Proteins and Nucleic Acids: Restriction Endonucleases

The observation was made in the 1950s that bacteriophage cultured in one strain of bacteria easily infect other cultures of that strain but have difficulty infecting other strains of bacteria. The molecular basis for that observation, the system for modification and restriction of DNA, found in many microorganisms, was first described in 1962 (*1,2*). In this system, there are two activities. One activity recognizes specific DNA sequences and modifies those sequences in a non-destructive method. The second activity, the restriction activity, recognizes the same DNA sequence and destroys those sequences which are *not* modified. In this way, bacteria are able to mark and distinguish self-DNA from non-self DNA.

The restriction-modification system (the modification/restriction order is largely interchangeable) then functions as a small, two activity, immune system for bacteria who are largely threatened by genetic information (whether introduced by phage or not). In the original observation of phage infection, phage grown in one host are assembled from DNA modified by that host. Thus, in cultures of the same strain, their DNA passes as self and allows for infection. In other strains, with a different modification system, this phage DNA is correctly recognized as foreign and degraded. As it turns out, many microorganisms (mostly bacteria, no known multicellular organisms) have modification and restriction systems that differ in their specificity. Many organisms with modification and restriction systems have more than one, with two or three being common and five to seven marking the upper end of the range (*3*). With this many potential modification and restriction systems, several thousand modification and restriction systems have been described.

Modification and Restriction Systems

Nomenclature There is a standard method for nomenclature in modification and restriction systems (*3,4*). As an example, we can use a very famous system, *Eco*RI. The first letter of the system is drawn from the first letter of the genus name of the organism—*Escherichia*. Note both the capitalization and the italic suggesting that it is a genus name. The next two letters (lower case, italic) are drawn from the species name of the organism—*coli*. The capital letter (if present) is a designation of strain of organism in which this system is found (the strain is RY13, the thirteenth strain isolated by R. Yoshimori). For multiple modification and restriction systems in the same organism, Roman numerals are used to identify the particular modification and restriction systems from that strain and are assigned in the order in which the system was discovered. Thus, *Eco*RI is the first modification and restriction systems identified in *E. coli* strain RY13. As modification and restriction systems have two activities, you may see a either capital R or M in front of the standard name. This designation refers to either the restriction activity or the modification activity specifically. This bit of the nomenclature isn't particularly common unless you are studying the biological role of the system. Typically a component of these systems is used in isolation for molecular biology, and that is the most typical context in which you find mention of these systems.

Functions There are three distinct functions in the modification and restriction systems. The first is recognition. These systems recognize a small nucleotide sequence in DNA (typically 4 to 8 nucleotides) and act on, or around, these sequences. Those two actions are modification and restriction. Modification is accomplished by methylation of (typically) one base in the recognition sequence. The restriction function is accomplished by cleaving the phosophodiester bond in the sugar-phosphate backbone of the molecule (an endonuclease activity). This eliminates the information content in the genetic material without, necessarily, degrading the entire molecule to nucleotides. The general description of these

activities is necessarily a bit vague as there are several broad variations on modification and restriction systems. Each variation is identified as a type, and there are multiple examples of each type.

Types There are four types of modification and restriction systems (*3*). The types are designated as type I, II, IIs or III. The most common type is type II, and type II is also the most useful system to exploit for other purposes. After a brief description of each type, the remainder of the laboratory will focus on type II modification and restriction systems.

Type I modification and restriction systems are the largest and most complex enzymes. The three activities are three separate polypeptides than work in complex with one another to perform all three activities. In type I modification and restriction systems, the recognition site is not were the cleavage occurs, and the cleavage can occur at variable distances from the recognition site. Thus, the products of this reaction are variable and type I modification and restriction systems are not typically used in molecular biology applications.

Type II modification and restriction systems are the most common type (found in a wide variety of microorganisms), and they are also the most useful type. In type II modification and restriction systems, the methylase and endonuclease are two separate proteins which function independently. Both molecules have recognition activity and both act at or on the recognition sequence. The recognition sequences is symmetrical and can either be contiguous or separated a short sequence in which base composition is unimportant. Both modification and cleavage are symmetrical as well. Typically, the methyltransferase acts as a monomer methylating both strands in series. The endonucleases, however, function as homodimers and can cleave the phosphodiester bond in both strands in parallel, thus improving cleavage efficiency. Type II endonucleases can produces cleavages with either a single strand overlap (what are known as 'sticky ends') or fully duplexed DNA (these are known as 'blunt ends'). Either way, these endonucleases leave the $3'$ hydroxyl group and the $5'$ phosphate on the newly generated termini. These groups allow for the regeneration of the phosphodiester bond under the activity of a DNA ligase.

Type IIs modification and restriction systems are similar to the type II systems in that modification and restriction are the function of two separate activities. The major differences, however, are that the restriction enzyme is twice the size of the type II enzyme, recognize asymmetric sequences and functions as a monomer. These endonucleases cleave DNA at a site up to twenty bases distant from the recognition sequence and produce a staggered end.

Type III modification and restriction systems, specificity is conferred only by the modification activity. The methylase can act on its own while the endonuclease is dependent on dimerization with the methylase. The functional component of the restriction system is a heterodimer of both M and R activities. Like type IIs, these modification and restriction systems cleave the DNA at a site distant from the recognition sequence.

Of the four types, type II is the most regular and predictable and the endonuclease activity functions completely independently of the methylation activity. As a result, type II restriction endonucleases have been used since the about 1970 to specifically cut or digest DNA for a variety of applications. It should be noted that the Nobel Prize in Physiology or Medicine was awarded in 1978 for the discovery an application of restriction endonucleases to three individuals already cited in this introduction. Those individuals are Werner Arber, Daniel Nathans and Hamilton O. Smith. As is often the case, the discussion of the work of Nobel laureates could use a little simplification for the "average" reader. Below is the story told by Silvia Arber, about age ten in 1978, of her father's work (*5*).

The tale of the king and his servants

When I come to the laboratory of my father, I usually see some plates lying on the tables. These plates contain colonies of bacteria. These colonies remind me of a city with many inhabitants. In each bacterium there is a king. He is very long, but skinny. The king has many servants. These are thick and short, almost like balls. My father calls the king DNA, and the servants enzymes. The king is like a

book, in which everything is noted on the work to be done by the servants. For us human beings these instructions of the king are a mystery.

My father has discovered a servant who serves as a pair of scissors. If a foreign king invades a bacterium, this servant can cut him in small fragments, but he does not do any harm to his own king.

Clever people use the servant with the scissors to find out the secrets of the kings. To do so, they collect many servants with scissors and put them onto a king, so that the king is cut into pieces. With the resulting little pieces it is much easier to investigate the secrets. For this reason my father received the Nobel Prize for the discovery of the servant with the scissors.

Young Silvia's analogy about the molecular scissors describes very accurately the application of type II restriction endonucleases to the study of DNA. As we consider more features of the activity of modification and restriction systems, we will focus our attention on the type II restriction endonuclease.

Recognition Sites The recognition sites of type II restriction endonucleases are typically four or six nucleotide in length. These sequence are either continuous or separated by a short linker, they are symmetrical. The activity of type II restriction endonucleases are symmetrical as well. Figure 1 illustrates a small selection of type II restriction endonucleases recognition sequences as well as illustrating the cleavage activity of these enzymes.

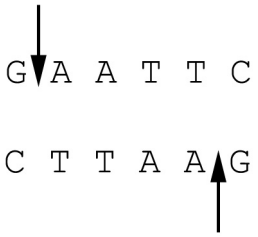

Figure 1: Recognition sequence of *Eco*RI. The arrows indicate the phosphodiester bond cleaved by R *Eco*RI. Note that this sequence is symmetrical, having a two-fold rotational axis of symmetry. The structure of the enzyme matches the structure of the substrate.

Pay particular attention to the symmetry of both the recognition sequence as well as the cleavage product. Recall that type II restriction endonucleases are homodimeric protein—two identical molecules come together to form a symmetrical dimer. That structure matches the symmetry of the substrate (the recognition sequence). The symmetry seen here is a two-fold axis of rotational symmetry. That is, if you could pin the recognition sequence in the middle, rotation about that point through one hundred eighty degrees would produce the same structure. The symmetry inherent in the recognition sequence is also in the enzyme allowing type II restriction endonucleases to be able make the cuts to both strands of the DNA in concert.There are more than thirty-five hundred characterized type II restriction endonucleases (*6*). A quick calculation would suggest that there aren't this many unique four nucleotide sequences and not many more unique six nucleotide sequences. Thus, you might expect that two or more modification and restriction systems recognize the same DNA sequence. That is indeed the case. For example, both EcoRI and FunII recognize the sequence `G/AATTC` and both enzymes cleave 3' of the guanosine (as indicated by the / in the sequence). These two enzymes are referred to as isoschizomers, meaning they have identical specificities (literally, they 'cut the same way'). In considering isoschizomers, some recognition sequences are found frequently (EcoRV `GAT/ATC`) while others are found far less often (XbaI, `TCT/AGA`). This allows for the preparation of DNA molecules from different sources with different enzymes to be easily 'pasted' together due to the complementarity of their ends.

Catalytic Cycle The catalytic cycle of the restriction endonuclease follows a three step mechanism (7). In locating the recognition sequence, restriction endonucleases must open their structure and encapsulate DNA. The restriction endonuclease has some low, non-sequence specific affinity for DNA. In this general interaction, the affinity appears to largely be due to the presence of a molecule of the size, shape and polarity of DNA. This non-specific binding releases few waters of hydration from either the DNA or the restriction endonuclease. Also, there appear to be few specific contacts made between the molecules.Following the enclosure of the DNA, the restriction endonuclease will scan the DNA to find a recognition sequence of greater binding affinity. Being tethered to the substrate allows for facilitated diffusion along the DNA molecule, greatly increasing the potential for finding a recognition sequence. There are several suggestions as to how the restriction endonuclease moves along the DNA that may involve any or all of the mechanisms of sliding, saccade (jumping) or intersegment transfer. Sliding (one nucleotide at a time) or jumping (several nucleotides at a time) are inherently one dimensional movement along the DNA axis. Intersegment transfer is made possible by the two DNA binding sites in each molecule. As one site remains secure to a non-specific binding site, the other is free to extend to another location. This might free the restriction endonuclease from an essentially one-dimensional search, allowing greater latitude and flexibility in the quest for a recognition sequence.The final element of the catalytic cycle is the chemical catalysis to hydrolyse a phosphodiester bond. The general mechanism, common to many type II restriction endonucleases begins with the activation of a water molecule by general acid base catalysis. The general base is most often a lysine side chain, but that side chain is not universally conserved. There appears to be an obligatory need for at least one divalent cation (typically Mg^{2++}) in this mechanism, likely involved in the coordination of the water molecule, one or more side chains and at least one phosphate group in the DNA backbone. The activated water/hydroxide anion then makes nucleophilic attack on the phosphorous leading to a pentavalent transition state. The final element of the catalytic cycle is the elimination of the 3' hydroxyl group and the reaction is complete. It is most important to note that this mechanism preserves the $3'$ hydroxyl group and the $5'$ phosphate as these groups are required for the regeneration of a phosphodiester bond by DNA ligase.

Star Activity With restriction endonucleases having low affinity for all DNA sequences and the catalytic mechanism involving only the activation of a water molecule, restriction endonucleases can, under certain conditions, cleave DNA with much lower specificity. This altered recognition specificity (for *Eco*RI, the sequence `T/AATTC` is recognized at about 1/10,000 the affinity of `G/AATTC` under ideal conditions, but there affinity of the enzyme for these two sites can be relatively equal under some 'extreme' solution conditions) leads to the cleavage of DNA at sites other than those expected by the typical recognition sequence. This is known as star activity and is usually abbreviated as *Eco*RI*. Star activity can be induced (or avoided) by controlling solution conditions. Recognition is dependent on ionic strength, divalent cation identity and concentration as well as the presence of cosolvents, such as glycerol which induced *Eco*RI* activity.

Basic Digestion Protocol

With purified, commercial enzyme, the exhaustive restriction digest is a simple protocol. The element combined are 1 μg of DNA, 5u of restriction endonuclease, 2 μg of acetylated bovine serum albumin (BSA is included to increase the protein concentration as some enzymes function poorly in very dilute concentrations) and the appropriate buffer for the enzyme.

For the first experiment, digest bacteriophage λ DNA with *Eco*RI. λ DNA is a large linear DNA that is often used as a set of molecular weight markers for gel electrophoresis following restriction digest. There are several common restriction endonucleases that cut λ DNA into a set of fragments of broad range (the classic set of broad molecular weight

markers is λ DNA cut with *Hin*DIII which produces fragments of 23,130, 9,416, 6,557 4,361, 2,322, 2,027, 564 and 125 base pairs. This digest will be the molecular weight markers in this experiment.). Digestion of λ DNA with *Eco*RI produces a similar set of fragments.

Table 1 below is the basic protocol for the digestion of λ DNA with *Eco*RI (buffer H is best buffer choice for *Eco*RI). Add each of the components to a sterile microcentrifuge tube in the order listed, gently mixing and pulsing in the microcentrifuge after each horizontal line.

Table 1: Classic restriction endonuclease protocol for the digestion of 1 μg of λ DNA with *Eco*RI in buffer H, 20 μL total volume.

Reagent	Volume (μL)
Sterile Water	14.0
10X Buffer H	2.0
Acetylated BSA, 10 μg/μL	0.2
λ DNA, 0.303 μg/μL	3.3
λ *Eco*RI, 10 u/μL	0.5

For an exhaustive digest (every recognition sequence cut) *Eco*RI will need at least one hour at 37 ^{o}C. Turn on the dry bath, and set the temperature to 37 ^{o}C (it will take 15 minutes to come to temperature equilibrium), and incubate your reactions for 1–3 hours (the temperature in the dry bath will fluctuate a bit during that hour—you do not need to be terribly concerned).

When your reaction is complete, prepare your sample for electrophoresis. Using the large gel apparatus, fit the gel tray to the casting block, level and lock in place. Prepare 200 mL of 1 % agarose gel in 1X TBE (5X TBE is 1.1 M Tris, 900 mM borate and 25 mM EDTA at pH 8.3) by gently heating in the microwave (*Caution: solution will be very hot*). Cool to less than 45 ^{o}C and add SYBR Safe DNA gel stain to 1X. Pour your gel, remove bubbles, insert the comb and allow to harden. When solid, move get to running apparatus, submerse in 1X TBE and run your gel; run time will be close to two hours. When the gel is complete, photograph your gel for analysis.

(Note: Don't do the above experiment—the data are already available)

Analysis The first element of the analysis of your data is to construct a standard curve relating DNA size in base pairs to the migration distance of each band for your molecular weight standards. To measure migration distance, measure from the center of the well to the center of each band in any units you find convenient (cm or pixels are the likely choices). The sizes of the DNA fragments from a *Hin*DIII digest of λ DNA are listed above. The standard curve is plotted as semi-log plot, where the fragment size in base pairs is plotted as a function of migration distance. There are two ways in which you can construct the semi-log plot. You can either plot the fragment size on a log scale or plot the $\log_{10}$ of the fragment size. Either way, your standard curve should approximate a straight line (the plot will have significant curvature if you don't make the semi-log plot).

When you have the standard curve, use it to determine the size of λ DNA in base pairs (*hint: there is a really easy way to calculate the size of λ DNA using only what you know thus far*). Now, λ DNA is larger than any of its fragments, but don't worry too much about it being outside the range of your standard curve.

Next, use the standard curve to determine the sizes of the fragments of λ DNA generated by *Eco*RI.

Questions

Question 1: How many recognition sequences for *Eco*RI (`G|AATTC`) are there in λ DNA?

Question 2: Given the choice, would you prefer to use *Eco*RI or *Hin*DIII cut λ DNA as a set of broad molecular weight markers? Why?

Question 3: Restriction mapping is the technique of locating the restriction sites on the DNA molecule that produced the bands you saw in your gel. Draw five (5) potential restriction maps for *Eco*RI on λ DNA. Can you definitely say that one of your maps is the correct one? Why or why not?

Question 4: Use Webcutter to verify your answers to Questions 1 and 3. When interpreting your results from the gel, remember that there will be measurement errors inherent in that analysis, and the number will be similar but will not match exactly to the base pair. How did you do?

Double Digests

The classic method for restriction mapping using exhaustive digests is the double digest. In the double digest, a DNA is cut with independently with two different restriction endonucleases. These two experiments allow you to determine the number of cut sites for each enzyme and the sizes of the fragments generated. The key to this experiment, then, is that both enzymes are used in concert in a double digest experiment—both enzymes are used in the same reaction and all sites are cut. In the results of the double digest, there are more bands than either of the single digests. However, some bands might be in common with either of the single digests. These data allow you to place the restriction sites in relative positions with respect to one another.

Prepare a restriction digest of λ DNA using *Eco*RI according to the basic protocol. In a seperate tube, prepare a restriction digest of λ DNA using *Hin*DIII. In a third tube, prepare an *Eco*RI and *Hin*DIII double digest in the restriction buffer of your choice (*hint: see the Restriction Enzyme Activity document for assistance*). Incubate these three digests at 37 ^{o}C for two hours.

There is a separate packet introducing two restriction mapping problems, one for a circular plasmid, the other for a short linear DNA. Produce restriction maps of both of these molecules, identifying the total size of the DNA and the positions of each restriction site.

Questions

Question 5: Which enzyme would you choose to digest λ DNA if you were interested in making molecular weight markers to use for your gels while constuction plasmids? (*hint: plasmids are typically 3500-6000 base pairs in length*)

Question 6: Describe your general method for solving these problems. A list of steps might be an easy method. What information comes from each experiment?

Question 7: Are you confident that your solution is the only possible solution? Explain

Reaction Conditions

The assumption of the double digest is that the two enzymes you choose to use together will both work optimally under the same set of solution conditions. When you purchase a restriction endonuclease, it is supplied with the ideal buffer. That buffer has been designed to produce the optimal activity of that enzyme. The major elements that vary between restriction endonuclease buffers is the pH (typically 7.5 or 7.9), NaCl concentration (0.05-1.5 M) and the concentration of the divalent cation Mg^{2+} (usually around 50 mM).

In addition to the unique buffers (such as Buffer H for *Eco*RI) there are many buffers which are designed to be used with a wide variety of enzymes and to provide a reasonable set of conditions for double digests. Using the buffer designations of Promega (Promega Corporation, Madison, WI. The other large supplier of restriction endonucleases is New England Biolabs, Ipswich, MA), the 'universal' buffers (4-CORE) are Buffers A, B, C & D. Promega also offers a MULTI-CORE buffer double digests and similar experiments.

Measure the efficiency of *Eco*RI toward λ DNA in each of the 4-CORE buffers and MULTI-CORE buffer using Buffer H as a control reaction. Prepare each of these six digests according to the prototype given in Table 1, substituting each of the other five buffers for Buffer H; the prototype reaction as written (with Buffer H) will serve as your positive control. In order to accentuate the differences in efficiency, use only 2 U of *Eco*RI activity in your digests. When your reactions are prepared, incubate them together at 37 ^{o}C for two hours. When your reactions are finished, prepare a sample of each for electrophoresis and run the gel. Collect an image of your gel and qualitatively assess the efficacy of digestions based on the intensity of the bands. Consider the digest performed in Buffer H to be your standard reference.

Questions

Question 8: If you were stranded on a desert island and your escape depended on making a restriction digest with *Eco*RI, which buffer would you choose to use, assuming your freezer was out of Buffer H?

Question 9: Which buffer would you choose for an *Eco*RI/*Hin*DIII double digest (*Hint: see the document "Relative Activity of Restriction Enzymes in Promega's 10X Buffers"*)? Why?

Question 10: Would you change your answer if you knew that *Eco*RI displays *Eco*RI* activity in MULTI-CORE buffer? Would you need to be concerned about *Eco*RI* activity in λ DNA?

Question 11: Which buffer combination would you use for and *Eco*RI/*Acc*III double digest? How would you account or compensate for the difference in assay temperatures?

Partial Digests

The restriction mapping problem can be made a good bit easier to solve (deterministic) if you are able to redesign your experiment. One of the ways in which you can do that is by performing a partial digest. A partial digest is a set of conditions under which the restriction endonuclease cuts each DNA molecule, on average, only once. Note that this is *on average* only once. Some molecules will be cut twice (or more) some molecules not be cut at all.

In order to shift the basic protocol from a complete digest to a partial digest, there a few things we can safely manipulate. One of those is simply the number of units of restriction endonuclease activity used in the experiment. The lower the enzyme concentration, the fewer cuts in the same unit of time. Another way to lower the enzyme efficiency is to cool the reactions below the ideal temperature of activity for your enzyme. Often a serial dilution of *Eco*RI is used on ice ($\approx$4 ^{o}C) to determine ideal conditions for the partial digest. Also, these reactions are run for significantly less than the 1–3 hours used for complete digests.

Question 12: Would you want to consider using a restriction endonuclease different buffer in this protocol? Explain.

A seperate packet details the restriction mapping of a cloned 4 kb fragment of DNA. Using that protocol and data, map the 4 kb segment of DNA for each of the three restriction

endonucleases using the data present in that question.

Question 13: Why was this method so much easier? Explain

Methylation Sensitive Restriction Endonucleases

There are restriction endonucleases that are not only specific for DNA sequence, but also methylation state. A classic example is DpnI which cleaves the DNA sequence GA|TC only when the adenine has been N^6-methylated. Enzymes of this form are often referred to as Type IIM restriction endonucleases.

The enzyme DpnI is a key player in a strategy for rapid site-directed mutagenesis. Site-directed mutagenesis involves changing one or more bases in a DNA sequence in order to change a codon, introduce a base (or bases) or remove a base (or bases). Typically, mutagenesis involves purification of intermediate states of your DNA constructs (and work over several days). The use of DpnI eliminates the purification step(s) and reduces the mutagenic protocols to a single day (sorry, you do still have to confirm that you've made what you were hoping for, and that'll still take several days).

The quintessential commercial expression of this method is the Aligent QuikChange method (originally marketed by Stratagene—corporate take–over). The QuikChange kit provides all the buffers and enzymes necessary, all you need to do is supply complimentary mutagenic primers of your own design. The mutagenic primers are used in a polymerase chain reaction that amplifies the entire plasmid. DpnI is then used to treat the products of that reaction.

Question 14: Using the illustration of the protocol, why does DpnI eliminate the native genotype prior to the cloning reaction?

Investigating the Structures

To begin with, let's load a structure of *Eco*RI into RasMol and look at some basics (*12*). When RasMol loads, two windows are created. The graphics windows is the large black one, the command line is the white window which opens minimized. Arrange RasMol so that you can see both windows (the command line only needs to be a couple lines high, the large the graphics window the better). Fetch the file 1ERI.pdb from the Protein Data Bank and save it in the same folder as the `RasMol.exe` file (*13,14*). Commands meant to be typed at the RasMol prompt will be given with the prompt and set in fixed-width font such as

```
RasMol>load 1eri.pdb
```

Alright, to begin, go ahead an run the above command. The command prompt should list some structure information and the structure of *Eco*RI and a short section of DNA will appear in the graphics window. You will need to rotate the graphics window (left click and drag) in order to see the DNA. To make the DNA a bit more obvious

```
RasMol>select *B
RasMol>spacefill
```

There, a nice view of *Eco*RI and DNA. What, you thought *Eco*RI was a homodimer? Well, it is. This structure is of the asymmetric unit, which is one polypeptide and one strand of DNA. You could generate the other half, but it will be easier to see the details with fewer atoms.

Let's begin with the active site. *Eco*RI has a sequence of `PD...EAK` at the active site which is a common motif in type II restriction endonucleases. The residue numbers are 90, 91...111, 112 and 113, with Glu 111 begin the general base at the active site.

```
RasMol>select 90-91, 111-113
```

```
RasMol>wireframe 40
RasMol>spacefill 120
```
This will draw the residues of the active site in a ball and stick representation. Notice how Glu 111 is very near the phosphate of the first A in the recognition sequence (*hint: when you click on an atom, RasMol will identify that atom in the command window*).

Question 15: Using the method mentioned above, what is the sequence of this DNA molecule (*hint: there are 13 nucleotides in this molecule*)?

Beyond the active site, we should focus on the large number of contacts between the protein and the DNA. Rotate the molecules until you are looking, more or less, down the long axis of the double helix. The first thing to note is that the structure of *Eco*RI wraps around the DNA and follows the curvature of the double helix. Also note that most of these contacts are with atoms of the backbone—atoms which are common to all DNA sequences.

Question 16: Does this observation support the suggestion that *Eco*RI has non-sequence specific affinity for DNA?

Role the molecule toward the inside of the double helix so that the sequence specific contacts are visible. You should see that *Eco*RI does make some contacts with bases in the major groove.

Question 17: To which end of the `GAATTC` sequence does *Eco*RI make the majority of its sequence-specific protein-DNA contacts? Does this help rationalize *Eco*RI* activity? What might you expect for a restriction endonuclease that doesn't exhibit appreciable star activity?

Finally let us turn our attention to the secondary and tertiary structure of the molecule.
```
RasMol>select *A
RasMol>spacefill off
RasMol>wireframe off
RasMol>cartoons
```
Question 18: To which top-level classification of the SCOP database does this structure belong?

Looking into the major groove, you can see that an α helix lies in that structure. This helix, instead of cutting across the major groove, runs largely parallel to the long axis of the DNA. In addition, one β strand also runs parallel to that helix and the DNA molecule. This strand contains the catalytic center, Glu 111.

Another notable feature of this structure is the parallel β sheet in the core of the protein. The parallel β-strand is not as common, but see how it is enabled by the bridging helices. *Eco*RI is 276 amino acids in length, this is just a bit shorter than the average protein in *E. coli*. These residues make a structure which is both complimentary to the shape of DNA as well as highly treacherous toward unmodified foreign kings.

References

(1) Arber W and Dussoix D. 1962. Host specificity of DNA produced by *Escherichia coli*: I. Host controlled modification of bacteriophage λ. *J Mol Biol.* **5:** 18-36.

(2) Dussoix D and Arber W. 1962. Host specificity of DNA produced by *Escherichia coli*: II. Control over acceptance of DNA from infecting phage λ. *J Mol Biol.* **5:** 37-49.

(3) Wilson GG and Murray NE. 1991. Restriction and modification systems. *Ann Rev Genet.* **25:** 585-627.

(4) Smith HO and Nathans D. 1973. A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. *J Mol Biol.* **81:** 419-423.
(5) Odelberg W, Ed.. 1979. **Les Prix Nobel**. Nobel Foundation, Stockholm.
(6) Roberts RJ, Vincze T, Posfai J and Macelis D. 2010. REBASE–a database for DNA restriction and modification: enzymes, genes and genomes. *Nucl Acids Res.* **38:** D234-D236.
(7) Pingoud A, Fuxreiter M, Pingoud V and Wende W. 2005. Type II restriction endonucleases: structure and mechanism. *Cell Mol Life Sci.* **62:** 685-707.
(8) Goldstein L and Waterman MS. 1987. Mapping DNA by stochastic relaxation. *Adv Appl Math.* **8:** 194-207.
(9) Perona JJ. 2002. Type II restriction endonuclease. *Methods.* **28:** 353-364.
(10) Wells RD, Klein RD and Singleton CK. 1981. **The Enzymes, 3rd Edition**. Acadmic Press, Waltham, MA.
(11) Sayle RA and Milner-White EJ. 1995. RASMOL: biomolecular graphics for all. *Trends Biochem Sci.* **20:** 374.
(12) Kim Y, Grable JC, Love R, Greene PJ and Rosenberg JM. 1990. Refinement of Eco RI endonuclease structure: a revised protein chain tracing. *Science.* **249:** 1307-9.
(13) Berman HM, Henrick K and Nakamura H. 2003. The protein data bank. *Nucleic Acid Res.* **28:** 235-42.

Prepared @ 09:11 on 01–17–2024.