

## Paper: Restriction Endonucleases

### 1. What is the significance of this research topic?

Restriction enzymes (REase) protect bacteria and archaea from infection by viruses DNA molecules. In 1952-1953, Luria and Human observed that bacteriophages have a different ability to grow on different host strains. Once established on one strain, the phages would grow easily on this strain but were “restricted” initially to grow on new strains. Arber and Dussoix postulated, that certain bacterial strains contain an endonuclease that is able to cleave DNA but also have a strain modification protecting them from their own endonuclease. Foreign DNA, such as the one from an infecting phage, is degraded by the endonuclease, restricting its ability to infect its host (hence the term “restriction endonuclease”). In some cases, a small portion of the phage DNA, is modified prior to degradation by the endonuclease. This modified DNA phage is able to successfully replicate, and infect other hosts, but since the new host does not contain the protecting system modification as the first initial host, the modified phage cannot replicate in the new host [1].

- **Class 1**, REase *EcoKI*, was the first to be purified. Enzymes of type I, are very large (with a molecular weight of about 300-400 kDa). They consist in 5 subunits of 3 different proteins, and require the cofactors  $Mg^{2+}$ , ATP and S-adenosyl methionine (SAM) for activity. Because these enzymes cuts DNA randomly far from their recognition sequence, they cannot be easily used as cloning reagents.
- In 1970's, a breakthrough came after two seminal papers from Smith's laboratory. Smith and al. described the properties of *Haemophilus influenzae* endonuclease, *HindII*, which, unlike *EcoKI*, requires only  $Mg^{2+}$  for activity. *HindII*, like any **type II enzyme**, cleaves the two strands of DNA at a fixed location into two fragments with “blunt” ends: for *HindII* within its recognition sequence (GTY|RAC) (Y=pyrimidine (C or T), R=purine (A or G)). An important feature of the cut is its palindromic nature, a nucleotide sequence on one strand segment is now complementary of segment on the other strand. Pieces of DNA cut with the same restriction enzymes would therefore all have the same length and could be joined to create hybrid DNA molecules. Type II enzymes are smaller (molecular weight about 100-20kDa) and have identical subunits.
- **Type III and IV** enzymes are rare. Type III enzyme is a large molecule (mol. weight ~ 200KDa), they do require  $Mg^{2+}$  and ATP for activation but not SAM. They cleave outside of their recognition sequences and require two DNA sequences within the same DNA in opposite directions, and rarely give complete digests. Type IV enzymes recognize methylated DNA.
- These days there are more than 3,600 enzymes representing more than 250 different specificities (latest numbers can be found at <http://rebase.neb.com/rebase/rebase.html>).

Restriction enzymes (REase) played a critical role in recombinant DNA technology (**rDNA**): when Boyer and Cohen used DNA ligase to join a DNA molecule with “sticky” ends produced by cleavage with *EcoRI* restriction endonuclease, to a plasmid DNA molecule also cleaved by *EcoRI*, they launched the era of recombinant DNA (rDNA). Numbers of discoveries have stemmed from gene cloning and rDNA. Enzymes themselves have been improved by rDNA. Every commercial

enzyme has been purified from *an E. coli*. Small quantities of genes are isolated from the desired proteins, moved to *E. coli*, transcribed, translated again into proteins at optimal rates, high yield and high levels of purity by avoiding contaminating factors [1].

The restriction enzymes have transformed biology and medicine, increasing significantly the understanding of many forms of life.

## **2. Who is working in this area?**

The companies like NEB (New England Biolabs) or Promega are still in activity today, they were among the first companies to discover, purify, produce restriction enzymes. With time they have accumulated an impressive catalog of restriction endonucleases but also PCR amplification products, nucleic acid purification products, RNA reagents, DNA assembly, cloning and mutagenesis kits and genome editing tools.

More recent new comers are: Life in Genomics, Arbor Sciences, Watchmaker Genomics and Eton Science. Other companies were founded to exploit REases to produce high value molecule-based therapeutics like Biogen which engineered the first “recombinant” vaccine for hepatitis B or Genentech which was able to create human insulin in *E. coli*.

In academia, restriction enzymes are ubiquitous in any biological or life science research.

Recently Fordyce and Herschlag labs created a high-throughput microfluidic kinetics (HT-MEK) chip that dramatically speeds up enzyme analysis. HT-MEK could accelerate allosteric targeting drugs, and allows to reverse-engineer enzymes and design for example enzymes that can degrade plastics into nontoxic pieces.

## **3. What methods are used to study the concepts described in the paper?**

Restriction Enzymes identifies their target sequence with very high accuracy, for example the 8-bp *NotI* enzyme binds to only one sequence: GCGGCCGC among the 65K existing DNA sequences of this length. They recognize and bind to short sequence of base pairs in DNA.

Binding between the protein and the DNA bases in the recognition site is held by 15-20 hydrogen bonds mediated through specific amino acids (ASP and GLU). Thus, the high specificity of the binding, is due to the amino-acids that form the binding-sites which accommodate, and even attract, the correct base pair at each binding position and will obstruct, and even repel, the incorrect base pair. Each recognition site can properly accommodate only one cognate base pairs. Once the cognate sequence is identified, water molecules are expelled, and conformational changes occur in both the enzyme and DNA. Under non optimal conditions, restriction enzymes can bind to next best site, which is referred to as **star activity**. Reaction parameters which increase star activity include pH, type of present ions, ionic strength, metal cofactors other than  $Mg^{2+}$ , high DNA/enzyme ratios and presence of volume excluders (glycerol, ethylene glycol, etc.). Cleavage rates with star activity also decrease. To this date, there are only postulations about how restriction enzyme cleaves the DNA to which it is bound. Two models have been proposed: the substrate-assisted catalysis and the two-metal-ion mechanisms in which  $Mg^{2+}$  and  $Mn^{2+}$  play significant roles [2][3].

#### 4. Does the review article lead to new questions or hypotheses in this technical area (by the authors, by other researchers)?

Concerns have been raised about the ethics of creating synthetic humans. At Asilomar conference in 1975 on rDNA, it was decided to continue rDNA work but appropriate rules in the form of physical and biological containment procedures needed to be define. More recently the consensus is that the research community needs to exert critical thinking about its work and be more open about the risk of the emerging technological advances. Regulation is needed but should not be overbearing so not to discourage financial investments.

#### 5. What are some practical applications of the research discussed in the article?

One of the first application of REases, was their use as molecular weight markers for gel electrophoresis; for example, the restriction enzyme *HindIII* digestion of *E. coli* produces 8 restriction fragments with precise molecular weights. Restriction enzymes also allowed the detection of homologous DNA sequences, or polymorphisms. Restriction length polymorphism (RFLP) analysis, which has been superseded by faster and cheaper techniques, was initially used in genome mapping, localization of genes for genetic diseases, and genetic fingerprinting. Restriction enzymes, *Mbo I* and *Hpa II*, played a critical role in the identification of CpG islands which are DNA methylation regions crucial for gene expression. Today single-nucleotides polymorphism (SNP) analysis is preferred, and sequence changes can also be analyzed faster with polymerase chain reaction (PCR). Terminal restriction fragment length polymorphism (TRFLP) detection, characterizes bacterial, fungi species or microbial communities; PCR products are digested using restriction enzymes.

Restriction enzymes (REase) have been the workhorse of molecular biology for [4]:

**Cloning:** initially Stanley Cohen and his colleagues incorporated foreign DNA into natural plasmids to create cloning-plasmid vectors that self-propagate in *E. coli*. Restriction enzymes have been useful tools in PCR for confirming that insertions have taken place successfully.

**DNA mapping:**

In the 1970s, Nathans mapped the functional units of the SV40 genome, by discovering that individual gene could be mapped onto the endonuclease “restriction maps”. They also found that genetic mutations could be detected using restriction endonuclease digestion. This work evolved into SNP and insertion/deletions (Indels) detection.

**Markers of epigenetic modifications:** Restriction Landmark Genome Scanning (RLGS) is a gel electrophoresis mapping techniques in which *NotI*, *Ascl*, *EagI* or *BssHII* enzymes identify changes in the methylation patterns of the human genome during the development of cancer cells.

**Synthetic biology:** aims to redesign, or create biological systems to enhance existing or create new abilities. Some DNA assembly technologies are based on REase-ligation methodology, and REases abilities to cleave DNA outside of the recognition sequence (exonucleases activity).

**DNA libraries:** the use of type IIS enzymes as tagging enzymes, have allowed the identification and quantification of a large numbers of mRNA transcripts. For example, REases (like *AclI* and *USER*) have been used to insert tags into sample DNA, which is amplified to form long, single-stranded DNA “nanoballs” that are templates in chip-based sequencing-by-ligation methodology.

**Nicking enzymes:** Some type IIS enzymes have two different catalytic sites. By inactivating one catalytic site at a time, these enzymes are converted strand-specific DNA nicking enzymes (cleaving only one DNA strand) and are useful molecular investigating tools for research on altering DNA, generation of reporter plasmids with modified structures, and creation of a DNA motor to transport DNA payload.

**In vivo gene editing:** to correct mutations that causes genetic disease (or any other ethical genetic mutation): site-specific-cleavages were realized using Zinc Finger Nucleases (ZFNs) and Transcription-Activator-like Effector Nucleases (TALENs). These techniques have since been replaced by the more precise CRISPR-Cas enzyme system, which can cut DNA within a long target sequence (~30 bp) and can be more easily reprogrammed for new targets. The CRISPR-Cas uses a small piece of RNA to recognize the intended DNA sequence, and guide the Cas 9 enzyme to cut the DNA at the targeted sequence.

## **6. How does this topic relate to other areas of cell biology, bioengineering, or medicine?**

After more than 40 years of revolution in biotechnology enabled partly by restriction enzymes, the cost of DNA sequencing and synthesis have decreased dramatically. More than ever we have access to a large amount of genetic information and powerful sophisticated genetic engineering tools. Potential ripple effects of these new sciences and engineering technologies are and will further impact medicine, food, agriculture, energy production, manufacturing, national security, environment and transform completely our society.

By better understanding restriction enzyme activity, their action on targeted DNA (cleavage mechanisms), they could allow to boost research in [5]:

- **Production on demand of small molecules:** the process of generating any bespoke molecule.
- **In healthcare:**
  - Using patient's own cells to differentiate into specific cell types, to self-assemble into new tissues, or repair damaged organs
  - Production of theranostic cell lines which sense a disease state and produce therapeutic responses.
  - Production of sensor cells to better understand gene expression and regulation.
  - Optimization of antibodies or vaccines to deliver large genetic payloads to targeted tissues with fewer side effects and less resistance from receiving host.
  - Genetically engineer viruses to correct defective genes in inherited diseases (like treating sickle cells).
  - Creation of disease fingerprinting using biomarkers

[1] W. A. M. Loenen, D. T. F. Dryden, E. A. Raleigh, G. G. Wilson, and N. E. Murray, "Highlights of the DNA cutters: a short history of the restriction enzymes," *Nucleic Acids Research*, vol. 42, no. 1, pp. 3–19, Jan. 2014, doi: 10.1093/nar/gkt990.

[2] A. Pingoud and A. Jeltsch, "Recognition and Cleavage of DNA by Type-II Restriction Endonucleases," *Eur J Biochem*, vol. 246, no. 1, pp. 1–22, May 1997, doi: 10.1111/j.1432-1033.1997.t01-6-00001.x.

- [3] B. Das, K. Banerjee, and G. Gangopadhyay, “On the Role of Magnesium Ions in the DNA-Scissoring Activity of the Restriction Endonuclease ApaI: Stochastic Kinetics from a Single Molecule to Mesoscopic Paradigm,” *J. Phys. Chem. B*, vol. 125, no. 16, pp. 4099–4107, Apr. 2021, doi: 10.1021/acs.jpcb.0c10643.
- [4] “How restriction enzymes became the workhorses of molecular biology.” <https://www.pnas.org/doi/10.1073/pnas.0500923102> (accessed Apr. 07, 2022).
- [5] M. El Karoui, M. Hoyos-Flight, and L. Fletcher, “Future Trends in Synthetic Biology—A Report,” *Front. Bioeng. Biotechnol.*, vol. 7, p. 175, Aug. 2019, doi: 10.3389/fbioe.2019.00175.