Molecular biology and respiratory disease · 2   *Series editor: R A Stockley*

Applications to the study and treatment of respiratory disease: methods in molecular biology

C A Owen, R A Stockley

In the first article in this series we covered basic gene structure and function. The aim of this second introductory article is to provide an outline of the techniques commonly used in molecular biology and to provide examples of the application of these techniques to the investigation and diagnosis of respiratory disease.

Restriction endonucleases

Restriction endonucleases are bacterial enzymes that cleave specific base sequences of double stranded DNA. They are bacterial in origin and digest foreign DNA to prevent its incorporation into the genome of the bacterium. More than 350 restriction enzymes have been isolated, and between them they recognise some 85 distinct DNA sequences. They are named by three letter abbreviations derived from the bacterial species from which they have been isolated; for example, Eco R I is one of the enzymes derived from the R strain of *Escherichia coli*.

There are two classes of restriction enzymes. *Class 1 enzymes* recognise a specific nucleotide pair sequence and then cleave the DNA at a non-specific site a fixed distance away from this recognition site. *Class 2 enzymes* cleave the DNA only at the recognition site and always in a specific manner. Thus class 2 enzymes are more important in recombinant DNA technology because the cleavage is predictable. Class 2 enzymes are referred to as "sequence specific"—that is, they recognise only specific sequences, one for each enzyme. The sequences are usually three to six base pairs in length and possess a twofold rotational symmetry (palindrome). The number of cuts that each enzyme makes in a particular DNA molecule depends on the number of times the particular recognition sequence is present in the DNA. Some enzymes cut both strands of DNA between adjacent nucleotide pairs to form so called *blunt* ends but most make staggered cuts in the sequence to produce *sticky* ends (see table). This latter effect is useful in joining one fragment of DNA to another fragment derived from a different source when both have been cleaved by the same enzyme (for example, in the construction of cloning vehicles). The overlapping complementary sequences at the ends of the fragments will join spontaneously by hydrogen bonding *(anneal)*. Indeed, when DNA fragments are produced

with "blunt" ends it is often necessary to lengthen them in an asymmetrical manner in order to produce artificial "sticky" ends to facilitate annealing. Because of their specificity, restriction enzymes are also important in the detection of restriction fragment polymorphisms (see below).

Nucleic acid hybridisation and DNA probes

Double stranded DNA can be dissociated (*denatured*) by heating or by treatment with an alkaline solution. If the temperature is subsequently lowered or the solution neutralised the two strands will reassociate (*anneal*). This process is highly specific and will occur only between complementary strands. If two DNA sequences from different sources are denatured and allowed to reassociate, base pairing will take place only between any complementary sequences to form a duplex (*hybridisation*).

Probes are fragments of single stranded DNA or RNA that have been replicated in bacteria (*cloned*) to yield millions of identical copies. These fragments are labelled and used to identify (probe) complementary sequences in a mixture of nucleic acids. If a complementary sequence is present the probe will hybridise with the target sequence and the resulting duplex will be identified by the label on the probe. Probing may be carried out with nucleic acid that has been chemically extracted from tissues, digested into fragments by restriction

Sequence specificities and sites of cleavage of some commonly used restriction endonucleases

Enzyme	Recognition sequence with position of cut	Product
Bam H I	5'-GGATCC-3' 3'-CCTAGG-5'	-G GATCC- -CCTAG⁺ G- Sticky ends
Eco R I	5'-GAATC-3' 3'-CTTAG-5'	-G + AATC- -CTTA G- Sticky ends
Hae III	5'-GGCC-3' 3'-CCGG-5'	5'-GG + CC-3' 3'-CC GG-5' Blunt ends
Sma I	5'-CCCGGG-3' 3'-GGGCCC-5'	5'-CCC + GGG-3' 3'-GGG CCC-5' Blunt ends

Lung
Immunobiochemical
Research Laboratory,
General Hospital,
Birmingham
C A Owen
R A Stockley

Address for reprint requests:
Dr R A Stockley,
Clinical Teaching Block,
General Hospital,
Birmingham B4 6NH.

enzymes, separated by gel electrophoresis, and blotted on to nitrocellulose. Alternatively, the probe can be used directly on histological sections or whole chromosomes (*in situ hybridisation*).

There are two basic requirements for hybridisation with a probe—namely, that the target sequences are accessible and that both are in the form of a single strand.

DNA probes can be constructed in three ways:

1 chemical synthesis alone to obtain a predetermined nucleotide sequence (for example, synthetic *oligonucleotide probes*—see below);

2 cleavage of genomic DNA and subsequent identification and amplification of the fragment of interest in an appropriate vector;

3 treatment of messenger RNA (mRNA) with reverse transcriptase to produce *complementary DNA* (cDNA). This is the most commonly used method and is usually applied to mRNA obtained from tissues producing the protein of interest. As a result fewer DNA species have to be screened than when the whole genome is used. The cDNA probes consist only of the coding sequences (exons) and no introns are included.

DNA probes are used to determine the chromosomal locations of genes, analyse gene structure, diagnose single gene defects by the detection of restriction fragment polymorphisms, and determine the tissue specificity and activity of gene transcription (see below).

Reverse transcription

Reverse transcriptase is an enzyme that was originally isolated from a group of viruses whose genome consists of RNA rather than DNA. The enzyme converts the viral genome into DNA (a process termed *reverse transcription*), which enables it to be inserted within the host cell DNA and hence lead to viral replication by the cell.

The enzyme is most commonly used to synthesise complementary DNA probes from mRNA. The mRNA of interest is purified from a tissue in which it is being actively translated (for example, reticulocytes are used to obtain haemoglobin mRNA). The microsomal fraction of the tissue is separated and the total cellular RNA is purified and passed down a column containing either poly-T (thymine) or poly-U (uracil) residues linked to a solid phase support. These residues bind the poly-A (adenine) containing (coding) mRNA while other types of RNA pass through. The mRNA is subsequently eluted from the column and reverse transcriptase is used to synthesise a complementary DNA (cDNA) on the single stranded mRNA template. An oligonucleotide complementary to the poly-A tail (oligo dT) is used as a "primer" DNA molecule. It hybridises with the poly-A tail and is extended by reverse transcriptase to produce a DNA copy of the mRNA. The RNA is then destroyed by treatment with alkali and the single stranded cDNA is again copied to give a double stranded DNA molecule with the help of the

enzyme *DNA polymerase*. A final enzyme, called *S1 nuclease*, is then added to remove the remaining short sequence of single stranded DNA, producing a matched double stranded cDNA molecule for subsequent cloning.

Cloning

Cloning refers to the insertion of fragments of DNA into bacteria to produce multiple identical copies (clones). DNA is not taken up efficiently in its native state by bacteria. Thus the fragments are inserted into specialised forms of DNA (*vectors*) that can enter and replicate within the bacteria. The choice of vector usually depends on the size of the DNA fragment to be cloned: *plasmids* are 0–10 kilobases, bacterial viruses or *bacteriophages* 5–20 kilobases, and *cosmids* are hybrid vectors composed of plasmid and bacteriophage DNA sequences (40–50 kilobases).

Plasmids are used preferentially for the cloning of gene probes when possible. They are self replicating units of circular double stranded DNA, which have a single cleavage site for one of the bacterial restriction enzymes and usually carry genes encoding antibiotic resistance. For example, the synthetic plasmid pBR322 carries resistance genes for both ampicillin and tetracycline and within the latter gene is the single restriction site for the enzyme Bam H1 (fig 1). Incubation of the plasmid with this enzyme will result in cleavage of the plasmid at this site alone, leaving "sticky" ends. If the DNA to be cloned (insert) is also digested with Bam H1 similar "sticky end" sequences are produced. When the two forms of DNA are mixed some of the sticky ends of the plasmid will associate, by hydrogen bonding, with the DNA to be inserted. The two fragments can be joined covalently by the enzyme DNA ligase to form a new "recombinant" DNA plasmid molecule. The recombinants are then introduced into bacterial cells for propagation by a process called *transformation*. The few bacteria which have been transformed by the recombinant molecules are identified by the presence and type of antibiotic resistance that the plasmid has introduced (fig 1). The transformed bacteria are then grown in bulk culture and lysed and the plasmid DNA is harvested by centrifugation. The restriction enzyme used to create the recombinant molecule is then used to extract the insert, which is separated from the plasmid DNA by gel electrophoresis. In this way a large quantity of identical copies of the original DNA fragment may be obtained.

Probing and blotting

SOUTHERN BLOTTING

Southern blotting, devised by E M Southern, is a procedure in which a radiolabelled DNA probe is used to identify a specific sequence of DNA in a complex mixture of DNA fragments that have been transferred on to a nitrocellulose membrane.[1]

In brief, genomic DNA is extracted from blood leucocytes or other tissues and digested by various restriction enzymes to produce
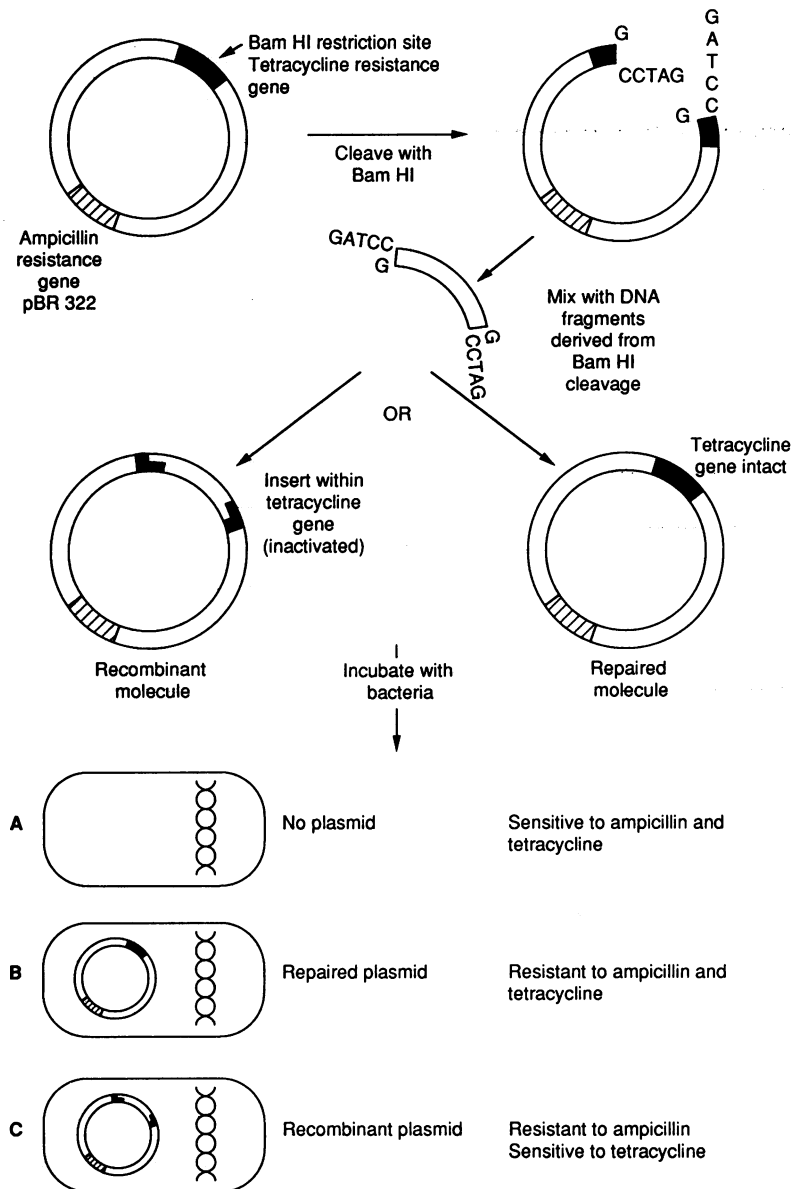
Figure 1 Cloning of a fragment of DNA by construction of a recombinant plasmid and transformation by competent bacteria. Selection of recombinant plasmids is made by exploiting the antibiotic resistance genes.

of appropriate restriction sites. The fragments will be identified on the Southern blot and their sizes are determined by reference to molecular weight markers run simultaneously on the electrophoresis gel. If this process is repeated with several restriction enzymes, alone or in various combinations, a restriction map of the gene can be constructed and the molecular basis for certain genetic diseases may become apparent.

Deletion of a gene will result in the abolition of any hybridisation signal. Partial gene deletion or single base mutations may result in the loss or creation of restriction sites for a restriction enzyme and therefore lead to a change in the size pattern of fragments generated and identified on the Southern blot. For example, in sickle cell anaemia a single base substitution (GAG → GTG) results in the loss of a recognition site for the enzyme Mst 11.[2] The abnormal gene can therefore be detected by digesting DNA with Mst 11, which results in the production of larger fragments than normal. These can be detected by Southern blot hybridisation with a β globin DNA probe.

Southern blotting has also been performed with oligonucleotide probes, which can detect even single point mutations in intact DNA sequences. Oligonucleotides complementary to both the normal and the mutant forms of the gene are synthesised. These oligonucleotides are radiolabelled and used to probe Southern blots of DNA prepared from samples from normal subjects or those who are either heterozygous or homozygous for the DNA defect. Under appropriate conditions the probes will hybridise with their complementary sequences and distinguish the normal and mutant alleles. This technique has been used to detect the point mutation of the Z variant of the α_1 antitrypsin gene.[3]

multiple DNA fragments of varying length, which are separated according to size by gel electrophoresis (fig 2). The gel is soaked in alkali to denature the double stranded DNA fragments and it is then transferred to an overlying nitrocellulose membrane. The membrane thus acquires an imprint of the separated single stranded DNA fragments. The DNA is fixed on to the membrane by baking and exposed to a solution containing the radiolabelled single stranded probe, which binds through hybridisation to any complementary DNA sequences. The membrane is washed to remove unbound probe and the signal is developed by autoradiography to show the positions of the fragments bearing sequences complementary to that of the probe that has been used.

Southern blotting can be used to analyse gene structure by the construction of *restriction maps*. When genomic DNA is digested with a restriction enzyme it will cleave the gene of interest into several fragments, whose number and nature depend on the presence and position

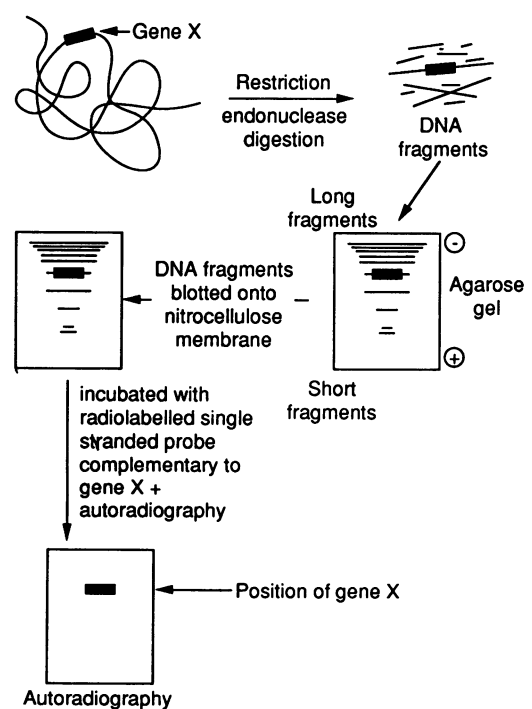*Figure 2 The Southern blot technique.*

RESTRICTION FRAGMENT LENGTH POLYMORPHISMS

The human genome is highly polymorphic (variable between individuals). On average, 1 in 250 of the bases is polymorphic and 1 in 6 of these polymorphisms results in the loss or gain of a recognition site for one of the restriction enzymes. Most of the polymorphisms are situated in the 95% of the genome that is non-coding and therefore do not affect the phenotype of the individual. The polymorphisms are revealed as changes in the length of DNA fragments produced by the restriction enzymes on Southern blotting when an appropriate DNA probe is used. To be clinically useful, however, the polymorphism must be rare in the healthy population and common in the disease state. Thus if the restriction fragment length polymorphism is closely and stably linked to the abnormal gene (linkage disequilibrium) it can be used as a genetic marker for the disease even if the true gene defect is not known. For example, in the United States among the black population the sickle cell mutant is associated in 60% of cases with a polymorphism close to the gene, which eliminates a site for the enzyme Hpa 1,[4] though the use of this marker is inferior to analysis with Mst 11 (see above).

Unfortunately the use of restriction fragment length polymorphism is limited because few diseases are the result of single mutations, and even these may not alter restriction sites.

LOD scores

When diseases are not known to be linked to specific genes the search for genetic linkage may be tedious and time consuming. The process depends on the identification of a gene polymorphism that is inherited together with the disease. This means the application of restriction fragment length polymorphism to large families where the disease can be identified and where several generations are available for study. The process is tedious as the gene is unknown and hence all genes may have to be studied. Furthermore, because the specific defect is unknown a linked polymorphism has to be found and this could occur with any one of the 350 or so restriction endonucleases. Thus it is exceptionally fortunate when a linkage is found relatively early during familial studies, as recently seen with IgE responsiveness and chromosome 11.[5] In this instance a linkage was found when only the 17th gene probe and one restriction enzyme (Taq 1) were used.

The process of genetic linkage depends on the disease and an identified polymorphism cosegregating: individuals with disease show the polymorphism whereas those without do not. But because genetic material may cross over during meiosis the polymorphism may become separated from the abnormal gene, and this is more likely the further they are apart. Separation of the polymorphism from the abnormal gene and hence the disease is due to recombination. The proportion of siblings in whom this occurs determines the *recombination fraction*. For instance, if in 10 siblings five have disease X and polymorphism Y, but only one sibling without the disease has polymorphism Y, the recombination fraction is probably 1 out of 10 or 0·1. This value is used to describe how far apart the gene for disease X and the polymorphism Y occur in the genome. In this example the two loci are 10 centimorgans (cM) or map units apart.

The next step consists of mathematical calculations of the likely value of the recombination fraction, which will depend on many factors, including the number of siblings studied and the disease penetration. This is simple but tedious and is usually calculated by means of a computer program.[6] The likelihood (or odds) that this recombination fraction is correct is conventionally expressed as its logarithm—the $\log_{10}$ of the odds or LOD score.

The LOD scores for several families are then added together and the sum is then plotted for each recombination fraction. The maximum likely estimate of the recombination fraction is the value that corresponds to the peak of the LOD scores. Thus in the example above a LOD score greater than 3·0 means that the odds are more than 1000:1 that a linkage occurs at this distance (10 centimorgans) between the gene and the polymorphic locus. In other words, the recombination fraction is very likely to be close to 0·1 and hence the restriction fragment length polymorphism to be within 10 centimorgans of the abnormal gene.

Once this association has been determined with confidence it is then possible to get progressively closer to the defective gene by the use of overlapping probes.

Northern and dot blotting of RNA

Northern blotting is similar to Southern blotting, except that the single stranded DNA probe is hybridised to complementary *RNA* molecules that have been blotted on to nitrocellulose after separation by size. Alternatively, cellular RNA may be directly "dotted" on to nitrocellulose and hybridised to the probe (dot blotting). These techniques may be used to assess the tissue specificity of gene expression. In addition, the Northern and dot blots can be semiquantified by densitometry scanning of the autoradiographs obtained, and this has facilitated the study of factors that alter the rates of gene transcription.

The major problem in assessing changes in the amount of RNA present in cells relates to the possible variability in purification yield. To take account of this possibility the results are corrected for the total amount of DNA or RNA that has been transferred to the nitrocellulose membrane. Alternatively, sample results are interpreted by comparison with a gene that shows constant expression (for example β actin).

Oligonucleotide directed mutagenesis

Oligonucleotide directed mutagenesis is a technique used to alter the gene being studied in a specific way to gain insight into protein function and even to produce designer proteins (see

below). Any nucleotide in a fragment of DNA can be specifically mutated to one of the other three nucleotides. The method requires that the DNA fragment to be mutated is available in pure single stranded form and is cloned into an appropriate vector. The next step is to synthesise chemically an oligonucleotide 7–20 base pairs in length that is complementary to the natural DNA but has a single base mismatch. The oligonucleotide will hybridise to the native DNA under carefully controlled conditions and can then be used as a template for priming DNA extension to form a heteroduplex molecule. The heteroduplex is then transformed into a bacterium and the bacterial DNA polymerase subsequently "repairs" the base mismatch to yield either the original or the mutant sequence. Further aliquots of the oligonucleotide sequence can be radiolabelled with phosphorus-32 and then under more stringent conditions used as a probe to screen for the mutant colonies. DNA sequencing is then used to confirm that the procedure has been successful.

Site directed mutagenesis has been used to produce new proteins. Any alteration in function will help to elucidate the primary structure-function relationships of the protein. For example, site directed mutagenesis of the gene that codes for the beta adrenergic receptor has shown that two cysteine residues on adjacent loops of the protein are important for maintenance of the tertiary structure (probably by the formation of a disulphide bond between the two loops) and hence the binding of ligand to the receptor.[7] Furthermore systematic substitution of neutral residues for acidic amino acids within the ligand binding domain of the beta receptor, by site directed mutagenesis of the gene, has resulted in identification of the specific amino acid residues which bind to the protonated amine group of the adrenergic ligands.[8]

Site directed mutagenesis has also been used to design new molecules for a defined purpose. For example, site directed mutagenesis of the reactive centre of α_1 antitrypsin (Met 358→Val) has resulted in the production of a mutant α_1 antitrypsin species that is resistant to oxidative inactivation but retains the elastase inhibitory properties of the native protein. This modified inhibitor has been considered to have therapeutic potential in conditions where a reduction in α_1 antitrypsin function due to oxidation is thought to be critical in the pathogenesis of chronic diseases.[9]

Expression vectors and cell free translation systems

DNA from individuals with a naturally occurring mutation or DNA manipulated in vitro by site directed mutagenesis may be introduced into a cell to determine the effects of the mutation on gene expression. The cells most commonly used are *Xenopus* (toad) oocytes and cells from the fertilised eggs or embryos of *Xenopus laevis*, the mouse, or the fly *Drosophila melanogaster*. *Xenopus* oocytes are most often used because they are readily available in large numbers. Each fully grown oocyte is a large cell (0·8–1·2 mm) with a correspondingly large nucleus, thus making microinjection of DNA into the nucleus technically easy. Single stranded DNA is converted into a duplex form in the oocyte nucleus and then assembled into chromatin. The oocyte nucleus contains eukaryotic RNA polymerases, which transcribe the exogenous DNA. When protein coding genes are injected into the nucleus and correct initiation of transcription occurs stable RNA is produced, which is transported into the cytoplasm. The RNA is translated into protein, which is then secreted by the oocyte into the incubation medium. In this way it is possible to study the effects of a mutation in a gene sequence on transcription, translation, and secretion of the protein product.

The oocyte may also be used as a cellular expression system for the study of mRNA. Injection of mRNA into the cytoplasm leads to protein synthesis, modification, and secretion. Alternatively, isolated mRNA species may be studied in cell free translation systems (for example, rabbit reticulocyte lysate or wheat germ lysate systems). Again translation of the mRNA occurs, leading to the formation of the protein product.

The mechanism for α_1 antitrypsin deficiency associated with the homozygous *ZZ* state has been partly elucidated by these techniques. Messenger RNA isolated from the liver of normal subjects and individuals homozygous *(ZZ)* for the α_1 antitrypsin defect has been used in cell free translation systems and in the *Xenopus* oocyte secretory system. The studies show that the low plasma concentrations of α_1 antitrypsin are the result of a block in the transport of the protein from the rough endoplasmic reticulum through the secretory pathway and not the result of impaired synthesis from mRNA.[10]

DNA sequencing

The development of techniques to determine the sequence of DNA molecules has enabled rapid advances in our knowledge of gene structure and function to be made and has allowed the defects underlying many genetic disorders to be determined. In addition, a knowledge of the basic sequence of DNA is a prerequisite for subsequent manipulations of DNA molecules by site directed mutagenesis. Once a sequence has been changed the sequence of the product has to be checked to confirm that the required mutation has been obtained.

The most widely used method for DNA sequencing was developed by Sanger *et al*[11] and uses an enzymic dideoxynucleotide technique (fig 3). The DNA molecule to be sequenced is initially cloned in a bacteriophage. A synthetic oligonucleotide with a sequence complementary to about 15 bases of the phage sequence 3' to the inserted DNA acts as a primer for DNA synthesis. The *Klenow fragment* of DNA polymerase is then used to synthesise a series of DNA molecules complementary to the DNA being sequenced. The reaction is carried out in the presence of dideoxynucleoside triphosphates.

Figure 3 Schematic representation of DNA sequencing (Sanger method). The sequence that is read off the sequencing gel is complementary to the target sequence.
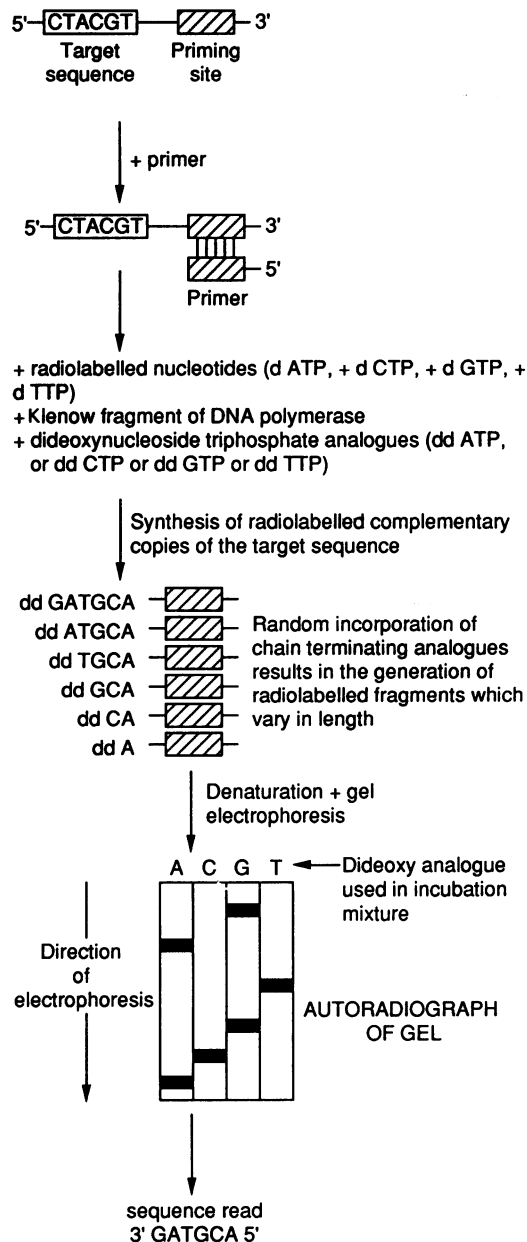
The Klenow fragment of DNA polymerase lacks the 5'→3' exonuclease activity of the intact enzyme and thus can also use dideoxynucleoside triphosphate molecules as substrates. These analogues lack a 3' OH group and once they have been incorporated into the DNA chain they can no longer act as a substrate for further chain elongation, so that the growing DNA chain is terminated. Four reaction mixtures are set up, each with all four nucleotide triphosphates, one of which is labelled with phosphorus-32 and a low concentration of one of the four dideoxynucleoside triphosphates. This results in the production of a population of partially synthesised radioactive DNA molecules each with a common 5' end and each varying in length to a base specific 3' end as one of the dideoxynucleotides becomes incorporated. The use of four mixtures means that a dideoxynucleotide should be incorporated into every nucleotide position in at least one of the lengthening sequences, thereby stopping the process. Each of the DNA fragments produced will differ in size from the others. The DNA in each of the reaction

mixtures is then denatured and electrophoresed in adjacent lanes. The radioactive bands of DNA are detected by autoradiography and the sequence can then be read directly from the autoradiograph as the DNA chain sequence was stopped.

As stated at the beginning of the first article, these two introductory articles can do no more than outline some of the principles, methods, and terminology used in molecular biology. With the increasing use of these techniques in diagnosis and pathogenesis it is becoming critical for everyone with a scientific interest in respiratory medicine to understand the basic language of the science. We hope that these and the subsequent articles in this series will make the transition as painless as possible.

1 Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975;98:503–17.
2 Orkin SH, Little PFR, Kazazian HH, *et al.* Improved detection of the sickle mutation by DNA analysis. *N Engl J Med* 1982;307:32–6.
3 Kidd VJ, Globus MS, Wallace RB, *et al.* Prenatal diagnosis of α_1 antitrypsin deficiency by direct analysis of the mutation site in the gene. *N Engl J Med* 1984;310:639–42.
4 Kan YW, Dozy AM. Polymorphisms of DNA sequence adjacent to human beta-globin structural gene: relation to sickle mutation. *Proc Natl Acad Sci* 1978;75:5631–5.
5 Cookson WOCM, Sharp PA, Faux JA, *et al.* Linkage between immunoglobulin E responses underlying asthma and rhinitis and chromosome 11q. *Lancet* 1989;i:1292–4.
6 Lathrops GM, Lalouel JM. Easy calculations of LOD scores and genetic risks on small computers. *Am J Hum Genet* 1984;36:460–5.
7 Dixon RAF, Sigal IS, Candelore MR, *et al.* Structures required for ligand binding to the beta adrenergic receptor. *EMBO Journal* 1987;6:3269–75.
8 Strader CD, Sigal IS, Register RB, *et al.* Identification of residues required for ligand binding to the beta adrenergic receptor. *Proc Natl Acad Sci* 1987;84:4384–8.
9 Courtney M, Jallat S, Tessier LH, *et al.* Synthesis in *E coli* of α_1 antitrypsin variants of therapeutic potential for emphysema and thrombosis. *Nature (Lond)* 1985; 313:149–51.
10 Verbanac KM, Heath EC. Biosynthesis, processing and secretion of M and Z variant α_1 antitrypsin. *J Biol Chem* 1986;261:9979–89.
11 Sanger F, Nicklen S, Coulsen AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci* 1977; 74:5463–7.

Glossary

Anneal Joining together of two complementary strands of nucleic acids by hydrogen bonding between opposite base pairs.

Bacteriophage Bacterial virus that can be used as a cloning vehicle.

Centimorgan (cM) Unit of measurement for the distance between two loci on a genetic map.

Cloning Introduction of fragments of DNA into bacteria to produce millions of identical copies (clones).

Cosmid Hybrid cloning vector composed of plasmid and bacteriophage, which is used to clone large fragments of DNA.

Denaturation Dissociation of two complementary strands of nucleic acid by heating or treatment with alkali.

Hybridisation See "Anneal."

Klenow enzyme Fragment of the enzyme DNA polymerase that catalyses DNA synthesis but lacks the 5'→3' exonuclease activity of the original enzyme.

Map unit (mu) See "Centimorgans."

Northern blotting Technique for identification of the presence and size of mRNA.

Oligonucleotide directed mutagenesis Technique in which a specific nucleotide of a gene is mutated for studying the structure-function relationships of complex proteins.

Penetrance The frequency with which a gene mani-

fests itself in the phenotype of an individual.

Plasmid Self replicating circular, double stranded DNA molecule found in bacteria: the most commonly used cloning vehicle.

Polymorphism One of two or more alternative nucleotide sequences at a genetic locus.

Probe Single stranded DNA or RNA used to identify complementary single stranded DNAs or RNAs in complex mixtures of nucleic acids.

Recombinant DNA molecule New DNA molecule constructed in vitro from two or more different DNA sequences.

Restriction map Map of the location of all possible cleavage sites for restriction endonucleases relative to each other on the DNA being analysed.

Restriction endonuclease Enzyme that recognises a specific short nucleotide sequence of DNA and cleaves the DNA at or close to this site.

Restriction fragment length polymorphism DNA polymorphism that results in either the loss or the gain of a restriction site for a restriction enzyme and hence alters the pattern of fragments obtained after digestion of the DNA with this restriction enzyme.

Reverse transcriptase Enzyme that catalyses the synthesis of single stranded complementary DNA (cDNA) with messenger RNA as the template.

Southern blotting Technique to identify the presence and size of DNA molecules.

Transformation Process in which extracellular DNA is taken up by a cell.