© 1996 Oxford University Press

Nucleic Acids Research, 1996, Vol. 24, No. 21 4171–4175

Effects on *Nae*I-DNA recognition of the leucine to lysine substitution that transforms restriction endonuclease *Nae*I to a topoisomerase: a model for restriction endonuclease evolution

Kiwon Jo and Michael D. Topal*

Lineberger Comprehensive Cancer Center and Department of Pathology, University of North Carolina Medical School, Chapel Hill, NC 27599-7295, USA

Received July 29, 1996; Revised and Accepted September 24, 1996

ABSTRACT

Substituting lysine for leucine at position 43 (L43K) transforms *Nae*I from restriction endonuclease to topoisomerase and makes *Nae*I hypersensitive to intercalative anticancer drugs. Here we investigated DNA recognition by *Nae*I-L43K. Using DNA competition and gel retardation assays, *Nae*I-L43K showed reduced affinity for DNA substrate and the ability to bind both single- and double-stranded DNA with a definite preference for the former. Sedimentation studies showed that under native conditions *Nae*I-L43K, like *Nae*I, is a dimer. Introduction of mismatched bases into double-stranded DNA significantly increased that DNA's ability to inhibit *Nae*I-L43K. Wild-type *Nae*I showed no detectable binding of either single-stranded DNA or mismatched DNA over the concentration range studied. These results demonstrate that the L43K substitution caused a significant change in recognition specificity by *Nae*I and imply that *Nae*I-L43K's topoisomerase activity is related to its ability to bind single-stranded and distorted regions in DNA. A mechanism is proposed for the evolution of the *Nae*I restriction–modification system from a topoisomerase/ligase by a mutation that abolished religation activity and provided a needed change in DNA recognition.

INTRODUCTION

The restriction enzyme *Nae*I provides an important model for protein–DNA interactions: *Nae*I must bind two DNA sequences to cleave (1–5), and has either endonuclease activity or topoisomerase activity depending on whether amino acid 43 is, respectively, leucine or lysine (6). A 10 amino acid region, near the NH_2-terminus of the putative *Nae*I substrate site (7), matches the consensus for the active site of DNA ligase I with one important difference: the lysine (K) that forms the adenylated intermediate essential for catalysis by the DNA ligase active site is not present in *Nae*I. Instead, there is a leucine (L43) at this position. Changing L43 to K43 transforms *Nae*I activity to that of a topoisomerase and recombinase even though *Nae*I has no detectable sequence similarity with the topoisomerase and recombinase protein families (6). *Nae*I and *Nae*I-L43K form covalent intermediates with the 5′ end of the cleaved DNA (6,8), and thus the topoisomerase activity appears to result from endonuclease and ligase domains coupled through the covalent intermediate. The *Nae*I ligase-like domain may catalyze phosphodiester bond formation using the energy of the activated protein–DNA intermediate in an analogous manner to DNA ligase catalysis of phosphodiester formation using DNA-adenylate as the activated intermediate.

The topoisomerase activity of *Nae*I-L43K is hypersensitive to intercalative anticancer drugs, whereas the endonuclease activity of wild-type (wt) *Nae*I is not (8). This switch to intercalative drug sensitivity brought on by the L43K change suggests possible changes in substrate recognition that are exaggerated by drug intercalation. Therefore, we investigated the DNA recognition properties of *Nae*I-L43K using competition and gel retardation experiments. Here we report that the L43K substitution maintained the native dimeric structure of *Nae*I but changed substrate recognition from double-stranded DNA in wt *Nae*I to a preference for single-stranded DNA and mismatched DNA in *Nae*I-L43K.

MATERIALS AND METHODS

Synthetic DNAs

Oligodeoxyribonucleotides were synthesized using an Applied Biosystems 380B synthesizer. Synthesized DNAs were purified by 12% PAGE followed by phenol extraction, ethanol precipitation and gel permeation chromatography using G-25 Sephadex. For double-stranded DNAs, two complementary oligodeoxyribonucleotides were annealed and electrophoresed on 10% nondenaturing polyacrylamide gel to remove remaining single-stranded DNA. The annealed DNAs were further purified by phenol extraction, ethanol precipitation and chromatography with G-25 Sephadex.

Inhibition of *Nae*I-L43K DNA relaxation activity

Plasmid pBR322 (11.6 nM) and purified *Nae*I-L43K (56 nM) were incubated with the DNA fragments (5 nM) being tested as inhibitors at 37°C in 15 µl containing 10 mM Tris–HCl (pH 8.0),

*To whom correspondence should be addressed. Tel: +1 919 966-5040; Fax: +1 919 966-3015; Email: mdtopal@med.unc.edu

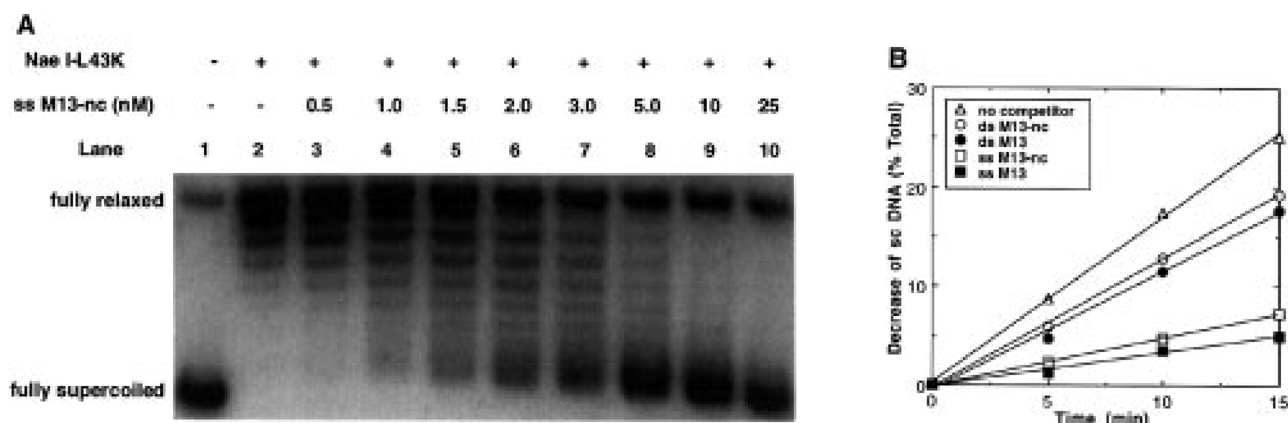

Figure 1. Effect of single- and double-stranded DNA inhibitors on DNA relaxation activity of mutant *Nae*I-L43K. (**A**) Inhibition of *Nae*I-L43K (0.23 μM) relaxation of pBR322 (11.6 nM) by single-stranded DNA without cognate sequence (ss M13-nc). Reactions were incubated for 1 h under conditions described in Materials and Methods. Lane 1, DNA control; lanes 2–10, inhibition of DNA relaxation by increasing concentrations of single-stranded M13-nc as indicated. (**B**) Apparent initial rates of relaxation of pBR322 by *Nae*I-L43K in the presence of 5 nM of each DNA competitor shown. Reaction conditions are described in Materials and Methods. Reactions were stopped at indicated time intervals and quantitations of fully supercoiled DNA were done by densitometry of bands after electrophoresis as in (A).

20 mM NaCl, 5 mM $MgCl_2$, bovine serum albumin (0.1 mg/ml) and 5 mM β-mercaptoethanol. The reactions were stopped by addition of SDS to a final concentration of 1%. Products were resolved by 0.8% agarose gel electrophoresis and visualized by EtBr staining. Results were plotted as apparent initial rates as described in Results and Discussion.

Gel retardation assay

DNA inhibitors (5 nM), 5′-end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP, were incubated with increasing amounts of *Nae*I-L43K (as indicated) under the conditions described above, supplemented with 10% glycerol, except 5 mM $MgCl_2$ was replaced with 5 mM $CaCl_2$ to eliminate cleavage (5). Immediately after incubation for 20 min at room temperature, binding reactions were loaded onto a 6% (29:1) polyacrylamide–bisacrylamide gel and electrophoresed for 2 h at 200 V in a buffer containing 5 mM $CaCl_2$, 25 mM Tris–HCl (pH 8.3) and 190 mM glycine. Autoradiography revealed protein–DNA complexes.

RESULTS AND DISCUSSION

DNA inhibition of *Nae*I-L43K and *Nae*I

To investigate the substrate specificity of *Nae*I-L43K, different DNA fragments were screened for their relative abilities to inhibit *Nae*I-L43K relaxation of supercoiled pBR322 (Table 1). A saturating concentration of *Nae*I-L43K (0.23 μM) was used to completely relax pBR322 (Fig. 1A, lane 2). Single-stranded DNA fragments (36mers, Table 1), with or without a *Nae*I cognate site, efficiently inhibited the topoisomerase activity (Fig. 1A): the occurrence of bands indicative of relaxed topoisomers was inhibited completely at 10 nM (concentration of molecules) single-stranded DNA. This concentration was less than that of substrate pBR322 (11.6 nM) present in the reaction. Although production of topoisomers was completely inhibited, some fully relaxed DNA remained. This is consistent with the production of some nicked DNA during the relaxation reaction (6) and suggests that single-stranded DNA may inhibit religation as well as cleavage.

Table 1. DNA fragments tested as competitive inhibitors

name	sequence
ss M13[a]	TTTCTCGCCACGTTC**GCCGGC**TTTCCCCGTCAAGCT
ds M13[a]	TTTCTCGCCACGTTC**GCCGGC**TTTCCCCGTCAAGCT AAAGAGCGGTGCAAG**CGGCCG**AAAGGGGCACTTCGA
ss M13-nc[b]	TTTCTCGCCACGTTC**GAAGAA**TTTCCCCGTCAAGCT
ds M13-nc[b]	TTTCTCGCCACGTTC**GAAGAA**TTTCCCCGTCAAGCT AAAGAGCGGTGCAAG**CTTCTT**AAAGGGGCACTTCGA
ds M13-C1, C2, C5[c]	TTTCTCGCCACGTTCGA**AG**AATTTCCCCGTCAAGCT AAAGAGCGGTGCAAGCT**TC**TTAAAGGGGCACTTCGA
ds M13-mismatch[d]	TTTCTCGCCACGTTCG**CCGGC**TTTCCCCGTCAAGCT AAAGAGCGGTGCAAGC**TTCTT**AAAGGGGCACTTCGA

[a]The *Nae*I cognate sequence is bold. The sequence containing the *Nae*I site is identical to that containing the unique *Nae*I site in bacteriophage M13 DNA.
[b]The noncognate fragment, designated as -nc, differs from cognate by substituting GAAGAA (bold) for the GCCGGC *Nae*I cognate sequence.
[c]The location of one- (C1), two- (C2), and five extra cytosine(s) (C5) residues in the top strand are between the AT and GC base pairs in bold.
[d]The mismatched region is in bold.
Double- and single-stranded are indicated as ds and ss, respectively. Preparation of DNA fragments is described in Materials and Methods.

To compare the relative ability of DNAs (Table 1) to inhibit *Nae*I-L43K topoisomerase activity, apparent initial reaction rates for *Nae*I-L43K were determined in the presence of 5 nM of each DNA competitor. Apparent initial rates of pBR322 relaxation were linear up to at least 15 min. Apparent initial rates (line slopes in Fig. 1B) of topoisomerase activity were determined because *Nae*I-L43K produces DNA molecules with different amounts of supercoiling (6). The apparent initial rate is defined as a loss of fully supercoiled pBR322 DNA, and it is proportional to the steady-state concentration of the complex between *Nae*I-L43K and fully supercoiled pBR322. Thus, the effect of inhibitor on this initial rate depends on the ability of the inhibitor to bind enzyme. The degree of inhibition caused by a competitive inhibitor depends on [S], [I], K_m and K_i: $v/V_{max} = [S]/(K_m(1 + [I]/K_i) + [S])$ (9). We used identical concentrations of pBR322 and

Table 2. Relative inhibition of *Nae*I-L43K relaxation of pBR322 by synthesized DNA fragments at 5 nM concentration

DNA fragment	initial rate (nM/min)	% inhibition
no competitor	0.19	0
ds M13-nc	0.15	21
ds M13	0.14	26
ds M13-C1	0.11	42
ds M13-C2	0.090	53
ds M13-C5	0.072	62
ds M13-mismatch	0.065	66
ss M13-nc	0.056	71
ss M13	0.041	78

Apparent initial rates (defined in text) are line slopes in Figures 1B and 5. % inhibition is decrease in initial rates from no competitor.

identical concentrations of the different inhibitors in each reaction to measure the initial rate. The results in Table 2 show that single-stranded DNA with a cognate site inhibited best, followed closely by single-stranded DNA without a cognate site. Double-stranded DNA inhibited significantly less efficiently (Fig. 1B).

Whereas *Nae*I-L43K was inhibited by low amounts of single-stranded DNA, wild-type *Nae*I was unaffected by single-stranded DNA even at ~50-fold excess over substrate concentration (Fig. 2). Only double-stranded M13 (Table 1) inhibited *Nae*I (Fig. 2), at concentrations consistent with the binding preferences of the substrate binding domain of *Nae*I (10).

Gel retardation analysis of *Nae*I-L43K and *Nae*I interaction with DNA

These results show that *Nae*I-L43K preferentially binds single-stranded DNA. Therefore, gel retardation analysis was used to demonstrate the *Nae*I-L43K–single-stranded DNA complex, and to quantitate the binding of *Nae*I-L43K and *Nae*I with single-stranded and double-stranded DNAs. The gel retardation results (Fig. 3) are consistent with the relative abilities of competitor DNAs to inhibit *Nae*I-L43K topoisomerase activity. Single-stranded DNA clearly bound *Nae*I-L43K (Fig. 3A), with 50% of single-stranded M13 DNA being shifted to a higher mobility

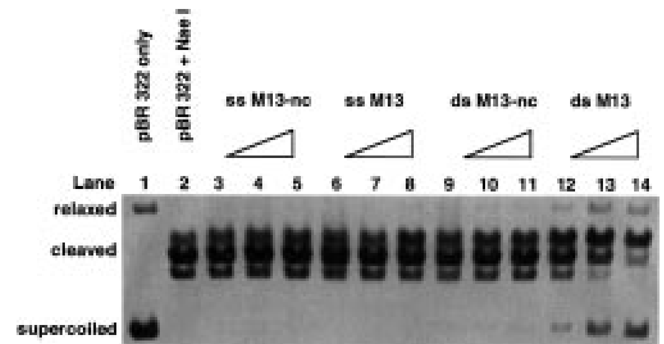

Figure 2. Effects of DNA inhibitors on DNA cleavage by wild-type *Nae*I. Cleavage reactions, identical to the DNA relaxation conditions described in Materials and Methods, except for 8 nM of *Nae*I, were incubated at 37°C for 10 min. Lane 1, pBR322 only; lane 2, with *Nae*I; lanes 3–14, *Nae*I cleavage reaction with 100, 300 and 500 nM of single-stranded M13-nc (lanes 3–5), single-stranded M13 (lanes 6–8), double-stranded M13-nc (lane 9–11) and double-stranded M13 (lanes 12–14), respectively.

species at 1.2 µM of *Nae*I-L43K protein. This means that the K_D (dissociation constant) for the *Nae*I-L43K–single-stranded DNA complex is ~1.2 µM. Single-stranded DNA with the cognate site showed the best binding (Fig. 3B) followed closely by single-stranded DNA with a noncognate site (ss M13-nc); double-stranded M13 and double-stranded M13-nc bound less tightly.

The enzyme–DNA complexes involve both monomeric and dimeric forms of *Nae*I-L43K as implied by the doublet of retarded bands (Fig. 3A), and shown for *Nae*I (5). Sedimentation of *Nae*I-L43K in a glycerol gradient demonstrated that *Nae*I-L43K topoisomerase activity sediments with bovine serum albumin (not shown), and so sediments as a dimer, as does wild-type *Nae*I (5). *Nae*I bound to the double-stranded M13 DNA fragment (Table 1) and *Nae*I-L43K bound to DNA containing a single-stranded region (M13-mismatch DNA fragment in Table 1) showed identical sedimentation patterns in glycerol gradients demonstrating that binding single-strand DNA does not inhibit formation of higher-order structures of *Nae*I-L43K (not shown). Therefore, inhibition of *Nae*I-L43K topoisomerase activity is not caused by conversion of *Nae*I to a monomeric protein.

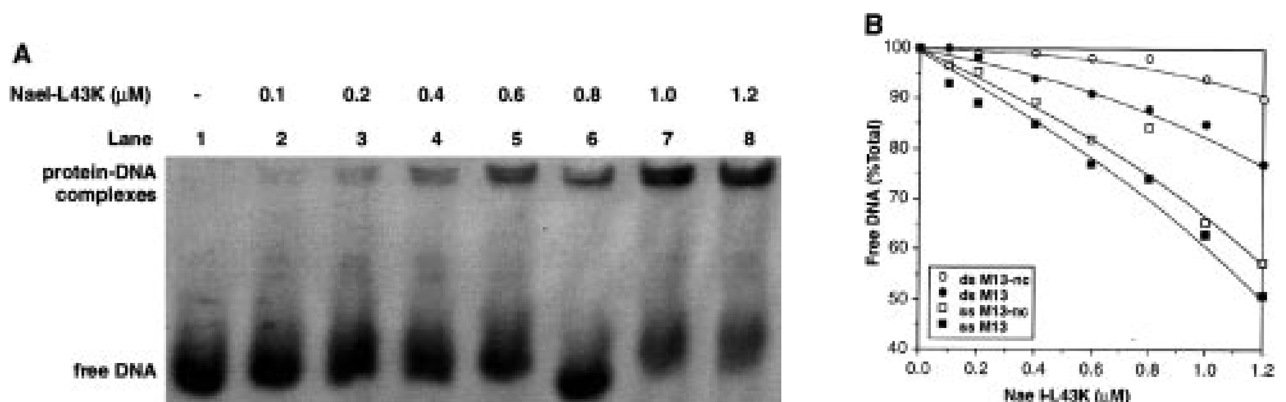

Figure 3. Analysis of protein binding by gel retardation. (**A**) *Nae*I-L43K complexes with single-stranded M13-nc. Reaction conditions are described in Materials and Methods. (**B**) Comparison of relative binding of *Nae*I-L43K to single- and double-stranded DNAs. Each 5′-end-labeled DNA (5 nM) was incubated with increasing amount of *Nae*I-L43K under the same condition as that in (A). After electrophoresis of the reaction products and autoradiography, the amount (% of total) of the free DNA at each concentration of *Nae*I-L43K was quantitated by densitometric analysis of the autoradiogram.

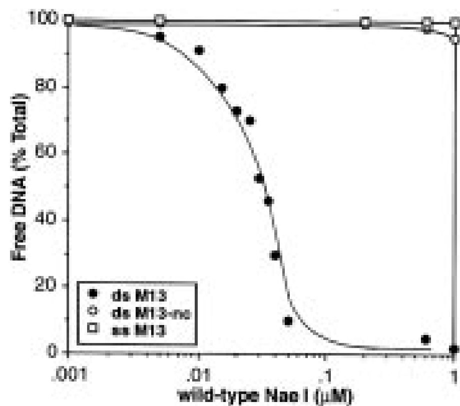

Figure 4. Comparison of relative binding of wild-type *Nae*I to single- and double-stranded DNAs. Increasing amounts of *Nae*I were incubated with 5 nM of each 5′-end-labeled DNA under the same conditions described in Figure 3. The concentration of *Nae*I (μM) is given on a logarithmic scale.

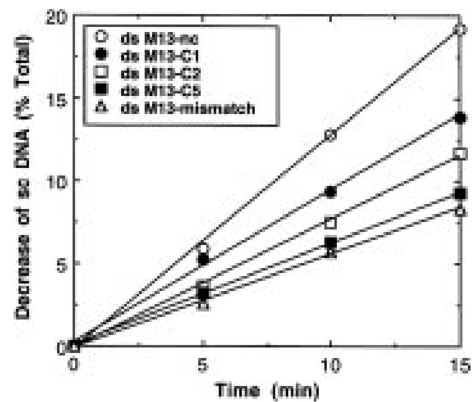

Figure 5. Bulges and mismatches in a double-stranded DNA fragment increase its ability to inhibit *Nae*I-L43K relaxation of pBR322. Apparent initial rates of relaxation of pBR322 in the presence of 5 nM of double-stranded M13-nc (control), bulged DNAs (ds M13-C1, C2, C5) and mismatched DNA (ds M13-mismatch) were compared under the conditions described in Materials and Methods.

Wild-type *Nae*I did not show any detectable binding ability with single-stranded DNA by gel retardation, and at 1 μM concentration formed a small amount (<5%) of complex with double-stranded M13-nc DNA (Fig. 4). The only DNA competitor that showed efficient binding to wild-type *Nae*I was double-stranded M13. Wild-type *Nae*I has a K_D for double-stranded M13 of ~30 nM (Fig. 4). Thus, the L43K mutation of *Nae*I reduced DNA binding by 40-fold, and radically altered substrate recognition enabling *Nae*I-L43K to recognize single-stranded DNA. It is possible that the ability to bind single-stranded DNA is caused by alteration of one of the two DNA-binding domains of *Nae*I. Thus, the remaining ability to bind double-strand DNA could remain with the second DNA binding domain. Alternatively, single-stranded DNA binding could be associated with a third binding site possibly associated with ligase function that is uncovered by a conformational change in the protein because of the L43K substitution.

Ability of mismatched DNA to inhibit *Nae*I-L43K

The ability of *Nae*I-L43K to recognize single-stranded DNA raises the question of whether this enzyme can recognize small regions of single-stranded or distorted DNA within the DNA helix. Therefore, DNA fragments containing mismatched DNA were tested as inhibitors of *Nae*I-L43K topoisomerase activity under apparent initial velocity conditions (Fig. 5). Mismatched insertions (bulges) of one-, two- and five-cytosine(s) were synthesized in double-stranded DNA fragment M13-nc to give double-stranded DNA fragments M13-C1(-C2)(-C5), respectively (Table 1). The DNA fragment with a mismatched region (ds M13-mismatch) contained -CCGGC-/-TTCTT- in the center of the DNA (Table 1). The DNA fragments were all found to have melting temperatures significantly >37°C (not shown), demonstrating that they were double-stranded under our reaction conditions. For example, the double-stranded M13-mismatch and double-stranded M13-C5 DNA fragments have melting temperatures (T_ms) of ~72 and 74°C, respectively. The apparent initial velocities of *Nae*I-L43K DNA relaxation in the presence of the mismatched DNAs and DNA without a mismatch were compared. The abilities of the DNA competitors to inhibit topoisomerase activity increased with the size of the mismatch in the double-stranded DNA helix (Fig. 5 and Table 2), and demonstrate the ability of *Nae*I-L43K to recognize mismatched DNA. Wild-type *Nae*I was not inhibited by any of the mismatched DNAs at DNA concentrations ≤900 nM (not shown). The ability of *Nae*I-L43K to recognize mismatched DNA relates *Nae*I to a potential superfamily of proteins able to recognize mismatched DNA that include, among others, RecA (11), p53 (12), the excision repair proteins *Escherichia coli* UvrABC (13), yeast RAD1–RAD10 (14,15) and human XPF-ERCC1 (16), and the mismatch repair proteins *E.coli* Mut S (17) and human MSH2 (18). Understanding how substitution L43K enables recognition of single-stranded and mismatched DNA may help understand how proteins differentiate distorted and single-stranded DNA from double-stranded DNA.

Substitution L43K radically transforms the activity of *Nae*I protein from endonuclease to topoisomerase (6). Investigation of the effects of substitution L43K on *Nae*I substrate specificity, showed that the topoisomerase activity of *Nae*I-L43K is inhibited by single-stranded and mismatched DNA, and demonstrated by gel retardation analysis complex formation between single-stranded DNA and *Nae*I-L43K. The change in *Nae*I binding affinity, and dramatic transformation of *Nae*I substrate recognition and activity demonstrates the importance of leucine 43 for *Nae*I recognition and function and implies that the topoisomerase activity of *Nae*I-L43K is related to its ability to bind distorted regions in DNA.

Evolution of restriction endonuclease

Given the transformation of *Nae*I from an endonuclease to a topoisomerase by a single amino acid change that restores an essential amino acid in the DNA ligase-like active site of *Nae*I, it is likely that *Nae*I evolved from a topoisomerase/ligase (6). The ability of *Nae*I-L43K to recognize single-stranded and mismatched DNA is similar to that of topoisomerase I. DNA relaxation by topoisomerase I is also strongly inhibited by single-stranded DNA fragments (19), and topoisomerase I must bind a single-stranded DNA region for activity (20–22). Topoisomerase I also appears to recognize mismatches: topoisomerase I nicks 5′ to all eight possible mismatched bases (23). The resulting changes in *Nae*I activity and DNA recognition imply that mutating lysine 43

to leucine in the topoisomerase/ligase was the fundamental change that enabled restriction enzyme *Nae*I to evolve by a loss in ligation activity that concomitantly provided a needed change in recognition specificity.

The relation of *Nae*I to a topoisomerase/ligase and the ability of a single amino acid change to switch off the religation function of that topoisomerase/ligase leads us to propose a framework for the evolution of the *Nae*I restriction–modification system that may be generally applicable to most restriction–modification systems and that provides the basis for future studies of restriction endonuclease evolution. Step 1 involves duplication of a *Nocardia* topoisomerase/ligase gene. Gene duplication has been widely noted and is commonly believed to be the first step in the generation of functionally novel proteins (24), although it is possible for the duplication to come later in the evolutionary pathway (25). The duplication of an essential gene was necessary for Step 2: the emergence of a spectrum of *Nocardia* mutants differing in the DNA binding site of the duplicated topoisomerase/ligase protein. The lack of toxicity of the topoisomerase/ligase is important for Step 2, because the religation activity of the enzyme means that the binding site of a DNA cleaving enzyme could evolve towards any DNA sequence without killing the host. This can be contrasted with trying to directly evolve the specificity of an existing restriction enzyme. Any relaxation or change in specificity of a restriction enzyme is lethal to the host unless the altered specificity happens to overlap the specificity of an existing methylase activity or the methylase and endonuclease evolve in tandem. The latter is highly unlikely because of the requirement for two rare occurrences. In Step 3, the topoisomerase/ligase mutants, under the selective pressure of predation by bacteriophage invasion, are tested by the single K43L substitution that turns off religation and enhances double-strand DNA binding. Those mutants that cleave host DNA are eliminated, whereas those mutants whose cleavage specificity overlaps at least partially with an existing methylation specificity that protects the host DNA from cleavage has a chance at survival as a restriction–modification system that confers a survival advantage against phage infection. The need for only a single amino acid change presents a facile means to test a significant spectrum of the topoisomerase/ligase mutants within a reasonable period of time.

The starting enzyme does not have to be topoisomerase/ligase. Any enzyme with DNA cleaving and religating activity is a possibility. For example, *Eco*RII is related to the integrase family of site-specific recombinases (26). This mechanism for the evolution of restriction–modification systems is directly applicable to the type IIe restriction enzymes, which must bind two DNA sequences to cleave. Loss of the second site, no longer required for restriction enzyme function, may explain the evolution of the single site type II endonucleases as well. Alternatively, enzymes that nick DNA as part, for example, of their DNA transposition or DNA repair function produce damage that can be readily repaired by a separate DNA ligase. Thus, these enzymes, which include enzymes that bind DNA at single sites are candidates for evolving single-site type II restriction enzymes using a similar pathway to that giving rise to the type IIe enzymes. For the type II enzymes the ligase activity does not have to be abolished since it resides in a separate polypeptide. Evolution from single-strand nicking to double-strand cleavage produces damage much less repairable by a separately existing DNA ligase as witnessed by the lethality of restriction endonucleases expressed in cells lacking a protecting methylase.

ACKNOWLEDGEMENTS

We thank Terri Maness for expert technical assistance, and Jack Griffith, Steve Matson and Aziz Sancar for reading parts of the manuscript before publication. This work was supported by grant GM52123 from the National Institutes of Health.

REFERENCES

1 Conrad, M. and Topal, M. D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9707–9711.
2 Oller, A. R., Vanden Broek, W., Conrad, M. and Topal, M. D. (1991) *Biochemistry* **30**, 2543–2549.
3 Topal, M. D., Thresher, R. J., Conrad, M. and Griffith, J. (1991) *Biochemistry* **30**, 2006–2010.
4 Yang, C. C. and Topal, M. D. (1992) *Biochemistry* **31**, 9657–9664.
5 Baxter, B. K. and Topal, M. D. (1993) *Biochemistry* **32**, 8291–8298.
6 Jo, K. and Topal, M. D. (1995) *Science* **267**, 1817–1820.
7 Holtz, J. K. and Topal, M. D. (1994) *J. Biol. Chem.* **269**, 27286–27290.
8 Jo, K. and Topal, M. D. (1996) *Biochemistry* **35**, 10014–10018.
9 Segel, I. H. (1975) *Enzyme Kinetics.* John Wiley & Sons, New York, pp. 100–111.
10 Yang, C. C., Baxter, B. K. and Topal, M. D. (1994) *Biochemistry* **33**, 14918–14925.
11 Wang, Y.-H., Bortner, C. D. and Griffith, J. (1993) *J. Biol. Chem.* **268**, 17571–17577.
12 Lee, S., Elenbass, B., Levine, A. and Griffith, J. (1995) *Cell* **81**, 1013–1020.
13 Sancar, A. (1994) *Science* **266,** 1954–1956.
14 Tomkinson, A. E., Bardwell, A. J., Bardwell, L., Tappe, N. J. and Friedberg, E. C. (1993) *Nature* **362**, 860–862.
15 Sung, P., Reynolds, P., Prakash, L. and Prakash, S. (1993) *J. Biol. Chem.* **268**, 26391–26399
16 Park, C.-H., Bessho, T., Matsunaga, T. and Sancar, A. (1995) *J. Biol. Chem.* **270**, 22657–22660.
17 Modrich, P. (1991) *Annu. Rev. Genet.* **25**, 229–253.
18 Alani, E., Chi, N.-W. and Kolodner, R. (1995) *Genes Dev.* **9**, 234–247.
19 Wang, J. C. (1971) *J. Mol. Biol.* **55**, 532–533.
20 Beran-Steed, R. K. and Tse-Dinh, Y.-C. (1989) *Proteins: Struct. Funct., Genet.* **6**, 249–258.
21 Lima, C. D., Wang, J. C. and Mondragón, A. (1994) *Nature* **367**, 138–146.
22 Wang, J. C. (1996) *Annu. Rev. Biochem.* **65**, 635–692.
23 Yeh, Y.-C., Liu, H.-F., Ellis, C. A. and Lu, A.-L. (1994) *J. Biol. Chem.* **269**, 15498–15504.
24 Ohno, S. (1970) *Evolution by Gene Duplication.* Springer-Verlag, New York.
25 Hughes, A. L. (1994) *Proc. R. Soc. Lond. B* **256**, 119–124.
26 Topal, M. D. and Conrad, M. (1993) *Nucleic Acids Res.* **21**, 2599–2603.