

Structural, functional, and evolutionary relationships between λ -exonuclease and the type II restriction endonucleases

RHETT A. KOVALL AND BRIAN W. MATTHEWS*

Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, University of Oregon, Eugene, OR 97403

Contributed by Brian W. Matthews, May 13, 1998

ABSTRACT λ -exonuclease participates in DNA recombination and repair. It binds a free end of double-stranded DNA and degrades one strand in the 5' to 3' direction. The primary sequence does not appear to be related to any other protein, but the crystal structure shows part of λ -exonuclease to be similar to the type II restriction endonucleases *PvuII* and *EcoRV*. There is also a weaker correspondence with *EcoRI*, *BamHI*, and *Cfr10I*. The structure comparisons not only suggest that these enzymes all share a similar catalytic mechanism and a common structural ancestor but also provide strong evidence that the toroidal structure of λ -exonuclease encircles its DNA substrate during hydrolysis.

DNA exonucleases are essential for DNA replication, recombination, and repair (1). λ -Exonuclease is a processive 5'-3' single-stranded DNA exonuclease that requires Mg^{2+} as a cofactor (2–4). The protein is encoded by bacteriophage λ and functions in recombination and repair of the viral chromosome (5–10). When a double-stranded break occurs, one of the first steps taken to repair the damage is the formation of long 3' single-stranded overhangs. These are created by the action of λ -exonuclease and then are processed by other proteins such as *RecA* to give a recombinant DNA molecule (11).

The crystal structure of λ -exonuclease shows the protein to be a homotrimer with a large, tapered channel that runs parallel to the threefold axis of symmetry (12). The channel appears large enough to accommodate the DNA substrate, providing a structural rationale for the processivity of the protein (12). The overall fold of λ -exonuclease is not similar to any other protein in the protein database. There is, however, a subdomain of the protein that structurally corresponds to parts of the type II restriction endonucleases *PvuII* and *EcoRV* (12).

Type II restriction endonucleases are dimeric Mg^{2+} -dependent enzymes that recognize and cut a specific 4- to 8-bp palindromic DNA sequence (13, 14). In conjunction with a methylase that recognizes the same DNA sequence, these two proteins make up a bacterial Restriction Modification System that is thought to protect the bacterium from foreign DNA, such as invading viral DNA. To date, there are crystal structures of five type II restriction endonucleases: *PvuII* and *EcoRV*, which cleave DNA to give blunt ends, and *EcoRI*, *BamHI*, and *Cfr10I*, which cut DNA to give recessed ends (15–20). There is very little homology at the primary sequence level between restriction endonucleases. However, as pointed out by Venclovas *et al.* (21), at the structural level these five enzymes share a common motif consisting of five β -strands sandwiched between two α -helices. The blunt-end cutters *EcoRV* and *PvuII* are more similar to each other than to the recessed-end cutters *EcoRI*, *BamHI*, and *Cfr10I* (15, 20).

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.

© 1998 by The National Academy of Sciences 0027-8424/98/957893-5\$2.00/0
PNAS is available online at <http://www.pnas.org>.

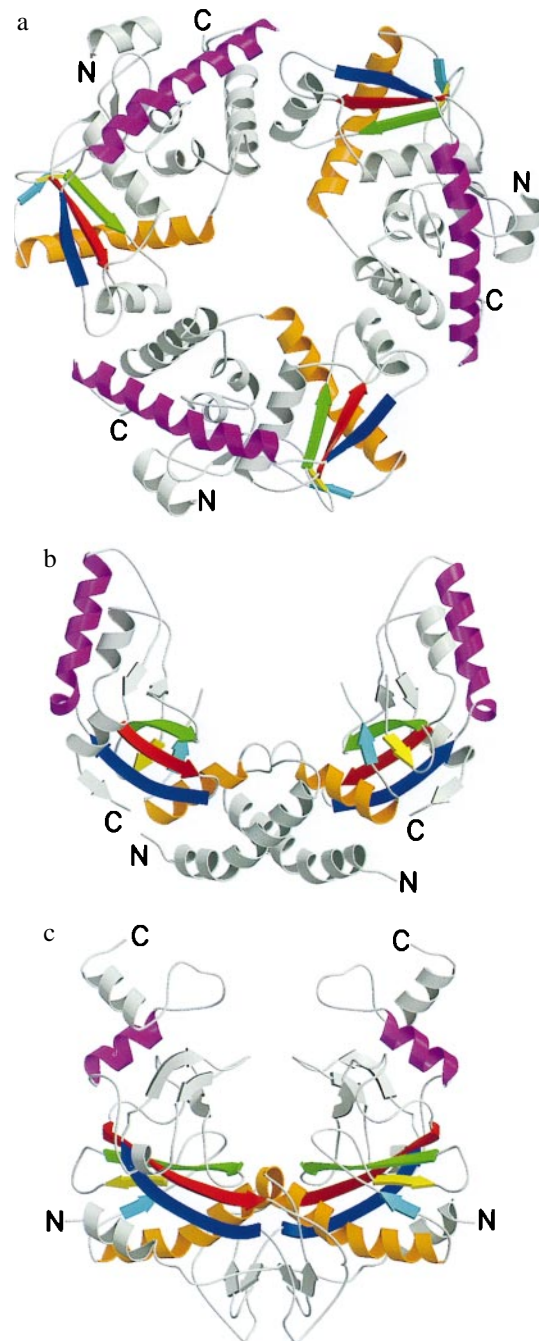


FIG. 1. Ribbon diagrams of λ -exonuclease (a), *PvuII* (b), and *EcoRV* (c) with corresponding secondary structure elements between the three proteins colored similarly. Figure created with MOLSCRIPT (26) and rendered in RASTER3D (27, 28).

*To whom reprint requests should be addressed. e-mail: brian@uoxray.uoregon.edu.

Table 1 Structural relationships between λ -exonuclease and type II restriction enzymes

	<i>PvuII</i> (15, 16)	<i>EcoRV</i> (17)	<i>EcoRI</i> (18)	<i>BamHI</i> (19)	<i>Cfr10I</i> (20)
Structurally equivalent residues	78	90	51	63	74
Deviation between C α atoms, Å	2.7	4.5	3.5	4.1	3.0
Chemically identical residues	12	10	2	8	6
Corresponding active-site residues					
λ -exonuclease: Glu-85	Glu-55	Glu-45	Asp-59	Glu-77	Glu-71
λ -exonuclease: Asp-119	Asp-58	Asp-74	Asp-91	Asp-94	Asp-134
λ -exonuclease: Glu-129	Glu-68	Asp-90	Clu-111	Glu-111	Glu-204
λ -exonuclease: Lys-131	Lys-70	Lys-92	Lys-113	Glu-11	Lys-190
Structurally equivalent residues					
λ -exonuclease: α E	83–98, 33–48	74–97, 34–57	79–97, 33–51	80–97, 56–73	74–96, 60–82
λ -exonuclease: β a	99–103, 49–53	98–106, 61–69	98–109, 53–64	100–103, 74–77	98–103, 89–94
λ -exonuclease: β c	—	—	111–117, 83–89	—	113–118, 128–133
λ -exonuclease: β d	121–124, 58–61	116–124, 71–79	119–124, 94–99	117–124, 92–99	119–123, 134–138
λ -exonuclease: β e	125–135, 66–76	125–132, 86–93	125–131, 104–110	125–133, 107–115	124–133, 183–192
λ -exonuclease: loop	—	134–137, 94–97	—	—	—
λ -exonuclease: α G	—	—	—	149–165, 118–134	148–166, 195–213
λ -exonuclease: loop	163–166, 91–94	165–168, 125–128	—	166–169, 135–138	—
λ -exonuclease: β f	167–177, 96–106	169–176, 130–137	—	170–174, 139–143	168–175, 227–234
λ -exonuclease: β g	183–191, 109–117	183–189, 167–173	—	—	—
λ -exonuclease: α H	192–207, 120–135	192–202, 209–219	—	—	—
λ -exonuclease: α H	—	205–210, 220–225	—	—	—

Each of the proteins listed was compared with λ -exonuclease using the Dali database server (22). The table gives the number of equivalent residues, the rms deviation, and the number of structurally equivalent residues that are chemically identical in the pairwise comparison of λ -exonuclease with each enzyme. The segments of structure that correspond are identified, with the residues from λ -exonuclease given first, followed by the structurally equivalent residues in the restriction enzyme. In addition to the overall regions of similarity, the server identified residues 154–157 in λ -exonuclease and residues 80–83 in *PvuII* and residues 153–156 in λ -exonuclease and residues 106–109 in *EcoRV* as equivalent. However, by visual inspection, it was clear that these short segments are not part of the overall correspondence, and they therefore were deleted from the analysis. Similar short segments were identified by Dali in the pairwise comparison of λ -exonuclease with *EcoRI*, *BamHI*, and *Cfr10I* and also were omitted.

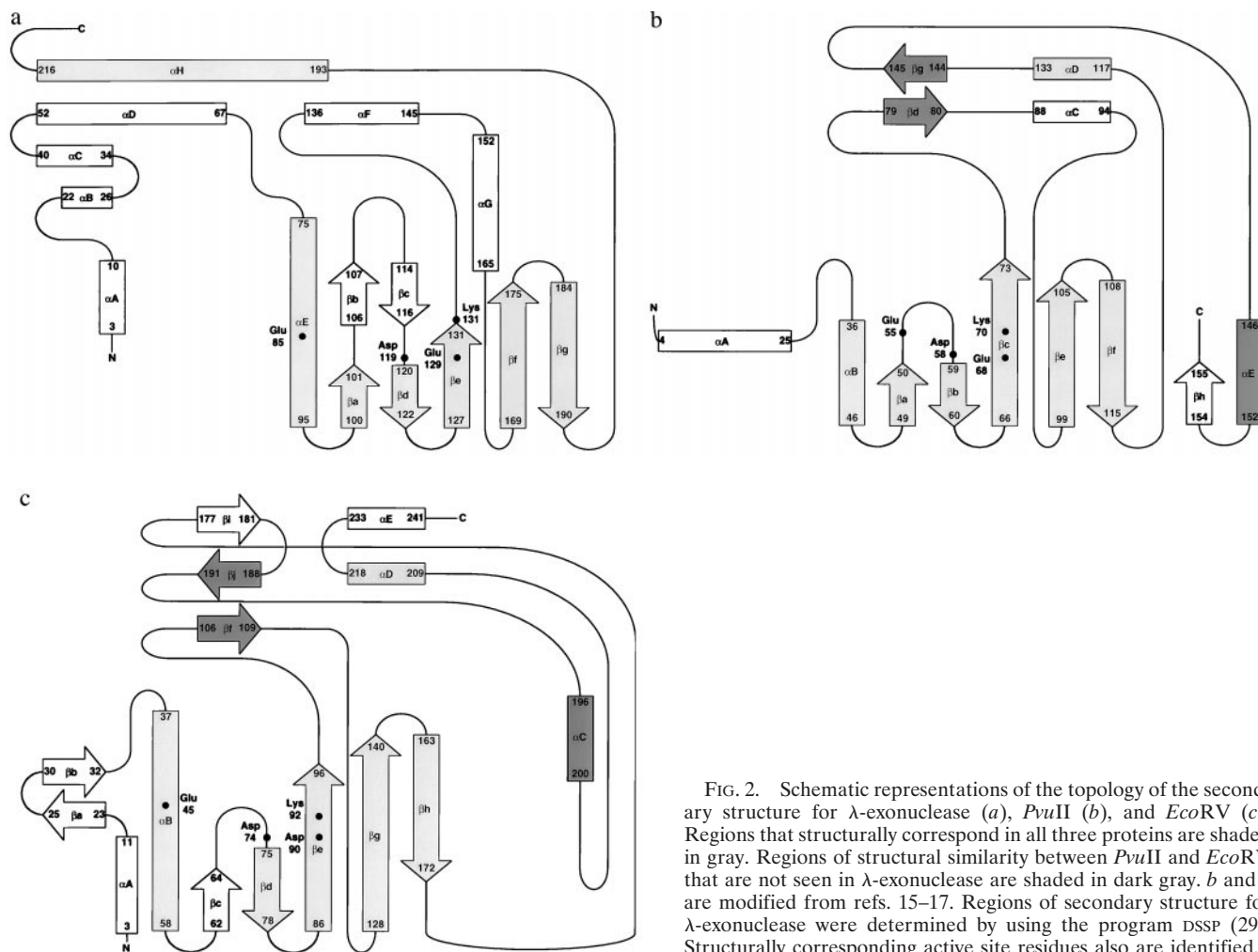


FIG. 2. Schematic representations of the topology of the secondary structure for λ -exonuclease (a), *PvuII* (b), and *EcoRV* (c). Regions that structurally correspond in all three proteins are shaded in gray. Regions of structural similarity between *PvuII* and *EcoRV* that are not seen in λ -exonuclease are shaded in dark gray. b and c are modified from refs. 15–17. Regions of secondary structure for λ -exonuclease were determined by using the program DSSP (29). Structurally corresponding active site residues also are identified.

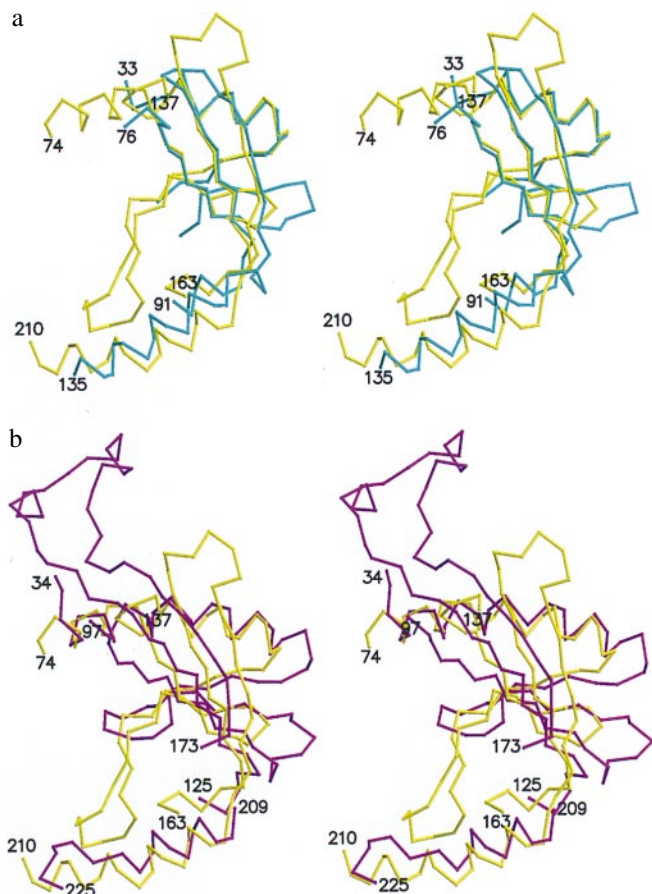


FIG. 3. Stereoview overlay of C^{α} atoms for λ -exonuclease (yellow) and *PvuII* (cyan) protein fragments (a) and λ -exonuclease (yellow) and *EcoRV* (magenta) protein fragments (b). The segments shown include all of the regions of structural similarity identified by the Dali database server (Table 1). Figure prepared by using MOLSCRIPT (26) and rendered with RASTER3D (27, 28).

Structural Similarity Between λ -Exonuclease and the Type II Restriction Enzymes. The structural relationship between λ -exonuclease (Brookhaven Data Bank ID Code 1AVQ) and the type II restriction endonucleases *EcoRV* and *PvuII* was identified by using the Dali structural database server (22) and is illustrated in Fig. 1. There are 78 equivalent C^{α} positions in *PvuII* and λ -exonuclease with a rms deviation of 2.7 Å. Twelve of the structurally equivalent residues are chemically identical (Table 1). Similarly, Dali identified 90 equivalent C^{α} positions in *EcoRV* and λ -exonuclease with an rms deviation of 4.5 Å. Ten of the 94 residues are identical.

The region of homology among λ exonuclease, *EcoRV*, and *PvuII* consists of five β -strands and two α -helices (Figs. 1–3). The region of correspondence begins with helix αE and strands βa , βd , and βe in λ -exonuclease. These are homologous to helix αB and strands βa , βb , and βc in *PvuII* and to helix αB and strands βc , βd , and βe in *EcoRV*. There is then an interruption in the structural correspondence between λ -exonuclease and both *EcoRV* and *PvuII*. The similarity begins again with strands of βf , βg , and helix αH . These regions are similar to βe , βf , and αD in *PvuII* and βg , βh , and αD in *EcoRV*.

Helix αB in *PvuII* and *EcoRV* (shaded orange in Fig. 1) forms parts of the dimer interface of these two proteins. However, the homologous helix in λ -exonuclease, αE , is located on the “back” of the protein (Fig. 1) and does not participate in any monomer–monomer contacts. The interface between monomers in λ -exonuclease is formed by helix αF and

a loop region (residues 175–184) between strands βf and βg . Because of its homology to helix αB in *PvuII* and *EcoRV* and its relatively exposed location, we speculate that helix αE in λ -exonuclease may be involved in other protein–protein interactions. In particular, *in vitro* data suggest that λ -exonuclease forms a complex with another bacteriophage λ -encoded protein termed “ β ” that also is involved in phage recombination (4). We suggest that the interface formed between the two proteins might include helix αE of λ -exonuclease.

It previously has been noted that, notwithstanding the similarity in the region of their active sites, the dimerization domains of *PvuII* and *EcoRV* differ substantially (16). As shown in Fig. 1, the present comparisons expand and extend this pattern of behavior. Not only are different regions of λ -exonuclease used to make the subunit–subunit interfaces but the character of these interfaces changes from “face-to-face” dimerization in the restriction enzyme to “head-to-tail” association in the λ -exonuclease trimer.

Relationships with *EcoRI*, *BamHI*, and *Cfr10I*. λ -exonuclease is a more distant structural relative to the recessed-end cutters *EcoRI*, *BamHI*, and *Cfr10I* than it is to *EcoRV* and *PvuII*. These type II restriction endonucleases were not identified as structural homologues by the Dali database search. They do, however, share the motif of five β -strands flanked by two α -helices that is common to λ -exonuclease, *PvuII*, and *EcoRV*. The Dali server (22) was used again to compare λ -exonuclease with *EcoRI*, *BamHI*, and *Cfr10I* (Table 1). In distinction to all the other comparisons, Dali identified helix αG of λ -exonuclease as corresponding to helix $\alpha 7$ of *Cfr10I* and helix $\alpha 4$ of *BamHI*. The correspondence between αG in λ -exonuclease and $\alpha 4$ in *BamHI*, however, is much less compelling than that of αG in λ -exonuclease with $\alpha 7$ in *Cfr10I*. Bozic *et al.* (20) have speculated that the helix $\alpha 7$ in *Cfr10I* may be involved in interactions with the DNA. This argument is based in part on the structural correspondence to helix $\alpha 4$ of *EcoRI*. Helix $\alpha 4$ of *EcoRI* does not, however, correspond structurally with helix αG of λ -exonuclease. Whether this helix in λ -exonuclease is important for protein–DNA interaction is not known.

Correspondence of Active Site Residues. Based on the above structural similarity, the residues Asp-119, Glu-129, and Lys-131 in the active site of λ -exonuclease can be identified as corresponding, respectively, to Asp-74, Asp-90, and Lys-92 in *EcoRV* and to Asp-58, Glu-68, and Lys-70 in *PvuII*. There are also analogous active-site residues for *EcoRI*, *BamHI*, and *Cfr10I* (Table 1). By using the program EDPDB (23), we made a comparison by overlaying all of the atoms of residues Asp-119, Glu-129, and Lys-131 in λ -exonuclease with all of the atoms in residues Asp-58, Glu-68, and Lys-70 in *PvuII*. A related comparison was made with Asp-74, Asp-90, and Lys-92 in *EcoRV* (Fig. 4A). To compare Glu-129 in λ -exonuclease with Asp-90 in *EcoRV*, the C_{γ} atom of Glu-129 was excluded. The rms discrepancy for the 26 atoms involved was 0.95 Å for *PvuII* and 1.17 Å for *EcoRV*.

It has been shown for all six proteins that two of the acidic active site residues form a binding site for the catalytically essential magnesium ion (12, 15–20). In the co-crystal structure of *EcoRV* bound to its specific DNA sequence, under some conditions, a second metal binding site is formed by Asp-90 and Glu-45 (24). In λ -exonuclease, there is a glutamic acid (Glu-85) that structurally corresponds to Glu-45 of *EcoRV*. Whether a second magnesium binding site is formed and, if so, whether it participates in catalysis remain uncertain. In general terms, however, the correspondence between the active sites strongly suggests that the mechanism of catalysis of λ -exonuclease is very similar to those of *PvuII* and *EcoRV*.

Implications for the Mechanism of Action of λ -Exonuclease. Several aspects of the present work have specific implications for the mechanism of action of λ -exonuclease. First, the structure comparisons confirm the presumed location of the

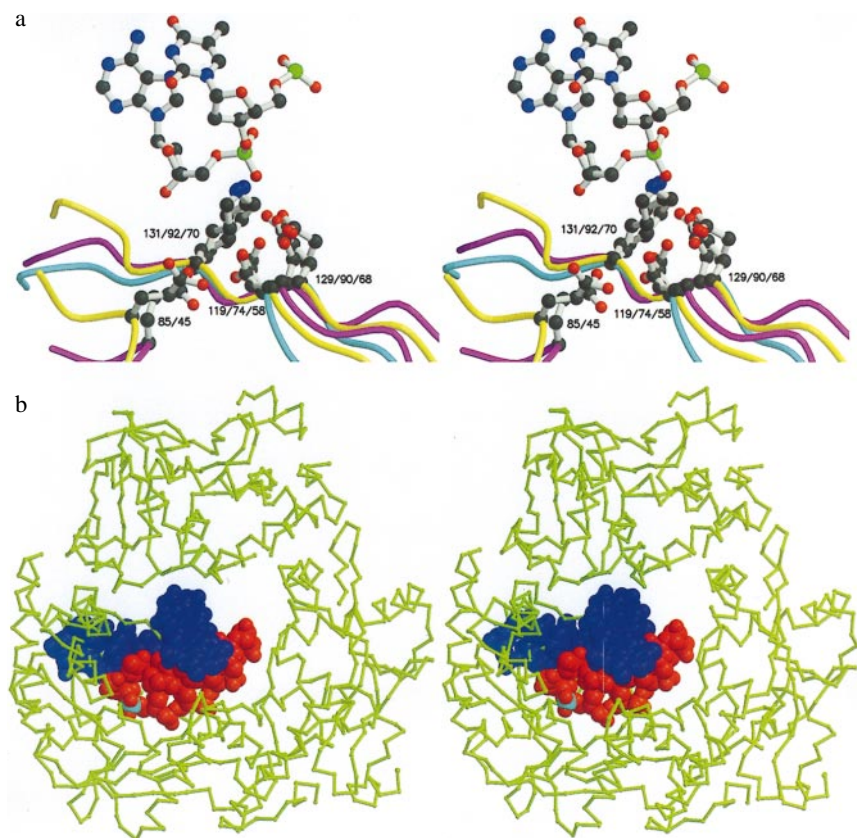


FIG. 4. (A) Stereoview showing the superposition of the active site residues of λ -exonuclease (yellow), (Glu-85, Asp-119, Glu-129, and Lys-131), *PvuII* (cyan) (Asp-58, Glu-68, and Lys-70), and *EcoRV* (magenta) (Glu-45, Asp-74, Asp-90, and Lys-92). The figure includes the central nucleotide (TA) from the co-crystal structure of *EcoRV* bound to its cognate DNA sequence (17). (B) Superposition on the trimeric structure of λ -exonuclease of the cognate DNA decamer from the *EcoRV*:DNA complex (17). To make the DNA resemble the substrate of λ -exonucleases, five nucleotides were removed from the 5' end of the strand that is hydrolyzed. The next phosphate that would be attacked is shown in cyan. The DNA was aligned based on the correspondence between the active site residues of *EcoRV* and λ -exonuclease (A). In the stereoview, the wider end of the central channel through the λ -exonuclease trimer, i.e., the end of the channel that is thought to accept the double-strand DNA substrate, is toward the viewer. Figure created by using MOLSCRIPT (26) and rendered with RASTER3D (27, 28).

active site (12) and show it to be on the inside of the toroidal trimer. Second, the comparisons are consistent with the presumed alignment of the hydrolyzed DNA strand. Fig. 4A includes the central dinucleotide (TA) from the *EcoRV*:DNA complex aligned within the active site of λ -exonuclease, based on the structural similarity of the two proteins. The location and mode of binding are as expected for cleavage by λ -exonuclease in the 5' to 3' direction. Extending this result, we took the entire DNA decamer as seen in complex with *EcoRV* (17), trimmed part of one strand so as to correspond to the substrate for λ -exonuclease, and used the structural alignment between *EcoRV* and λ -exonuclease to place this segment of DNA on the structure of λ -exonuclease. In this superposition (Fig. 4B), the double-strand part of the DNA is located within the wider end of the channel that passes through the λ -exonuclease toroid and the single-strand end of the DNA is located within the narrow end of the channel. This alignment of the DNA, determined by superposition of corresponding structural elements of λ -exonuclease and *EcoRV*, is very similar to that suggested by consideration of the λ -exonuclease structure alone (12). It provides independent support for the idea that the λ -exonuclease toroid encircles its DNA substrate during hydrolysis.

The structural homology between λ -exonuclease and the type II restriction endonucleases *PvuII*, *EcoRV*, *EcoRI*, *BamHI*, and *Cfr10I* also suggests that each of these proteins diverged from a common structural ancestor. Recently, Ban and Yang (25) showed that the similarity between *PvuII* and *EcoRV* extends to include the restriction enzyme *FokI* and the monomeric sequence-specific endonuclease *MutH*. It is ironic that the type II restriction endonucleases, which protect bacteria by hydrolyzing invading viral DNA, and an exonuclease that is involved in recombination and repair of the same phage DNA, both appear to have evolved from a common structural ancestor.

We thank Dr. Michael Quillin for comments on the manuscript and Doug Juers for advice on the structural comparisons. This work was

supported in part by National Institutes of Health Grant GM20066 to B.W.M.

- Linn, S. M., Lloyd, R. S. & Roberts, R. J., eds. (1993) *Nucleases* (Cold Spring Harbor Lab. Press, Plainview, NY).
- Radding, C. M. (1966) *J. Mol. Biol.* **18**, 235–250.
- Little, J. W. (1967) *J. Biol. Chem.* **242**, 679–686.
- Carter, D. M. & Radding, C. M. (1971) *J. Biol. Chem.* **246**, 2502–2512.
- Cassuto, E. & Radding, C. M. (1971) *Nat. N. Biol.* **229**, 13–16.
- Cassuto, E., Lash, T., Sriprakash, K. S. & Radding, C. M. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 1639–1643.
- Stahl, F. W., Kobayashi, I. & Stahl, M. M. (1985) *J. Mol. Biol.* **181**, 199–209.
- Stahl, F. W. & Stahl, M. M. (1985) *J. Genet.* **64**, 31–39.
- Thaler, D. S., Stahl, M. M. & Stahl, F. W. (1987) *J. Mol. Biol.* **195**, 75–87.
- Takahashi, N. & Kobayashi, I. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2790–2794.
- Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. (1994) *Microbiol. Rev.* **58**, 401–465.
- Kovall, R. & Matthews, B. W. (1997) *Structure* **277**, 1824–1827.
- Aggarwal, A. K. (1995) *Curr. Opin. Struct. Biol.* **5**, 11–19.
- Pingoud, A. & Jeltsch, A. (1997) *Eur. J. Biochem.* **246**, 1–22.
- Athanasiadis, A., Vlassi, M., Kotsifaki, D., Tucker, P. A., Wilson, K. S. & Kokkinidis, M. (1994) *Nat. Struct. Biol.* **1**, 469–475.
- Cheng, X., Balendiran, K., Schildkraut, I. & Anderson, J. E. (1994) *EMBO J.* **13**, 3927–3935.
- Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K. & Wilson, K. S. (1993) *EMBO J.* **12**, 1781–1795.
- Kim, Y., Grable, J. C., Love, R., Greene, P. J. & Rosenberg, J. M. (1990) *Science* **249**, 1307–1309.
- Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I. & Aggarwal, A. K. (1994) *Nature (London)* **368**, 660–664.
- Bozic, D., Grazulis, S., Siksnys, V. & Huber, R. (1996) *J. Mol. Biol.* **255**, 176–186.
- Venclovas, C., Timinskas, A. & Siksnys, V. (1994) *Proteins* **20**, 279–282.

22. Holm, L. & Sander, C. (1993) *J. Mol. Biol.* **233**, 123–138.
23. Zhang, X.-J. & Matthews, B. W. (1995) *J. Appl. Cryst.* **28**, 624–630.
24. Kostrewa, D. & Winkler, F. K. (1995) *Biochemistry* **34**, 683–696.
25. Ban, C. & Yang, W. (1998) *EMBO J.* **17**, 1526–1534.
26. Kraulis, P. J. (1991) *J. Appl. Cryst.* **24**, 946–950.
27. Bacon, D. J. & Anderson, W. F. (1988) *Mol. Graphics* **6**, 219–220.
28. Merritt, E. A. & Murphy, M. E. P. (1994) *Acta Cryst. D* **50**, 869–873.
29. Kabsch, W. & Sander, C. (1983) *Biopolymers* **22**, 2577–2637.