How restriction enzymes became the workhorses of molecular biology

Richard J. Roberts*
New England Biolabs, 32 Tozer Road, Beverly, MA 01915

Restriction enzymes have proved to be invaluable for the physical mapping of DNA. They offer unparalleled opportunities for diagnosing DNA sequence content and are used in fields as disparate as criminal forensics and basic research. In fact, without restriction enzymes, the biotechnology industry would certainly not have flourished as it has. The first experiments demonstrating the utility of restriction enzymes were carried out by Danna and Nathans and reported in 1971. This pioneering study set the stage for the modern practice of molecular biology in which restriction enzymes are ubiquitous tools, although they are often taken for granted.

Today, it is difficult to imagine a time when our laboratory freezers were not well stocked with restriction enzymes, when DNA sequencing was not possible, or when genes were only accessible to the geneticists and could not be simply cloned out by recombinant DNA technology. Yet, in December 1971, a key paper appeared in PNAS that set the stage for much of what is now routine (1). In that paper, Kathleen Danna and Daniel Nathans of Johns Hopkins University (Baltimore) showed for the first time that the restriction enzyme called "endonuclease R," discovered by Hamilton Smith and Kent Wilcox (2), could be used to produce specific fragments of simian virus 40 (SV40) DNA. Moreover, the authors showed that these fragments could be nicely separated from one another by electrophoresis on a polyacrylamide gel. The resulting picture (Fig. 1) provided an immediate visual example of just how powerful the combination of restriction endonucleases and gel electrophoresis would be. Earlier that year, I had been fortunate to listen to a seminar given by Nathans at Harvard Medical School (Boston) and immediately began to think of the possibilities. It was a defining moment in my life when I realized that my half-formed plans for future research would be dropped and a new avenue pursued. It was because of this presentation that I developed my own lifelong passion for restriction endonucleases.

Looking back at the Danna and Nathans paper today, one is struck by the simplicity and elegance of the experiments. As with all great pioneering work, one can say, "But how obvious!" Yet, at the time, Smith, who had discovered and characterized endonuclease R, had not immediately recognized the value of an enzyme that could cleave DNA specifically. It was Nathans who made the key intuitive leap and then went on to demonstrate not only that the resulting fragments could be used to produce a physical map of SV40 (3), but also that this physical map allowed the mapping of the origin of replication (4) and the location of SV40 genes (5). These pioneering studies set the stage for modern molecular biology. Suddenly, everyone wanted to map DNA and use any available restriction enzymes to examine their favorite genome (in those days, usually a phage or viral DNA). When I moved from Harvard to Cold Spring Harbor Laboratory (Cold Spring Harbor, NY) in 1972, I first purified the few known restriction endonucleases and then began looking for more, ideally ones that would recognize different sequences and so permit the many applications that we routinely see today. Because of the ubiquitous distribution of these enzymes, we were successful beyond our wildest dreams (6).

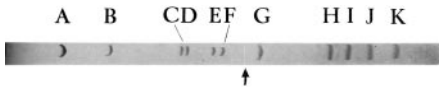

Fig. 1. Radioautogram of ^{14}C-labeled SV40 DNA cleaved with endonuclease R showing the 11 distinct fragments (figure 3 from ref. 1; courtesy of the Nathans family and Kathleen Danna).

Restriction Enzymes

The phenomenon of restriction and modification was first observed genetically in 1952–1953 by Luria and Human (7) and Bertani and Weigle (8), although they referred to it as host-induced, or host-controlled, variation. The authors observed that several different bacteriophages varied in their ability to grow on different host strains. However, once growth was achieved on one strain, the phages could continue to grow happily on this strain but were now restricted in their ability to grow on other strains. It was not until the 1960s that a theory to explain this phenomenon was proposed and then biochemically demonstrated by Werner Arber and his laboratory (summarized in ref. 9). Simultaneously, Matt Meselson and Bob Yuan also isolated a restriction enzyme from *Escherichia coli* K (10). The systems that Arber and Meselson characterized are now known as the type I systems. Although these enzymes recognize specific DNA sequences, they have the unfortunate property of cleaving DNA randomly, thus rendering the enzymes unsuitable for use as cloning and mapping reagents.

A significant breakthrough came in 1970 when the first of two papers from Smith's laboratory described an enzyme, endonuclease R, that was able to cleave bacteriophage T7 DNA into specific fragments (2). This was the first type II restriction enzyme, the sort that now populates our freezers, because it recognize specific sequences and also gives rise to very specific cleavage. Smith had been looking for an enzyme that might be involved in site-specific recombination in *Haemophilus influenzae* and thought at first that endonuclease R might be his long-sought quarry. With Tom Kelly, he went on to determine the DNA sequence recognized by endonuclease R and reported it as GTY ↓ RAC (11). This sequence seemed too short for a recombination enzyme, and during correspondence with his close friend Nathans, who ran the neighboring laboratory but was away on sabbatical, it became clear that this enzyme might have very practical uses for the analysis of DNA.

Polyacrylamide Gel Electrophoresis

Nathans realized that the sucrose gradients, which Smith had used to analyze the reaction products (Fig. 2), might not be the best way to try to characterize the specific fragments of DNA produced by cleavage with endonuclease R. The gradients simply lacked the resolution that would enable the fragments to be separated, characterized, and used for

This Perspective is published as part of a series highlighting landmark papers published in PNAS. Read more about this classic PNAS article online at www.pnas.org/classics.shtml.

Abbreviation: SV40, simian virus 40.

*E-mail: roberts@neb.com.

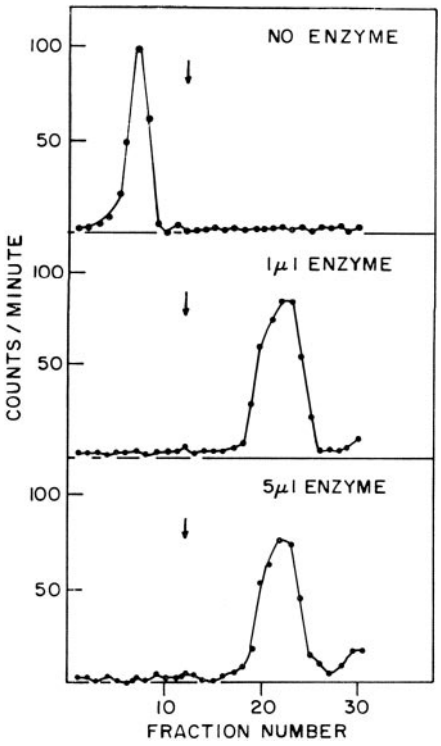

Fig. 2. Sucrose gradient analysis of ^{3}H-labeled SV40 DNA cleaved with endonuclease R (figure 1 from ref. 1; courtesy of the Nathans family and Kathleen Danna).

mapping. Nathans thus turned to another technique, polyacrylamide gel electrophoresis, whose use had been pioneered by Ulrich Loening (12) to separate RNA species. Danna, in Nathans' laboratory, quickly performed a simple experiment. She took a small amount of the purified SV40 DNA being studied in his laboratory, incubated it with some endonuclease R, and ran the resulting digest on one of Loening's polyacrylamide gels. In those days, most gels were prepared in glass tubing, not the slab gels that are common today. Once the gel had run, two methods could be used to find the resulting DNA, which in this case had been radioactively labeled. The first method was to expel the gel from the tube and place it in contact with an x-ray film for autoradiography. The result, using ^{14}C-labeled DNA, is shown in Fig. 1. The second method was to cut the gel into slices and use a scintillation counter to quantitate the radioactivity in each slice. The results of this kind of analysis, in this case, using DNA that had been labeled with ^{32}P, are shown in Fig. 3. One immediate advantage of autoradiography is readily apparent. The third and fourth peaks shown in Fig. 3 can be seen in the autoradiograph (Fig. 1) to each contain two separate fragments labeled CD and EF. The resolution necessary to separate them is simply not available by slicing gels. Danna and Nathans proceeded to work out very good length estimates for each of these fragments based on the percentage of the total SV40 DNA that was present and by using the known molecular mass of SV40 DNA of $\approx 3 \times 10^6$ kDa, which corresponds to ≈5,200 base pairs (13). The team very carefully compared sedimentation values, radioactivity measurements, and even EM length measurements to make sure that the results from each method were consistent.

Because the recognition sequence of endonuclease R was GTYRAC, Danna and Nathans expected cleavage to take place approximately once every 1,000 base pairs. It was a little surprising, therefore, that the 4,500 base pairs of SV40 DNA would be cut into 11 fragments. Danna and Nathans considered the possibility that the SV40 DNA was heterogeneous. However, the careful length measurements of the fragments precluded this possibility. We now know that it was, in fact, the original preparation of endonuclease R that was heterogeneous, because it contained not one restriction enzyme, but two. This discovery was made in several laboratories, including Smith's, as soon as DNAs other than T7 DNA were used as assay substrates. It turns out that, by chance, bacteriophage T7 DNA has recognition sites for only one of the two restriction enzymes present in the endonuclease R preparation. The original enzyme characterized by Smith is now called *Hin*dII (recognition sequence GTY ↓ RAC), whereas the second enzyme, for which there are no sites in T7 DNA, is *Hin*dIII. Its recognition sequence is A ↓ AGCTT (14), and this enzyme has six sites in SV40 DNA, rather more than one would have anticipated by chance.

Mapping DNA

Another feature of the Danna and Nathans paper that helped to make it a classic was that it clearly foresaw several potential applications of restriction endonucleases that later proved to be of considerable utility. For instance, the authors clearly saw the possibility of using these fragments to prepare a physical map of the SV40 genome, a feat that was later accomplished in Nathans' laboratory (3). The authors also showed that it was possible to localize the origin of replication (4) and to position the early and late genes of SV40 onto this "restriction map" (5), and that any individual gene could be mapped by testing for biological activity during transformation experiments (15). Even more insightful was the realization that informative mutants could be made by

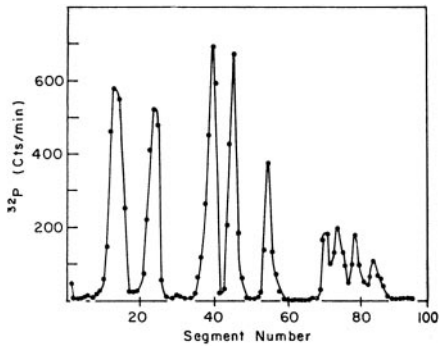

Fig. 3. Electrophoretic analysis of ^{32}P-labeled SV40 DNA cleaved with endonuclease R. After electrophoresis, individual slices from the gel were quantitated by scintillation counting (figure 2 from ref. 1; courtesy of the Nathans family and Kathleen Danna).

deleting one or more of the specific fragments (16). These deletions and others were easily visualized by restriction enzyme analysis because fragments were either missing completely or rendered shorter if the deletion was located within them. This feature quickly became a standard use of restriction enzyme maps. In a sentence that foreshadows the current diagnostic use of restriction endonuclease digestion, the authors noted: "By this means, we have found that the DNA of small-plaque, large-plaque, and minute-plaque SV40 strains show specific differences in the mobility of particular DNA fragments." The length variations, now known as restriction fragment length polymorphisms (RFLPs), found in human minisatellites and used so successfully by Alec Jeffreys for forensic purposes (17), are one of the applications known even to the general public.

Agarose Gels and More Restriction Endonucleases

One momentous feature of the paper was the realization that gel electrophoresis provided a wonderful assay by which one might hope to find new restriction endonucleases. That this turned out to be true is testified by the current collection of known restriction endonucleases, which now numbers >3,600 individual enzymes representing >250 different specificities (18). Of course, electrophoresis in tube gels was soon superseded by polyacrylamide slab-gel electrophoresis, again first introduced by Loening. Even more useful were the agarose gels first described in 1972 (19, 20) and the use of ethidium bromide to stain the DNA in them, which permitted nonradioactive DNA to be visualized (21). Initially, these agarose gels also were run in tubes, tapered at the end to stop the slippery agarose from sliding

out, and then run later on vertical slabs, when special frames were made to prevent slippage and intricate combs were developed to allow the even loading of samples (22). These gels were, in turn, replaced by the horizontal slab gels we run today (23). Indeed, it was the development of agarose slab-gel electrophoresis that enabled my laboratory in the early 1970s to isolate large numbers of restriction endonucleases. At first, we gladly shared these enzymes with the academic community, but the demand quickly became far too high. In 1975, New England Biolabs started to make the sale of these enzymes its major product line, and many other companies soon followed suit.

Sequencing DNA

Another important use for restriction enzyme fragments was in the early days of DNA sequencing. When Fred Sanger first developed methods for RNA sequencing, it was because there were many small RNA molecules such as tRNAs and 5S RNA on which to practice, which is always crucial when new methods are being developed. However, there were no comparable short DNA molecules. This situation changed when restriction fragments became available, and some of the early primers used by Sanger's laboratory were the tiny fragments produced when ϕX174 DNA was cleaved by restriction enzymes recognizing four base pairs (24). This methodology was underway in the early 1970s, and for a time my laboratory was continuously sending restriction enzyme samples to the Medical Research Council Laboratory of Molecular Biology (Cambridge, U.K.). The restriction enzyme maps also helped in the assembly of DNA sequences by providing useful landmarks and permitting manageable segments of DNA to be isolated and sequenced before assembly. The chemical sequencing methods developed by Wally Gilbert and Allan Maxam also depended heavily on restriction enzymes to provide the unique 5′ termini, which could be labeled with ^{32}P before chemical degradation (25).

Recombinant DNA and Biotechnology

There are many other developments that stem from this early work on restriction enzymes and SV40 DNA. Perhaps the most prominent development was the discovery of methods for preparing recombinant DNAs. When John Morrow and Paul Berg (26) discovered that SV40 DNA contained a single site for the restriction endonuclease *Eco*RI, which had been discovered in Herb Boyer's laboratory (27), the stage was set to convert SV40 DNA into a vector for recombinant DNA. The original method pioneered by Peter Lobban and Dale Kaiser (28) used bacteriophage λ as a vector, but Dave Jackson, Bob Symons, and Paul Berg (29) used the better method of adding poly(A) tails to an SV40 vector and poly(T) tails to the fragment to be cloned. Herb Boyer and Stan Cohen soon came up with the even better idea for making recombinant DNAs (30). They used DNA ligase to join a DNA molecule with "sticky" ends, produced by cleavage with *Eco*RI restriction endonuclease, to a plasmid DNA molecule that also had been cleaved by *Eco*RI, ensuring that the ends of each DNA molecule were complementary. This complementarity provided a very easy route to the preparation of recombinant DNA molecules and enabled any DNA to be cloned into the easily grown *E. coli*. From these early tools of the molecular biologists, we have now witnessed a multibillion-dollar biotechnology industry develop that can trace its origins back to the ideas embedded in the paper submitted by Danna and Nathans to PNAS in 1971. DNA diagnostics, DNA forensics, and the routine checking of DNA samples in the research laboratory all arose from the methods described in this classic paper.

Fig. 4. Dan Nathans (left) and Hamilton Smith (right) in the laboratory at Johns Hopkins University (Copyright 1978, Susie Fitzhugh).

The Nobel Prize

It is particularly fitting that when the Nobel Committee decided in 1978 who should be on the ticket for the discovery of restriction enzymes and their uses, they selected Nathans, who joined Arber and Smith (Fig. 4). Arber had provided the theoretical framework that described the biology of restriction and modification and had successfully isolated the very first type I restriction enzyme, *Eco*BI. Smith had discovered the first type II restriction enzyme, endonuclease R. But it was Nathans who first realized and then demonstrated just how powerful restriction enzymes could be when used for the physical mapping of SV40 DNA and its genes. This trio laid the groundwork that has led to the current addiction to restriction enzymes as routine, but essential, tools for molecular biologists.

1. Danna, K. & Nathans, D. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 2913–2917.
2. Smith, H. O. & Wilcox, K. W. (1970) *J. Mol. Biol.* **51,** 379–391.
3. Danna, K., Sacks, G. H., Jr., & Nathans, D. (1973) *J. Mol. Biol.* **78,** 363–376.
4. Nathans, D. & Danna, K. J. (1973) *Nat. New Biol.* **236,** 200–202.
5. Lai, C. J. & Nathans, D. (1975) *Cold Spring Harbor Symp. Quant. Biol.* **39,** 53–60.
6. Roberts, R. J. (1976) *CRC Crit. Rev. Biochem.* **4,** 123–164.
7. Luria, S. E. & Human, M. L. (1952) *J. Bacteriol.* **64,** 557–569.
8. Bertani, G. & Weigle, J. J. (1953) *J. Bacteriol.* **65,** 113–121.
9. Arber, W. (1965) *Annu. Rev. Microbiol.* **19,** 365–378.
10. Meselson, M. & Yuan, R. (1968) *Nature* **217,** 1110–1114.
11. Kelly, T. J., Jr., & Smith, H. O. (1970) *J. Mol. Biol.* **51,** 393–409.
12. Loening, U. E. (1967) *Biochem. J.* **102,** 251–257.

13. Crawford, L. V. & Black, P. H. (1964) *Virology* **24,** 381–387.
14. Old, R., Murray, K. & Roizes, G. (1975) *J. Mol. Biol.* **92,** 331–339.
15. Lai, C. J. & Nathans, D. (1974) *Virology* **60,** 466–475.
16. Lai, C. J. & Nathans, D. (1974) *J. Mol. Biol.* **89,** 179–193.
17. Jeffreys, A. J., Wilson, V. & Thein, S. L. (1985) *Nature* **316,** 76–79.
18. Roberts, R. J., Vincze, T., Posfai, J. & Macelis, D. (2005) *Nucleic Acids Res.* **33,** D230–D232.
19. Aaij, C. & Borst, P. (1972) *Biochim. Biophys. Acta* **269,** 192–200.
20. Hayward, G. S. & Smith, M. G. (1972) *J. Mol. Biol.* **63,** 383–388.
21. Sharp, P. A., Sugden, B. & Sambrook, J. (1973) *Biochemistry* **12,** 3055–3063.
22. Sugden, B., DeTroy, B., Roberts, R. J. & Sambrook, J. (1975) *Anal. Biochem.* **68,** 36–46.
23. McDonell, M. W., Simon, M. N. & Studier, F. W. (1977) *J. Mol. Biol.* **110,** 119–146.
24. Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* **94,** 441–448.
25. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
26. Morrow, J. & Berg, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 3365–3369.
27. Yoshimori, R. N. (1971) Ph.D. thesis (Univ. of California, San Francisco).
28. Lobban, P. E. & Kaiser, A. D. (1973) *J. Mol. Biol.* **78,** 453–471.
29. Jackson, D. A., Symons, R. H. & Berg, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 2904–2909.
30. Cohen, S. N., Chang, A. C. Y., Boyer, H. W. & Helling, R. B. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 3240–3244.