thing else), the shape becomes immutable, at least until the membrane breaks up, unperturbed by metabolic depletion or a crenating drug. Why, what else but that the shape is after all controlled by the cytoskeletal complex. And here too some reappraisal has been called for with the discovery in the red cell of more of the protein components characteristic of contractile systems (Fowler, V.M. & Bennet, V.J. *biol. Chem.* **259**, 5978; 1984).

So we have moved within the year from poverty to riches, to at least three entirely self-consistent explanations of how metabolic shape control operates. No matter that they are almost certainly mutually incompatible, perhaps even all wrong: they will undoubtedly inject good cheer and even perhaps a cautious optimism into toilers in the field. □

Walter Gratzer is in the MRC Cell Biophysics Unit, 26-29 Drury Lane, London WC2B 5RL.

Sequencing techniques

A close look at the genome

from P.F.R. Little

A recent paper¹ in *Proceedings of the National Academy of Sciences U.S.A.* from George Church and Walter Gilbert of Harvard University and Biogen describes a remarkably powerful new method, named genomic sequencing, which provides us with the most accurate view yet of a DNA sequence in a chromosome. It can be used to study methylation patterns directly in uncloned DNA and to analyse DNA sequences associated with protein in chromatin.

The method combines the chemical DNA-sequencing procedures of Maxam and Gilbert² with detection of DNA sequences by Southern blotting³. Figure 1 outlines the procedure by which a DNA molecule can be directly sequenced from unfractionated mouse DNA. It illustrates the sequencing of cytosine (C) residues but similar rationale and procedures are used for sequencing adenine (A), guanine (G) and thymine (T) bases.

The first step is complete cleavage of total cell DNA by a restriction enzyme; Fig. 1a shows a double-stranded molecule generated by digestion with *EcoRI*. This is followed by partial cleavage at C residues by hydrazine and fractionation of single-strand DNA fragments on 6 per cent polyacrylamide gels. The DNA is then transferred electrophoretically to nylon membranes and fixed by UV irradiation, and specific DNA sequences are detected by hybridization to a ³²P-labelled nucleic acid probe, as in conventional Southern blots. If we consider only the lower strand of the DNA molecule in Fig. 1a, then partial C cleavage will generate the single-strand fragments in b and c. The fragments in Fig. 1b can be detected by hybridization with probe 1 and will show up on an autoradiograph as a set of DNA fragments whose sizes correspond to the distance between the left-hand *EcoRI* site and the position of each C residue. The fragments in Fig. 1c are not homologous to probe 1 and so are not detected. The concept is no different from using end-labelled DNA in more conventional sequencing experiments. The novel feature, of course, is that the hybridization step enables one set of fragments to be picked out from the 10⁶

other *EcoRI* fragments in the mouse genome.

In the figure, the DNA detected by probe 1 contains no C residues — a rather unlikely event. What is the effect of cleavage at C residues within the DNA detected by the probe? The result would be fragments similar to those in Fig. 1b, but terminating at both ends in C residues, and these would generate on the autoradiogram a second set of fragments, starting at the internal cleavage site. This is not a problem if the frequency (P) of C cleavage is small. Since they all terminate on the left with an *EcoRI* site, the fragments in Fig. 1b will be generated at a frequency of P; by contrast, fragments terminating at a C at both ends will be generated at a frequency of P². Thus, internal C cleavage will generate only small amounts of fragments and, consequently, a faint band pattern. Church and Gilbert are able to show that cleavage

DNA fragments detected by genomic sequencing. Probe 1 is the DNA from the upper strand and is homologous to the lower strand. a, Double-stranded DNA molecule resulting from digestion of total cell DNA with *EcoRI*. Partial cleavage at C residues by digestion with hydrazine generates two sets of fragments, one of which will hybridize with probe 1 (b) and another which will not hybridize (c).

at an average of 1 per 500 nucleotides is optimal. The nature of probe 1 is also important. If it is too big, with too many internal Cs, then the cumulative effect of the alternative sequence sets will tend to obscure the primary set. Similarly, if the probe is too short it will hybridize non-specifically and generate other fragment sets. Church and Gilbert indicate that probes of 100–200 nucleotides 'work well'.

A, G and T bases are sequenced in a similar way and the complete sequence can be read directly from the hybridization autoradiograph. An advantage of this method is that filters can be rehybridized, for example with probe 2, 3 or 4. Probes 2 and 4 can be used to sequence the upper strand and probes 1 and 3 the lower. The procedure is remarkably sensitive: as little as 3 fg can be detected using highly radioactive RNA probes made by SP-6 RNA polymerase (reviewed in ref. 4) or single-strand DNA probes.

Church and Gilbert have used their new method to study differential methylation of C residues round the mouse immunoglobulin C_μ gene in several expressing and non-expressing cell lines. This is possible because 5-methylcytosine is relatively insensitive to hydrazine cleavage, so the appropriate fragments are absent or under-represented in genomic sequences. The method allows the analysis of all 5-methylcytosine bases, not just those that fall within the recognition sites of particular restriction enzymes, and should result in a much clearer understanding of the inverse correlation between methylation and gene activity.

More importantly, perhaps, genomic sequencing can be used to detect specific DNA sequences bound by protein. The protein protects the DNA against chemical or enzymatic attack ('footprinting') and a characteristic 'hole' in the sequence results (see ref. 5 for recent advances). This method was originally developed using cloned DNA sequences to study lac and λ repressor-operator and RNA polymerase-promoter interactions. Genomic sequencing can be used in a similar fashion and Church, in collaboration with Anne Ephrussi in Tonegawa's laboratory at MIT, has recently detected protein(s) interacting with the immunoglobulin enhancer at two distinct sites (personal communication).

For the future we can anticipate the use of non-radioactive probes and Church is exploring methods for rapid DNA sequencing. The results of more general application of genomic sequencing are awaited with much interest. □

1. Church, G. & Gilbert, W. *Proc. natn. Acad. Sci.* **81**, 1991 (1984).

2. Maxam, A.M. & Gilbert, W. *Proc. natn. Acad. Sci.* **74**, 560 (1977).

3. Southern, E.M. *J. molec. Biol.* **98**, 503 (1975).

4. Little, P.F.R. *Nature* **309**, 191 (1984).

5. Becker, M.M. & Wang, J.C. *Nature* **309**, 682 (1984).

P.F.R. Little is at the Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London SW3 6JB.