site). The crystal structure of D493Q−DNA was superimposed on that of H545Q−DNA in Figure 4B. The D493 and Q493 side chains exhibit different conformations, interacting with the +3 site and +4 site, respectively. Although the Q493 mutant made hydrogen bonds with the DNA, as predicted, it changed its side chain conformation to interact with the bases at the +4 site instead of the +3 site. The computational screening was based on the structure of nColE7 bound to the 8-bp DNA whose sequence differs from the 18-bp DNA in the crystal structure of the D493Q−DNA complex. This suggests that, compared to the wild-type D493 side chain, the Q493 side chain is more extended and flexible so it can interact with the +4 site upon binding to different DNA sequences.

Table 4. Free Energy Contributions (in kcal/mol) of the Interface Residues in the Wild-Type and D493Q/N nColE7 towards Binding DNA

DNA	GCGATCGC (PDB: 1PT3)			GGAATTCGATCGAATTCC[d] (PDB: 3FBD)	
nColE Residue *i*[a]	WT ΔG_i[b]	D493Q $\Delta\Delta G_i$[c]	D493N $\Delta\Delta G_i$[c]	WT ΔG_i[b]	D493Q $\Delta\Delta G_i$[c]
K490	−1.3 ± 0.3	0.2 ± 0.3	0.1 ± 0.4	−5.9 ± 0.9	−0.9 ± 1.2
S491	0.9 ± 0.7	−1.1 ± 0.5	−2.1 ± 0.5	0.9 ± 0.4	0.6 ± 0.6
493	**11.6 ± 2.2**	**−14.3 ± 1.5**	**−13.5 ± 1.5**	**14.4 ± 3.3**	**−12.9 ± 2.3**
R496	−12.1 ± 1.4	6.7 ± 2.1	5.4 ± 2.1	−16.8 ± 1.8	5.6 ± 1.7
K497	2.4 ± 1.2	−0.7 ± 0.8	1.5 ± 1.2	3.5 ± 2.2	3.9 ± 2.2
R520	−8.9 ± 1.2	−0.5 ± 1.4	0.4 ± 1.4	−7.7 ± 1.0	−2.8 ± 1.1
K525	−8.6 ± 1.0	0.6 ± 1.0	−0.1 ± 1.3	−8.6 ± 1.4	−1.3 ± 1.3
A526	−2.4 ± 0.5	−0.1 ± 0.5	0.2 ± 0.6	−2.5 ± 0.5	0.1 ± 0.4
S535	1.3 ± 0.7	0.3 ± 0.6	−1.5 ± 0.6	0.6 ± 0.7	0.3 ± 0.6
G536	−0.8 ± 0.1	−0.5 ± 0.3	−0.3 ± 0.3	−0.2 ± 0.4	−0.8 ± 0.4
K537	−2.9 ± 0.1	−4.7 ± 1.3	−4.3 ± 1.3	1.4 ± 2.3	−5.7 ± 1.7
R538	−6.5 ± 1.1	−2.4 ± 1.5	−3.2 ± 1.7	1.9 ± 3.2	−2.8 ± 2.8
S540	−4.5 ± 1.0	−0.3 ± 1.1	−0.2 ± 0.7	−5.1 ± 1.2	−0.6 ± 1.0
E542	4.9 ± 1.6	5.3 ± 2.3	3.6 ± 1.5	9.9 ± 2.4	−6.0 ± 2.0
L543	−1.8 ± 0.8	0.4 ± 0.9	0.3 ± 0.8	−1.8 ± 1.0	0.1 ± 0.9
H545	−5.6 ± 0.8	3.6 ± 0.8	2.6 ± 0.9	−10.4 ± 2.2	−3.9 ± 1.8
H569	−2.4 ± 0.9	0.0 ± 1.1	−2.5 ± 1.2	−7.6 ± 1.4	−3.4 ± 1.5
I570	−0.9 ± 0.1	−0.5 ± 0.2	−0.3 ± 0.2	−1.4 ± 0.2	0.2 ± 0.3
R574	−0.7 ± 1.0	0.3 ± 1.0	0.3 ± 0.7	−4.7 ± 1.2	−0.5 ± 1.3

[a] Interface residues are those that have a contact within 4 Å of a DNA atom for more than 1/2 of the sampled trajectory. [b] Free energy contribution (in kcal/mol) of interface residue *i* in the wild-type nColE7 toward binding DNA. [c] Free energy contribution (in kcal/mol) of interface residue *i* in the mutant nColE7 toward binding DNA relative to its contribution in the wild-type protein. [d] The first six bases in the 18-bp dsDNA, which extended past nColE7, were omitted in the calculation to reduce the system's size for the free energy calculations.

Physical Basis for the Improved DNA Binding of the D493Q/N Mutants. To provide a physical basis for the observed higher DNA binding affinity of the D493Q and D493N mutants, the overall free energy for wild-type or D493Q/N nColE7 binding to the 8-mer DNA, GCGATCGC, was decomposed into the contributions from each of the interface residues. The DNA-binding free energy contributions of the interface residues in the wild-type protein (ΔG_i) as well as those in the mutant nColE7 relative to the wild-type protein ($\Delta\Delta G_i$) are listed in Table 4. In the simulation of the wild-type protein bound to the 8-mer DNA, D493 is in vdW contact with the S491 hydroxyl group and forms four hydrogen bonds: two hydrogen bonds with two DNA bases (Cyt6, +3 site and Cyt10, +4′ site), an internal hydrogen bond with the R496 side chain, and a backbone−backbone hydrogen bond with K497 (Table 1). Despite these favorable contacts, D493 makes an unfavorable contribution to the overall DNA-binding free energy (ΔG_{D493} positive) mainly because of the repulsion between its negatively charged side chain and the negatively charged DNA phosphate backbone (Table 4). In the D493Q/N mutant, the mutant residue forms hydrogen bond(s) with the DNA directly and indirectly via R496, while maintaining the native internal hydrogen bonds and vdW contact (Table 1). As the mutant N/Q residue not only maintains hydrogen bonds to the DNA but also eliminates unfavorable electrostatic interactions with DNA, it makes a significant favorable contribution to the relative DNA-binding free energy ($\Delta\Delta G_{D493Q/N} \approx -14$ kcal/mol, Table 4), and the respective mutant nColE7 could bind DNA tighter than the wild-type protein.

To ensure that the physical basis found for the improved DNA binding of the D493Q/N mutants does not depend on the DNA sequence, the free energy decompositions were also performed using the D493Q−DNA X-ray structure solved in this study but with the first six bases in the 18-bp dsDNA, which extended past nColE7, omitted to reduce the system's size for the free energy calculations. Thus, the free energy contributions of the interface residues in the wild-type and D493Q nColE7 toward binding the 12-mer DNA (underlined), GGAATT**CGATCGAATTCC**, were computed. In the simulation of the wild-type protein bound to the 12-mer DNA, D493 forms three hydrogen bonds: one with the DNA base (Ade13, +3 site), one with the R496 side chain, and a backbone−backbone hydrogen bond with K497. However, it makes an unfavorable contribution to the overall DNA-binding free energy (ΔG_{D493} positive, Table 4), as in the case for the wild-type protein bound to the 8-mer DNA. In the MD structure of the D493Q mutant, the Gln side chain forms two hydrogen bonds with Thy23 and Ade22 bases (+4′ and +5′ site; see Supporting Table S2), while retaining the backbone−backbone hydrogen bond with K497. The extra protein−DNA hydrogen bond combined with the elimination of unfavorable electrostatic interactions with the phosphate backbone enables the mutant residue to bind DNA tighter than the wild-type residue ($\Delta\Delta G_{D493Q} \approx -12.9$ kcal/mol, Table 4); hence the D493Q mutant nColE7 also binds the 12-mer tighter than the wild-type protein. In summary, substituting the wild-type D493 with Gln maintained the hydrogen bonds within nColE7 and between nColE7 and DNA but eliminated the unfavorable repulsion between the negatively charged D493 carboxylate and the DNA phosphate groups. As a result, D493Q binds to DNA of different sequences with higher affinity than that of wild-type nColE7.

Discussion

In this study, four of the five designed nColE7 mutants bind DNA with 2- to 5-fold higher affinity and cleave DNA with 3- to 4-fold higher activity. Among the four higher-affinity mutants, D493Q, D493N, and D493R exhibit sequence preferences in DNA cleavage that differ from from those of the wild-type nColE7. In contrast to wild-type nColE7, which does not prefer to cleave regions rich in Gua at the +3 site (boxed in Figure 2), the three designed mutants, D493N, D493Q and D493R, prefer to cleave Gua-rich regions; hence the cleavage preference for a Gua residue at the +3 site is increased significantly (see Table 2). Analyses of the D493Q−DNA complex crystal structure and the DNA-binding free energies provide a physical basis for the observed improved affinity of the nColE7 mutant

D493Q for DNA. Both analyses show that the Q493 side chain hydrogen bonds directly to the DNA. The crystal structure of the D493Q−DNA complex confirmed that the hydrogen-bonding network within nColE7 and between nColE7 and DNA is preserved, whereas the free energy decompositions further show that unfavorable electrostatic interactions with the phosphate backbone are eliminated, thus contributing to the observed enhanced affinity. In general, for a nonspecific DNA-binding protein, replacing an acidic residue at the protein−DNA interface with a polar or positively charged residue such as Asn, Gln, and Arg that can maintain the native hydrogen-bonding network would likely improve the protein−DNA binding affinity.

Interestingly, D493 is conserved in the nuclease colicins from *Escherichia coli*, including ColE2, ColE7, ColE8, and ColE9.[29] However D493 is not strictly conserved in the related pyocins S1 and AP41 from *Pesudomonas aeruginosa*.[30,31] The residue in pyocins corresponding to D493 is an arginine, indicating an increased protein−DNA binding affinity. On the other hand, the residue in pyocins corresponding to K497 is an acidic glutamate, but it is either a lysine or an arginine in the other colicins. Therefore, it seems that colicins and pyocins avoid high-affinity binding to DNA by placing an acidic residue in the protein−DNA interface, viz., an aspartate in colicins or a glutamate in pyocins. These toxins may have evolved under selective pressure to cleave random sequences of DNA without specificity to induce cell death efficiently.

Why does the mutation at D493 produce such a change in the sequence cleavage preference of nColE7? The abundant crystal structures of nColE7 in complex with different sequences of DNA provide a clue to this issue. Four crystal structures, wild-type nColE7 in complex respectively with an 8-bp,[22] 12-bp,[23] and 18-bp[20] DNA and D493Q in complex with an 18-bp DNA (this study), were superimposed to line up the DNA at the scissile phosphates. We found that the protein-induced DNA conformational changes are similar in these structures regardless of the differences in DNA sequences and crystal packing. For each DNA, the major and minor groove depths are depicted in Figure 4B. One of the perceptible features of DNA conformational changes is that all of the complexes have a shallow minor groove and a deep major groove from the −2 to +2 sites (nucleotides 6−12 in Figure 4B) due to the binding of the nColE7 active site at the minor groove, which pulls the DNA closer to the enzyme. A dramatic decrease of the major groove depth is followed at the +3 site (nucleotide 13 in the figure) due to the interactions of nColE7 at the major groove in this region.

This conserved DNA conformational change observed among these complexes appears to be important for the interactions between nColE7 and DNA. The interactions of D493 with the DNA in the major groove contribute to the sharp change from a deep to a shallow major groove at the +3 site. The D493 carboxylate side chain can only hydrogen bond to adenine and cytosine with $-NH_2$ proton donor groups, but not to guanine or thymine with no $-NH_2$ group in the major groove. On the other hand, an amide side chain (or a basic Arg side chain) can serve as a proton donor and hydrogen bond with proton acceptor atoms in guanine at the major groove. Therefore, mutation of D493 to N/Q/R in nColE7 significantly increases its cleavage preference for Gua at the +3 site. This structural comparison highlights the importance of protein-induced DNA deformation in protein−DNA interactions, as noted previously.[20]

Conclusions

This study shows a systemic way of creating higher-affinity and higher-activity nucleases by combining crystal structural data with computational screening methods. The analysis of SASA was used as a starting screening condition to select residues at the protein−DNA interface. Virtual mutations of the selected residues to arginine, asparagine, and glutamine, the three most frequently found residues at protein−DNA interfaces, assisted by MD simulations and free energy calculations identified candidate mutants with improved DNA-binding activities. Using nColE7 as a model system, we show biochemically and structurally that this method was feasible and can be applied to nonspecific nucleases, which have never been tested in protein redesign due to the complexity of DNA binding with different sequences. In principle, this method can be applied to all DNA-binding proteins, and the results are expected to be better for site-specific DNA-binding proteins, which have a more defined DNA-binding site as compared to nonspecific DNA-binding proteins.

Methods

Molecular Dynamics Simulations. Starting Structures. The initial wild-type nColE7^{wt}−DNA complex was taken from PDB entry 1PT3, chain A, while the initial mutant nColE7^{mut}−DNA complexes were generated as described in the Results section. For the D493Q nColE7 bound to the DNA 12-mer, the initial structure was taken from the X-ray structure reported in this study.

Force Field. The simulations were carried out using the all-hydrogen CHARMM27 parameter set[32] for the protein and DNA atoms and the TIP3P model for the water molecules.[33] At the simulation pH of 7, all Asp and Glu residues were deprotonated and all Arg and Lys residues were protonated, while His residues were protonated according to the local environment predicted by the Whatif program.[25]

Simulation Protocol. Each DNA-bound structure was neutralized by placing sodium or chlorine counterions at positions of high electrostatic potential ≥4 Å from any protein or DNA atom and ≥7 Å from each other. The neutral protein•DNA complex was then solvated with a sufficient number of TIP3P water molecules to create a truncated octahedron with a square face distance of 90 Å. This resulted in 59 389 atoms for the fully solvated neutral nColE7^{wt}•DNA complex. The simulations were performed at 300 K using a 1 fs time step, periodic boundary conditions, vdW interactions shifted to zero at 12 Å, and electrostatic interactions treated by particle mesh Ewald summation.[34] Constraints on the terminal DNA base pairs were applied to prevent them from unwinding. Each simulation was continued until the protein backbone rmsd from the initial structure had stabilized (i.e., fluctuated less than 0.1 Å) for at least 300 ps, yielding a simulation period of 1.3 ns. Coordinates were saved every 0.1 ps from the last 100 ps of each trajectory for a total of 1000 coordinate sets for each complex.

Analyses. The saved coordinates were used to (i) generate an average structure, (ii) analyze hydrogen bonds and vdW contacts, and (iii) compute the binding free energies. A hydrogen bond is defined as having a mean heavy−heavy donor−acceptor distance of ≤3.5 Å and a mean heavy−light distance of ≤2.5 Å. The vdW

(29) Ko, T. P.; Liao, C. C.; Ku, W. Y.; Chak, K. F.; Yuan, H. S. *Structure* **1999**, *7*, 91–102.
(30) Sano, Y.; Matsui, H.; Kobayashi, M.; Kageyama, M. *J. Bacteriol.* **1993**, *175*, 2907–2916.
(31) Sano, Y.; Kageyama, M. *Mol. Gen. Genet.* **1993**, *237*, 161–170.
(32) MacKerell, J. A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(34) Darden, T.; Perera, L.; Li, L.; Pedersen, L. *Structure* **1999**, *7*, 55–60.

Scheme 1

$$\begin{array}{lll} & & G_{gas}(\mathrm{P}) \\ [\mathrm{P}]_{\mathrm{gas}} + & [\mathrm{DNA}]_{\mathrm{gas}} & \rightarrow \quad [\mathrm{P{\bullet}DNA}]_{\mathrm{gas}} \\ \uparrow -G_{solv}(\mathrm{p}) & \uparrow -G_{solv}(\mathrm{DNA}) & \quad \downarrow G_{solv}(\mathrm{P{\bullet}DNA}) \\ [\mathrm{P}]_{\mathrm{sln}} + & [\mathrm{DNA}]_{\mathrm{sln}} & \rightarrow \quad [\mathrm{P{\bullet}DNA}]_{\mathrm{sln}} \\ & & G^{o}(\mathrm{P}) \end{array}$$

contacts between the protein and DNA are defined as having a mean heavy−heavy donor−acceptor distance between 3.5 and 4.0 Å.

Free Energy Calculations. The DNA-binding free energy of wild-type or mutant nColE7 (denoted by P) in solution was based on Scheme 1:

It is given by

$$\Delta G^{\circ}(\mathrm{P}) = \Delta G_{gas}(\mathrm{P}) + G_{solv}(\mathrm{P{\bullet}DNA}) - G_{solv}(\mathrm{P}) - G_{solv}(\mathrm{DNA}) \quad (1)$$

where ΔG_{gas} is the standard free energy change per mole for the noncovalent association of the nColE7 and DNA in the gas phase at 300 K, and $-G_{solv}$ corresponds to the work of transferring the molecule in its solution conformation to the same conformation in the gas phase at 300 K. The ΔG^{o} calculations have been described in detail by Lo et al.[35] and are summarized in the Supporting Information.

Protein Expression and Purification. The wild-type nColE7−Im7 heterodimer was expressed using the vector pQE70 (Qiagen) in *E. coli* M15 cells as described earlier.[29] The nColE7 was then separated from Im7 and purified by chromatographic methods.[22] The nColE7 mutants K490R, D493R, D493N, D493Q, K497R, A562Q, and A526N were constructed using the Quick-Change Site-Directed Mutagenesis Kit (Stratagene, USA) and were expressed and purified in the same manner as that for wide-type nColE7. The molecular mass of each mutant was verified using mass spectrometry, and the measured molecular weight matched well with the calculated weights (data not shown).

Endonuclease Activity Measured by the FRET Method. Two complementary 16-nucleotide single-stranded DNA substrates were labeled respectively with fluorogenic material FAM (6-carboxyl-fluorescein) and TAMRA (6-carboxyl-tetramethylrhodamine) at the 5′ end (Biotech, Taiwan) to generate FAM-5′-CCACAGGTAGC-GACAG-3′ and TAMRA-5′-CTGTCGCTACCTGTGG-3′. These two single-stranded DNA were annealed together by heating to 90 °C for 15 min and cooling down slowly to room temperature. The wild-type and mutated nColE7 was then mixed respectively with the double-stranded labeled DNA in the buffer of 10 mM Tris-HCl (pH 8.0) and 1 mM $MgCl_2$ at 25 °C. The increased fluorescence emission intensity at 486 nm, resulted from the DNA cleavage by nColE7, was measured on a fluorescence plate reader (Wallac 1420 multilabel counter) over a period of 120 s at 25 °C. The protein concentration used in the measurement was fixed at 0.5 nM with varied DNA concentrations of 20, 50, 100, 150, 200, and 400 nM. The FRET measurements of each protein were repeated for at least three times, and the resulted kinetic data of K_{m} and k_{cat} are summarized in Table 2.

Footprint Assays. 260-bp DNA fragments were amplified from pQE30 plasmids by polymerase chain reaction. The purified DNA was then 5′-end-labeled with [γ-^{32}P]ATP by the T4 PNK (NEB, USA). The DNA fragments were digested respectively by the wild-type nColE7 and nColE7 mutants K490R, D493R, D493N, D493Q, K497R, A562Q, and A526N in a buffer containing 20 mM Tris-HCl (pH 8) and 2 mM Mg^{2+} at 37 °C for 5 min. The digestion reaction was then terminated by adding 10 mM EDTA with the loading dye. Thereafter, the digested DNA fragments were heated to 70 °C for 5 min and resolved on 8% denaturing polyacrylamide gel electrophoresis. The PAGE cleavage patterns were visualized, and the band intensities were measured on a phosphorimager (AlphaImage 2200). The cleavage percentages were calculated by first integrating each cleavage band, and then the band intensities were normalized by calculating the summation of $I_{\mathrm{base}}/N_{\mathrm{base}}$ divided by the summation of I_{all}, in which I_{base} is the band intensity cleaved at a specific base (A, T, C, or G), I_{all} is the band intensity cleaved at all bases, and N_{base} is the number of a specific base in the DNA.

Protein Crystallization, Data Collection, and Structural Determination. Crystals of the nColE7 mutant D493Q−DNA complex were obtained by the hanging drop vapor-diffusion method. The protein−DNA complex was prepared by mixing the nColE7 mutant D493Q (7 mg/mL) with an equal molar ratio of the 18-bp DNA: 5′-GGAATTCGATCGAATTCC-3′. The protein−DNA solution (1 μL) was further mixed with reservoir solution (1 μL) containing 15% PEG3350, 16.25% MPD, 0.15 M ammonium acetate, and 0.025 M sodium acetate at room temperature. The D493Q−DNA complex crystallized in a monoclinic $P2_1$ unit cell with two protein−DNA complexes per asymmetric unit. X-ray diffraction data were collected at the beamline 13C1 of the NSRRC in Hsinchu, Taiwan, and were processed and scaled by HKL2000.[36]

The structure was solved by molecular replacement using the structure of wild-type nColE7 in complex with an 18-bp DNA (PDB entry: 2IVH) as the searching model. The structure model was subjected to manual rebuilding with Coot and then refined with the CNS. The final complex model had an R factor of 20.1% and an R-free of 26.4% for 10 155/1166 reflections up to a resolution of 2.8 Å. Structural coordinates and diffraction structure factors have been deposited in the RCSB Protein Data Bank with PDB ID codes of 3FBD.

Acknowledgment. This work was supported by research grants from Academia Sinica and the National Science Council, Taiwan, ROC. Portions of this research were carried out at the National Synchrotron Radiation Research Center (BL-13B1 and BL-13C1), a national user facility supported by the National Science Council of Taiwan, ROC. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine.

Supporting Information Available: The comparison of hydrogen bonds in X-ray structures and simulated structures is provided in the two supplement Tables S1 and S2. The procedure for free energy calculation is described. The full author list of ref 32 is provided. This material is available free of charge via the Internet at http://pubs.acs.org.

JA907160R

(35) Lo, C.-H.; Chang, Y.-H.; Wright, J. D.; Chen, S.-H.; Kan, D.; Lim, C.; Liang, P.-H. *J. Am. Chem. Soc.* **2009**, *131*, 4051–4062.

(36) Otwinowski, Z. *Methods Enzymol.* **1997**, *276*, 307–326.