WHAT'S NEW

New Methods for Probing Nucleic Acids

H. Peter Spielmann, Jason D. Kahn, and John E. Hearst*

Introduction

Modern molecular biology depends on the ability to manipulate DNA. Cloning and other genetic engineering techniques require the specific cleavage of DNA at restriction sites (see below). One method of sequencing nucleic acids relies on base-specific chemical reactions, and several structure-specific ribonucleases are used to probe the secondary and tertiary structure of RNA. There is a great deal of interest in new techniques for studying the structure and function of DNA and RNA. Also, the advent of pulsed field gradient electrophoresis[1] has made it possible to analyze chromosome-size DNA, and this suggests the extension of probing techniques to this realm. (The size of a typical plasmid is ~ 5 kb, λ phage is 40 kb, the *E. coli* chromosome is 5 Mb, and the human X-chromosome is 50 Mb; common techniques are useful up to a size of perhaps 50 kb.) We will briefly review some advances in enzymatic and chemical mapping and discuss possibilities for studying structure in nucleic acids and for restricting large DNA's.

A New Restriction Technique

Restriction enzymes are some of the most important tools available for studying nucleic acids. They recognize specific sequences (often 4–6 base-pair palindromes) in DNA and cut both strands of the DNA at sites either within or near the recognition sequence. Restriction enzymes are essential for cloning and sequencing of DNA, but restriction enzyme methods are limited. The investigator is able to cut DNA only at sequences recognized by an available enzyme, and they are not active with single-stranded DNA or with RNA. Finally, there are a great

many 6 or even 8 basepair recognition sites in an entire genome (such as the human genome), and therefore for coarse mapping purposes, restriction digestion provides more fragments than desired. A method proposed[2] and tested[3] by Waclaw Szybalski addresses some of these limitations.

The Szybalski method essentially consists of creating artificial restriction sites at any position on single-stranded (ss) DNA. It uses a synthetic oligonucleotide containing the recognition sequence for a Class IIS restriction enzyme (one which cuts several base pairs away from its recognition site) and a linker complementary to the sequence one wants to cut (see Fig. 1). This molecule is annealed to the target ssDNA in the presence of the appropriate restriction enzyme. A double stranded substrate for the enzyme is thereby created, with the cut site on the target DNA.

This method can potentially be used to create a restriction system with any desired specificity, using a different synthetic oligonucleotide each time. The recognition element for the system as a whole can be any sequence; it is controlled by the variable part of the

Fig. 1. *An outline of the Szybalski restriction method.* (*a*) *The Fok I recognition site (thick black line) and the staggered cleavage points (arrows). The enzyme is schematically outlined (stippled). N represents any deoxynucleotide.* (*b*) *A 34-mer adapter oligonucleotide designed to react with Fok I and cleave M13 ss DNA.* (*c*) *Base pairing between the adapter and the target M13 ss DNA (italics). Reproduced from* (2), *by permission of Elsevier Science Publishers.*

above synthetic oligonucleotide. As Szybalski mentions, this enables specific restriction of natural ss DNA – there is no need to perturb the system by inserting linker sequences. Such a method could be especially useful for 'restricting' at perhaps 10 or 12 bp recognition elements. A restriction enzyme active on 12 bp sites in ds DNA would find application in cloning and in analysis of very large DNA's, and the Szybalski method for ss DNA could point the way in that direction.

The most serious limitation to this procedure is that it cannot be used easily on double-stranded DNA. Since the adapter-enzyme system cuts single-stranded DNA, to cut at a desired site on ds DNA requires denaturation or melting in some way to yield ss DNA. However, the adapter-target complex must be stable under the same conditions. There are several possible solutions to this problem. By analogy to the hybridization of a sequencing primer in 'supercoil sequencing',[4] Szybalski suggests annealing the adapter to reversibly denatured ds DNA. This would yield only a nick in one strand, so the procedure would have to be done once on each strand, or perhaps a nuclease such as S1 could be used to cut the other strand across from the nick (this would require quantitative renaturation to avoid destroying the rest of the DNA). Many of these proposals rely on very difficult or untested methods. In any case, the natural recognition sites on the ds DNA would have to be protected in some way, probably by specific methylation. It might also be possible to covalently link the restriction enzyme and the adapter oligonucleotide in such a way that the enzyme would be active only on the artificial substrate.

In general, methods based on restriction enzymes have some fundamental limitations. They provide information on the sequence of a nucleic acid but usually not its structure and conformation. Recently, chemical methods have been developed which may be more useful in this respect.

Chemical Mapping Methods

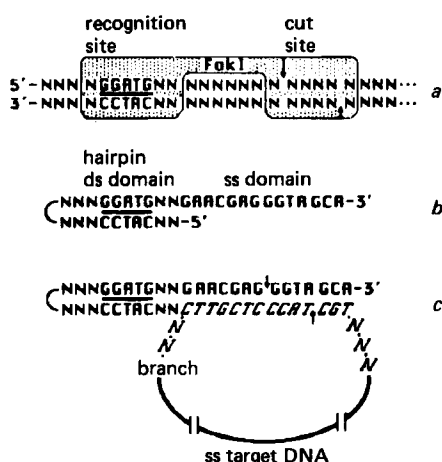Chemical methods of mapping DNA (or RNA) have also become quite

sophisticated. A familiar example is Maxam-Gilbert sequencing,[5] in which base-specific chemical reactions are used to modify DNA (or RNA) at specific bases and the sugar-phosphodiester backbone is then cut at the modified sites. There are many small molecules which have been used to probe DNA and RNA structure, to differentiate between single-stranded and double-stranded regions, and to study protein-nucleic acid interactions. Examples include the work of Mirzabekov and coworkers on DMS footprinting,[6] the use of psoralens to crosslink double-stranded regions,[7] and bromoacetaldehyde (BAA) as a probe for unpaired DNA bases.[8] Recently Johnston and Rich[9] have developed a method for identifying different structures in ds DNA using small molecules which modify Z-DNA. They have found, for example, that osmium tetroxide and hydroxylamine react preferentially with DNA bases near junctions between B and Z form DNA, and that diethyl pyrocarbonate reacts with purines inside Z-DNA regions. Barton has developed a family of cobalt and ruthenium complexes which bind Z-DNA. Upon UV irradiation, the cobalt complex Λ-[Co(DiP)$_3$]$^{3+}$ nicks Z-form DNA in supercoiled plasmids.[10] These chemical methods give information on structure as well as sequence, but in contrast to restriction enzymes they do not yield useful pieces of DNA after cleavage.

Another approach to mapping DNA has been developed by Peter Dervan and coworkers, who have attempted to make chemical probes which bind to specific nucleic acid sequences.[11] Several oligopeptide antibiotics bind to specific base pairs in DNA. Netropsin and distamycin are oligo N-methylpyrrolecarboxamides that bind the minor groove of B-DNA with a strong preference for A-T rich regions.[12] Tethering an Fe(II)·ETDA moiety to these molecules enables them to cut the DNA to which they bind. The addition of oxygen and a reducing agent such as dithiothreitol to the solution initiates the reaction. The reactive species is a diffusible hydroxyl radical (HO˙) which reacts with nearby sugars, so the procedure yields a family of DNA fragments. This lack of specificity gives the EDTA chelate its nickname, the 'wrecking ball'. Some of these molecules are shown in Fig. 2.

Notwithstanding the non-selective cleavage chemistry involved, molecules of this type can be sequence-specific and structure-specific probes. The above oligopeptides are specific for B-DNA and they bind most effectively to A-T rich regions, but their sequence specificity can be modified to make them prefer G-C regions. For example, Dervan has synthesized a bis (distamycin) phenoxazone (Fig. 2), as a groove binder-intercalator-groove binder, which binds the 10-bp sequence 5'-TATAGGTTAA-3'. The pheno-xazone ring preferentially intercalates between G-C base pairs and the distamycin segments bind in the A-T grooves. In this manner, specificity for 'words' composed of A-T and G-C 'letters' can be constructed.

Intermediate between Szybalski's approach and the small groove-binding wrecking balls is a class of DNA oligonucleotides with attached Fe(II)· EDTA moieties. These molecules hybridize to single-stranded DNA (and presumably RNA), and upon addition of DTT and molecular oxygen they cleave the DNA as above.[13] Again this leads to a family of fragments, but if such a system were used for coarse mapping of large DNA, this would not be a serious limitation. However, the oligonucleotide wrecking ball shares with the Szybalski method the problem of requiring ss DNA.

Conclusions

A long-term goal for mapping techniques is the sequence analysis of very large DNA's, for example human chromosomes. Also, because it is apparent that nucleic acids can be very polymorphic in structure, detection of special structures in DNA is of interest as well. The ideas and techniques above may be useful in both of these areas. In its present form, the Szybalski method of specifically cleaving ss DNA is not likely to be useful for large ds DNA, but the concept of an artificial restriction
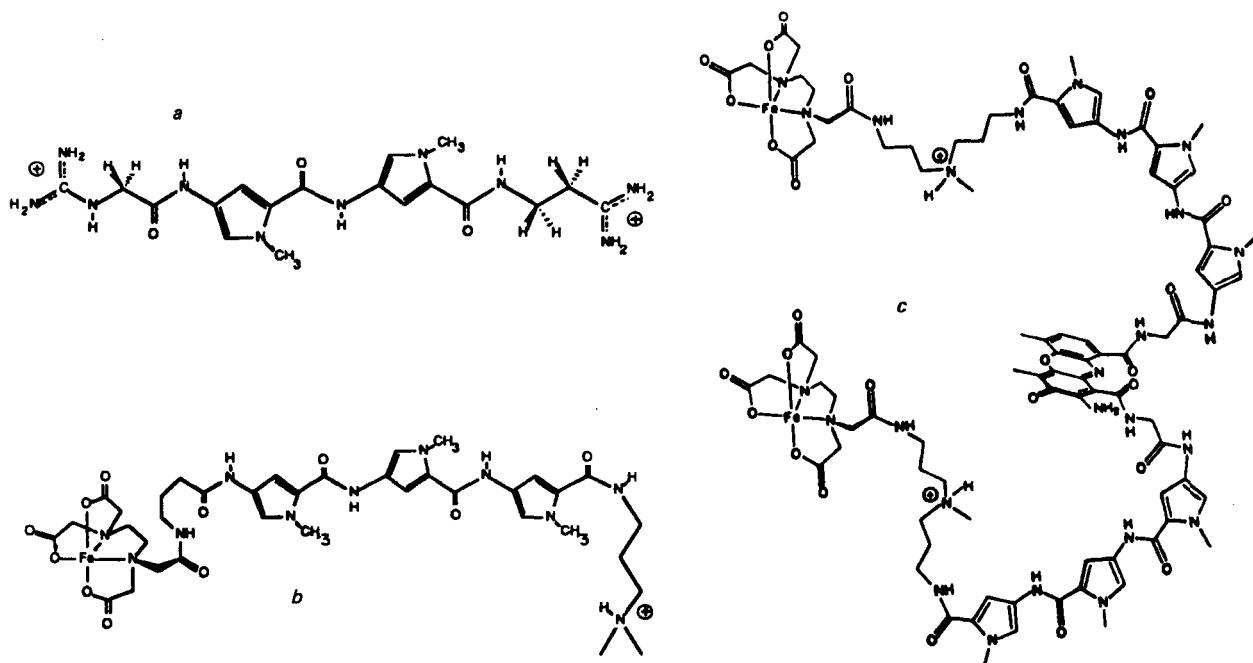

Fig. 2. *Several specific chemical probes; see* (12). (*a*) *The antibiotic netropsin.* (*b*) *Distamycin-EDTA . Fe(II).* (*c*) *Bis* [*Fe(II)*]. *EDTA-distamycin*] *phenoxazone; the wrecking ball is tethered to opposite ends of the distamycin moiety in* (*b*) *and* (*c*). (*c*) *reproduced from* (11), *by permission of the American Association for the Advancement of Science. Copyright* 1986 *by the AAAS.*

WHAT'S NEW

enzyme is certainly valuable. To define a unique sequence in human genomic DNA requires about 20 bp; an artificial restriction system or an oligomer with a 'wrecking ball' could be useful for cutting at such a large recognition sequence. This requires extemely specific hybridization with very low background; we favor crosslinking agents such as psoralen to make this possible.[14] It should be noted, however, that there are many difficulties in mapping large DNA which we have not addressed here.

Chemical methods are still being tested and improved. However, they have already provided valuable probes into nucleic acid structure, and chemical methods may be best for this purpose; enzymes cannot give much information on structure beyond differentiating ss and ds DNA. So far, the binding of chemical probes is not as sequence-specific as that of oligonucleotides or restriction enzymes, but as more small-molecule probes are synthesized we expect that the specificity will be improved. As the chemistry of DNA becomes better understood, it may be possible to develop an entirely chemical restriction system, thereby obviating the need for enzymes under some circumstances. A combination of chemical and enzymic methods could be the best of both worlds.

REFERENCES

1 SCHWARTZ, D. C. & CANTOR, C. R. (1984). Separation of yeast chromosome-sized DNA's by pulsed field gradient gel electrophoresis. *Cell* 37, 67–75.

2 SZYBALSKI, W. (1985). Universal restriction endonucleases: Designing novel cleavage specificities by combining adapter oligodeoxynucleotide and enzyme moieties. *Gene* 40, 169–173.

3 SZYBALSKI, W. & PODHAJSKA, A. J. (1985). Conversion of the *FokI* endonuclease to a universal restriction enzyme: Cleavage of phage M13mp7 DNA at predetermined sites. *Gene* 40, 176–182.

4 CHEN, E. Y. & SEEBURG, P. H. (1985). Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. *DNA* 4, 165–170.

5 MAXAM, A. M. & GILBERT, W. (1977). A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* 74, 560–564.

6 MELNIKOVA, A. F., BEABESASHVILLI, R. & MIRZABEKOV, A. D. (1978). A study of unwinding of DNA and shielding of the DNA grooves by RNA polymerase by using methylation with dimethyl sulfate. *Eur. J. Biochem.* 84, 301–309.

7 CIMINO, G. D., GAMPER, H. B., ISAACS, S. T. & HEARST, J. E. (1985). Psoralens as photoactive probes of nucleic acid structure and function: Organic chemistry, photochemistry, and biochemistry. *Annu. Rev. Biochem.* 54, 1151–1193.

8 KOHWI-SHIGEMATSU, T. & KOHWI, Y. (1985). Poly (dG)-poly(dC) sequences, under torsional stress, induce an altered DNA conformation upon neighboring DNA sequences. *Cell* 43, 199–206.

9 JOHNSTON, B. H. & RICH, A. (1985). Chemical probes of DNA conformation: Detection of Z-DNA at nucleotide resolution. *Cell* 42, 713–724.

10 BARTON, J. K. & RAPHAEL, A. L. (1985). Site-specific cleavage of left-handed DNA in pBR322 by *A*-tris (diphenylphenanthroline) cobalt(III). *Proc. Natl. Acad. Sci. USA* 82, 6460–6464.

11 DERVAN, P. B. (1986). Design of sequence-specific DNA-binding molcules. *Science* 232, 464–471.

12 KOPKA, M. L., YOON, C., GOODSELL, D., PJURA, P. & DICKERSON, R. E. (1985). The molecular origin of DNA-drug specificity in netrospin and distamycin. *Proc. Natl. Acad. Sci. USA* 82, 1376–1380.

13 DREYER, G. B. & DERVAN, P. B. (1985). Sequence-specific cleavage of single-stranded DNA: Oligodeoxynucleotide-EDTA. Fe(II). *Proc. Natl. Acad. Sci. USA* 82, 968–972.

14 CIMINO, G. D., GAMPER, H. B., ISAACS, S. T. & HEARST, J. E. Manuscript in preparation.

H. PETER SPIELMANN, JASON D. KAHN, and JOHN E. HEARST (*to whom all correspondence should be addressed*) *are at the Department of Chemistry, University of California, Berkeley, CA* 94720, *USA*.

The Airport Professor

Dublin – Where's Dublin?

That might be the response of a KLM agent on being asked for an airline ticket to Dublin. The (U.K.) *Sunday Times* of 27 July 1986 reports that: 'Air Lingus, the Irish State airline, has for six years had a secret agreement with KLM, the Dutch flag carrier, under which it agreed to pay KLM almost £500,000 *not* to fly to Dublin...one of an orchestral array of anti-competitive agreements between Europe's major airlines' which the European Commission has detailed in a letter to 10 major carriers, giving them two months to dismantle.