

See discussions, stats, and author profiles for this publication at:  
<https://www.researchgate.net/publication/226072152>

# Basic Techniques in Molecular Biology

Chapter · October 2007

DOI: 10.1385/1-59259-870-6:001

---

CITATIONS

8

READS

21,635

**1 author:**



[Ralph Rapley](#)

University of Hertfordshire

**155** PUBLICATIONS **314** CITATIONS

SEE PROFILE

## CHAPTER 1

# Basic Molecular Biology Techniques

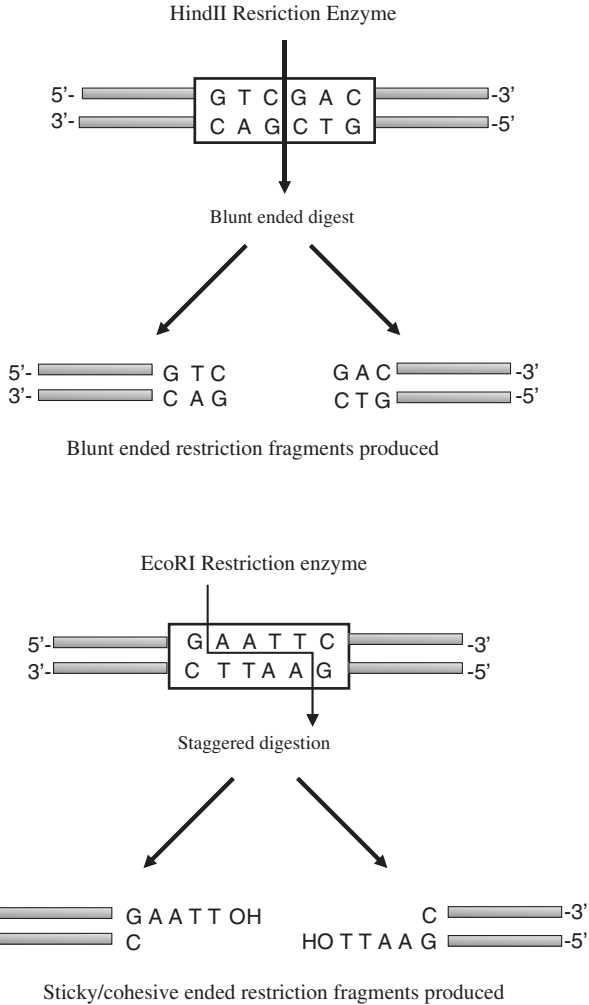
RALPH RAPLEY

School of Life Sciences, University of Hertfordshire, Hatfield, Hertfordshire  
AL10 9AB, UK

### 1.1 ENZYMES USED IN MOLECULAR BIOLOGY

The discovery and characterisation of a number of key enzymes have permitted the development of various techniques for the analysis and manipulation of DNA. In particular, the enzymes termed type II restriction endonucleases have come to play a key role in all aspects of molecular biology.<sup>1</sup> These enzymes recognise specific DNA sequences, usually 4–6 base pairs (bp) in length, and cleave them in a defined manner. The sequences recognised are palindromic or of an inverted repeat nature, that is, they read the same in both directions on each strand. When cleaved they leave a flush-ended or staggered (also termed a cohesive-ended) fragment depending on the particular enzyme used (Figure 1.1).

An important property of staggered ends is that those produced from different molecules by the same enzyme are complementary (or ‘sticky’) and so will anneal to each other (Table 1.1). The annealed strands are held together only by hydrogen bonding between complementary bases on opposite strands. Covalent joining of ends on each of the two strands may be brought about by the enzyme DNA ligase. This is widely exploited in molecular biology to allow the construction of recombinant DNA, *i.e.* the joining of DNA fragments from different sources.



**Figure 1.1** Examples of digestion of DNA by restriction endonucleases. The upper panel indicates the result of a restriction digestion forming blunt fragments with the enzyme *HindIII*. The bottom panel indicates the cohesive fragments produced by digestion with the enzyme *EcoRI*.

Approximately 500 restriction enzymes have been characterised that recognise over 100 different target sequences. A number of these, termed isoschizomers, recognise different target sequences but produce the same staggered ends or overhangs. A number of other enzymes have proved to be of value in the manipulation of DNA, as summarised in Table 1.2, and are indicated at appropriate points within the text.

**Table 1.1** Examples of restriction endonucleases that recognise different target sequences and the resulting fragments following digestion.

Name	Recognition Sequence	Digestion Products
Four Nucleotide Recognition Sequence		
<i>HaeIII</i>	$\begin{array}{c} \downarrow \\ 5' - \text{GGCC} - 3' \\ 3' - \text{CCGG} - 5' \end{array}$	$\begin{array}{ccc} 5' - \text{GG} & \text{CC} - 3' & \text{Blunt End} \\ 3' - \text{CC} & \text{GG} - 5' & \text{Digestion} \end{array}$
<i>HpaII</i>	$\begin{array}{c} 5' - \text{CCGG} - 3' \\ 3' - \text{GGCC} - 5' \end{array}$	$\begin{array}{ccc} 5' - \text{C} & \text{CCG} - 3' & \text{Cohesive End} \\ 3' - \text{GGC} & \text{C} - 5' & \text{Digestion} \end{array}$
Six Nucleotide Recognition Sequence		
<i>BamHI</i>	$\begin{array}{c} \downarrow \\ 5' - \text{GGATTC} - 3' \\ 3' - \text{GGCC} - 5' \end{array}$	$\begin{array}{ccc} 5' - \text{G} & \text{GATTC} - 3' \\ 3' - \text{CCTAG} & \text{G} - 5' \end{array}$
<i>EcoRI</i>	$\begin{array}{c} \downarrow \\ 5' - \text{GAATTC} - 3' \\ 3' - \text{CTTAAG} - 5' \end{array}$	$\begin{array}{ccc} 5' - \text{G} & \text{AATTC} - 3' \\ 3' - \text{CTTAA} & \text{G} - 5' \end{array}$
<i>HindIII</i>	$\begin{array}{c} \downarrow \\ 5' - \text{AAGCTT} - 3' \\ 3' - \text{TTCGAA} - 5' \end{array}$	$\begin{array}{ccc} 5' - \text{A} & \text{AGCTT} - 3' \\ 3' - \text{TTCGA} & \text{A} - 5' \end{array}$
Eight Nucleotide Recognition Sequence		
<i>NotI</i>	$\begin{array}{c} \downarrow \\ 5' - \text{GCGGCCGC} - 3' \\ 3' - \text{CGCCGGCG} - 5' \end{array}$	$\begin{array}{ccc} 5' - \text{GC} & \text{GGCCGC} - 3' \\ 3' - \text{CGCCGG} & \text{CG} - 5' \end{array}$

**Table 1.2** Comparison of the various labelling methods for DNA.

Labelling method	Enzyme	Probe Type	Specific Activity
5' end labelling	Alkaline Phosphatase Polynucleotide Kinase	DNA	Low
3' end labelling	Terminal Transferase	DNA	Low
Nick Translation	DNase I DNA Polymerase I	DNA	High
Random Hexamer	DNA Polymerase I	DNA	High
PCR	<i>Taq</i> DNA Polymerase	DNA	High
Riboprobes (cRNA)	RNA Polymerase	RNA	High

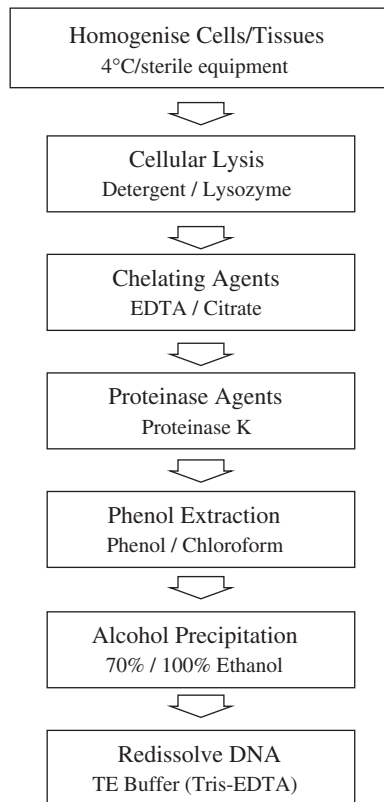
## 1.2 ISOLATION AND SEPARATION OF NUCLEIC ACIDS

### 1.2.1 Isolation of DNA

The use of DNA for analysis or manipulation usually requires that it is isolated and purified to a certain extent. DNA is recovered from cells by the gentlest possible method of cell rupture to prevent the DNA from fragmenting by mechanical shearing. This is usually in the presence of

EDTA which chelate the  $Mg^{2+}$  ions needed for enzymes that degrade DNA termed DNase. Ideally, cell walls, if present, should be digested enzymatically (*e.g.* lysozyme treatment of bacteria) and the cell membrane should be solubilised using detergent. If physical disruption is necessary, it should be kept to a minimum and should involve cutting or squashing of cells, rather than the use of shear forces. Cell disruption (and most subsequent steps) should be performed at 4 °C, using glassware and solutions which have been autoclaved to destroy DNase activity (Figure 1.2).

After release of nucleic acids from the cells, RNA can be removed by treatment with ribonuclease (RNase) which has been heat treated to inactivate any DNase contaminants; RNase is relatively stable to heat as a result of its disulfide bonds, which ensure rapid renaturation of the molecule on cooling. The other major contaminant, protein, is removed by shaking the solution gently with water-saturated phenol or with a



**Figure 1.2** Flow diagram of the main steps involved in the extraction of DNA.

phenol–chloroform mixture, either of which will denature proteins but not nucleic acids. Centrifugation of the emulsion formed by this mixing produces a lower, organic phase, separated from the upper, aqueous phase by an interface of denatured protein. The aqueous solution is recovered and deproteinised repeatedly, until no more material is seen at the interface. Finally, the deproteinised DNA preparation is mixed with two volumes of absolute ethanol and the DNA allowed to precipitate out of solution in a freezer. After centrifugation, the DNA pellet is redissolved in a buffer containing EDTA to inactivate any DNases present. This solution can be stored at 4 °C for at least 1 month. DNA solutions can be stored frozen, although repeated freezing and thawing tend to damage long DNA molecules by shearing.

The procedure described above is suitable for total cellular DNA. If the DNA from a specific organelle or viral particle is needed, it is best to isolate the organelle or virus before extracting its DNA, since the recovery of a particular type of DNA from a mixture is usually rather difficult. Where a high degree of purity is required, DNA may be subjected to density gradient ultracentrifugation through caesium chloride, which is particularly useful for the preparation of plasmid DNA.<sup>2</sup> It is possible to check the integrity of the DNA by agarose gel electrophoresis and determine the concentration of the DNA by using the fact that 1 absorbance unit equates to 50  $\mu\text{g ml}^{-1}$  of DNA and so

$$50 \times A_{260} = \text{concentration of DNA sample } (\mu\text{g ml}^{-1}) \quad (1)$$

Contaminants may also be identified by scanning UV spectrophotometry from 200 to 300 nm. A ratio of 260 nm:280 nm of approximately 1.8 indicates that the sample is free of protein contamination, which absorbs strongly at 280 nm.

### 1.2.2 Isolation of RNA

The methods used for RNA isolation are very similar to those described above for DNA; however, RNA molecules are relatively short and therefore less easily damaged by shearing, so cell disruption can be more vigorous. RNA is, however, very vulnerable to digestion by RNases, which are present endogenously in various concentrations in certain cell types and exogenously on fingers. Gloves should therefore be worn and a strong detergent should be included in the isolation medium to denature immediately any RNases. Subsequent deproteinisation should be particularly rigorous, since RNA is often tightly associated with proteins.<sup>3</sup> DNase treatment can be used to remove DNA and RNA can

be precipitated by ethanol. One reagent in which is commonly used in RNA extraction is guanadinium thiocyanate, which is both a strong inhibitor of RNase and a protein denaturant. It is possible to check the integrity of an RNA extract by analysing it by agarose gel electrophoresis. The most abundant RNA species are rRNA molecules, 23S and 16S for prokaryotes and 18S and 28S for eukaryotes. These appear as discrete bands on the agarose gel and indicate that the other RNA components are likely to be intact. This is usually carried out under denaturing conditions to prevent secondary structure formation in the RNA. The concentration of the RNA may be estimated by using UV spectrophotometry. At 260 nm, 1 absorbance unit equates to  $40 \mu\text{g ml}^{-1}$  of RNA and therefore

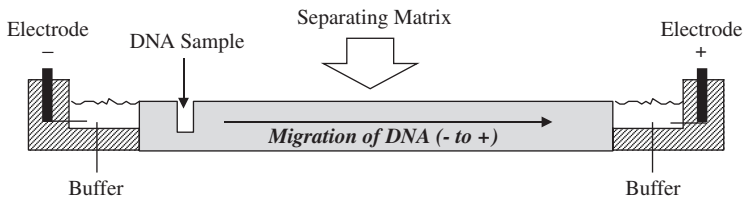
$$40 \times A_{260} = \text{concentration of RNA sample } (\mu\text{g ml}^{-1}) \quad (2)$$

Contaminants may also be identified in the same way as for DNA by scanning UV spectrophotometry; however, in the case of RNA a 260 nm:280 nm ratio of approximately 2 would be expected for a sample containing no contamination.

In many cases, it is desirable to isolate eukaryotic mRNA, which constitutes only 2–5% of cellular RNA, from a mixture of total RNA molecules. This may be carried out by affinity chromatography on oligo (dT)-cellulose columns. At high salt concentrations, the mRNA containing poly(A) tails binds to the complementary oligo(dT) molecules of the affinity column and so mRNA will be retained; all other RNA molecules can be washed through the column with further high-salt solution. Finally, the bound mRNA can be eluted using a low concentration of salt. Nucleic acid species may also be subfractionated by more physical means such as electrophoretic or chromatographic separations based on differences in nucleic acid fragment sizes or physicochemical characteristics.<sup>3</sup>

### 1.3 ELECTROPHORESIS OF NUCLEIC ACIDS

Electrophoresis in agarose or polyacrylamide gels is the most usual way to separate DNA molecules according to size (Figure 1.3). The technique can be used analytically or preparatively and can be qualitative or quantitative. Large fragments of DNA such as chromosomes may also be separated by a modification of electrophoresis termed pulsed field gel electrophoresis (PFGE). The easiest and most widely applicable method is electrophoresis in horizontal agarose gels, followed by staining with



**Figure 1.3** A typical setup required for agarose gel electrophoresis of DNA. The upper panel indicates a cross-section of the unit used for gel electrophoresis.

ethidium bromide. This dye binds to DNA by insertion between stacked base pairs (intercalation) and it exhibits a strong orange/red fluorescence when illuminated with ultraviolet light. Very often electrophoresis is used to check the purity and intactness of a DNA preparation or to assess the extent of an enzymatic reaction during, for example, the steps involved in the cloning of DNA. For such checks ‘mini-gels’ are particularly convenient, since they need little preparation, use small samples and give results quickly. Agarose gels can be used to separate molecules larger than about 100 bp. For higher resolution or for the effective separation of shorter DNA molecules, polyacrylamide gels are the preferred method.<sup>4</sup>

When electrophoresis is used preparatively, the piece of gel containing the desired DNA fragment is physically removed with a scalpel. The DNA may be recovered from the gel fragment in various ways. This may include crushing with a glass rod in a small volume of buffer, using agarase to digest the agarose thus leaving the DNA, or by the process of electroelution. In the latter method, the piece of gel is sealed in a length of dialysis tubing containing buffer and is then placed between two electrodes in a tank containing more buffer. Passage of an electric current between the electrodes causes DNA to migrate out of the gel piece, but it remains trapped within the dialysis tubing and can therefore be recovered easily. More commonly, commercial spin columns can be used which contain an isolating matrix used in conjunction with a bench-top microcentrifuge. The use of such standardised ‘kits’ in molecular biology is now commonplace. An alternative to conventional analysis of nucleic acids by electrophoresis is through the use of microfluidic systems. These are carefully manufactured chip-based units where microlitre volumes may be used and with the aid of computer analysis provide much of the data necessary for analysis. Their advantage lies in the fact that the sample volume is very small, allowing much of an extract to be used for further analysis.<sup>5</sup>

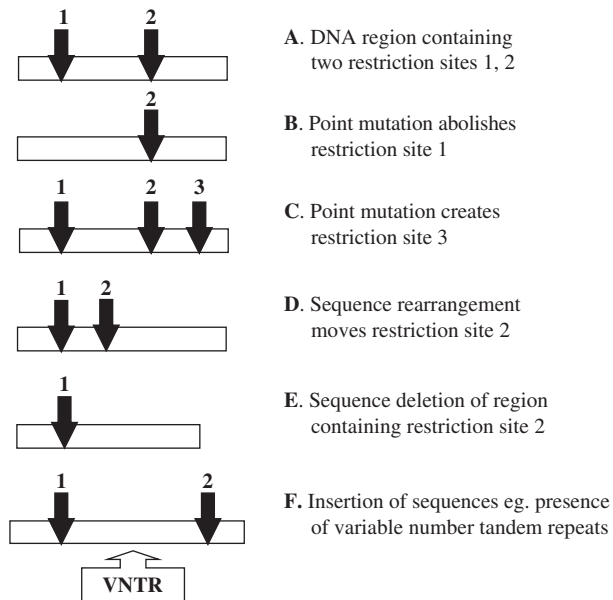


## 1.4 RESTRICTION MAPPING OF DNA FRAGMENTS

Restriction mapping involves the size analysis of restriction fragments produced by several restriction enzymes individually and in combination.<sup>6</sup> The principle of this mapping is illustrated, in which the restriction sites of two enzymes, A and B, are being mapped. Cleavage with A gives fragments 2 and 7 kilobases (kb) from a 9 kb molecule, hence we can position the single A site 2 kb from one end. Similarly, B gives fragments 3 and 6 kb, so it has a single site 3 kb from one end; but it is not possible at this stage to say if it is near to A's site or at the opposite end of the DNA. This can be resolved by a double digestion. If the resultant fragments are 2, 3 and 4 kb, then A and B cut at opposite ends of the molecule; if they are 1, 2 and 6 kb, the sites are near each other. Not surprisingly, the mapping of real molecules is rarely as simple as this and computer analysis of the restriction fragment lengths is usually needed to construct a map (Figure 1.4).

## 1.5 NUCLEIC ACID ANALYSIS METHODS

There are numerous methods for analysing DNA and RNA; however, many of them are solution based or more recently include the use of

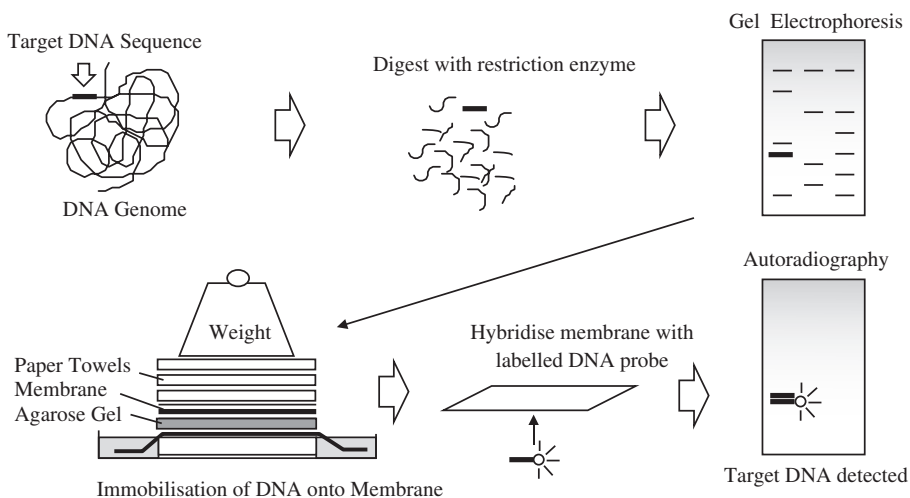


**Figure 1.4** Restriction fragment length polymorphisms (RFLP). The schematic panels A–F indicate the various fragments obtained following digestion as a result of differences in the position of restriction endonuclease target sequences.

chip-based array systems. Indeed, the lab-on-a-chip approach is developing rapidly and it is possible to envisage many detection and analysis methods being developed in this format in the future.<sup>7</sup> However, traditional methods are still employed in many laboratories and much is still made of producing a hard copy of digested and separated single-stranded DNA fragments attached to a matrix such as nylon for analysis with an appropriate labelled probe.

### 1.5.1 DNA Blotting

Electrophoresis of DNA restriction fragments allows separation based on size to be carried out; however, it provides no indication as to the presence of a specific, desired fragment among the complex sample (Figure 1.5). This can be achieved by transferring the DNA from the intact gel on to a piece of nitrocellulose or nylon membrane placed in contact with it.<sup>8</sup> This provides a more permanent record of the sample since DNA begins to diffuse out of a gel that is left for a few hours. First the gel is soaked in alkali to render the DNA single stranded. It is then transferred to the membrane so that the DNA becomes bound to the it in exactly the same pattern as that originally on the gel. This transfer, named a Southern blot after its inventor Ed Southern, can be performed electrophoretically or by drawing large volumes of buffer through both gel and membrane, thus transferring DNA from one to the other by



**Figure 1.5** The steps involved in the production of a Southern blot and the subsequent detection of a specific DNA sequence following hybridisation with a complementary labelled gene probe.

capillary action. The point of this operation is that the membrane can now be treated with a labelled DNA molecule, for example a gene probe. This single-stranded DNA probe will hybridise under the right conditions to complementary fragments immobilised on the membrane. The conditions of hybridisation, including the temperature and salt concentration, are critical for this process to take place effectively. This is usually referred to as the stringency of the hybridisation and it is particular for each individual gene probe and for each sample of DNA. A series of washing steps with buffer are then carried out to remove any unbound probe and the membrane is developed, after which the precise location of the probe and its target may be visualised. It is also possible to analyse DNA from different species or organisms by blotting the DNA and then using a gene probe representing a protein or enzyme from one of the organisms. In this way, it is possible to search for related genes in different species. This technique is generally termed Zoo blotting.

### 1.5.2 RNA Blotting

The same basic process of nucleic acid blotting can be used to transfer RNA from gels on to similar membranes. This allows the identification of specific mRNA sequences of a defined length by hybridisation to a labelled gene probe and is known as Northern blotting.<sup>9</sup> With this technique it is not only possible to detect specific mRNA molecules but it may also be used to quantify the relative amounts of the specific mRNA. It is usual to separate the mRNA transcripts by gel electrophoresis under denaturing conditions since this improves resolution and allows a more accurate estimation of the sizes of the transcripts. The format of the blotting may be altered from transfer from a gel to direct application to slots on a specific blotting apparatus containing the nylon membrane. This is termed slot or dot blotting and provides a convenient means of measuring the abundance of specific mRNA transcripts without the need for gel electrophoresis; it does not, however, provide information regarding the size of the fragments.

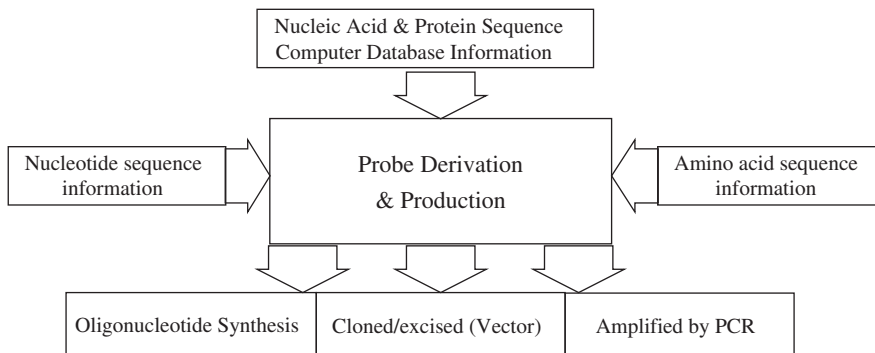
A further method of RNA analysis that overcomes the problems of RNA blotting is termed the ribonuclease protection assay. Here the RNA from a sample is extracted and then mixed with a probe representing the sequence of interest in solution. The probe and the appropriate RNA fragment hybridise to form a double-stranded sequence. RNase is then added, which cleaves any single-stranded RNA present but leaves the double-stranded RNA intact. The intact RNA can then be separated by electrophoresis and an indication of the size of the fragment generated. The efficient removal of the background of RNA and

the improved sensitivity make the ribonuclease protection assay a popular choice for the analysis of specific RNA molecules.

An important step in the field of RNA analysis was the development of RNAi (RNA interference), which inhibits gene expression. Here double-stranded RNA promotes the degradation of mRNA. Double-stranded RNA in the cell is cleaved by a dicer enzyme, resulting in the formation of small 21–25 bp interfering RNAs (siRNA). The siRNA are complementary to a target RNA strand. Small RNAi proteins are guided by the siRNA to the appropriate mRNA, where the target is then cleaved and is unable to be translated. Many areas are now benefiting from the adoption of this technique in the molecular biology and biotechnology fields.<sup>10</sup>

## 1.6 GENE PROBE DERIVATION

The availability of a gene probe is essential in many molecular biology techniques, yet in many cases is one of the most difficult steps (Figure 1.6). The information needed to produce a gene probe may come from many sources, but with the development and sophistication of genetic databases this is usually one of the first stages.<sup>11</sup> There are a number of genetic databases throughout the world and it is possible to search these over the internet and identify particular sequences relating to a specific gene or protein. In some cases it is possible to use related proteins from the same gene family to gain information on the most useful DNA sequence. Similar proteins or DNA sequences but from different species may also provide a starting point with which to produce a so-called heterologous gene probe. Although in some cases probes have already been produced and cloned, it is possible, armed with a DNA sequence from a DNA



**Figure 1.6** Alternative strategies for designing and producing a gene probe.

database, to synthesise chemically a single-stranded oligonucleotide probe. This is usually undertaken by computer-controlled gene synthesisers which link dNTPs together based on a desired sequence. It is essential to carry out certain checks before probe production to determine that the probe is unique, is not able to self-anneal or is self-complementary, all of which may compromise its use.<sup>12</sup>

Where little DNA information is available to prepare a gene probe, it is possible in some cases to use the knowledge gained from analysis of the corresponding protein. Thus it is possible to isolate and purify proteins and sequence part of the N-terminal end of the protein. From our knowledge of the genetic code, it is possible to predict the various DNA sequences that could code for the protein and then synthesise appropriate oligonucleotide sequences chemically. Due to the degeneracy of the genetic code, most amino acids are coded for by more than one codon, hence there will be more than one possible nucleotide sequence which could code for a given polypeptide. The longer the polypeptide, the greater is the number of possible oligonucleotides which must be synthesised. Fortunately, there is no need to synthesise a sequence longer than about 20 bases, since this should hybridise efficiently with any complementary sequences and should be specific for one gene. Ideally, a section of the protein should be chosen which contains as many tryptophan and methionine residues as possible, since these have unique codons and there will therefore be fewer possible base sequences which could code for that part of the protein. The synthetic oligonucleotides can then be used as probes in a number of molecular biology methods.

## 1.7 LABELLING DNA GENE PROBE MOLECULES

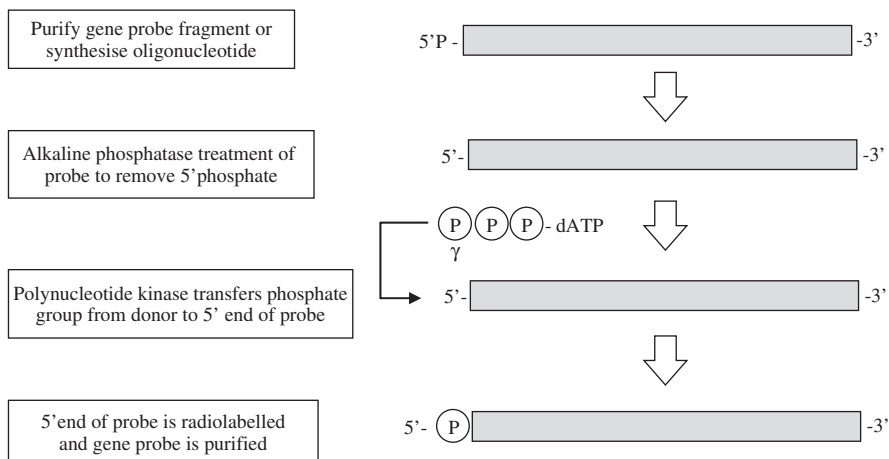
An essential feature of a gene probe is that it can be visualised by some means. In this way, a gene probe that hybridises to a complementary sequence may be detected and identify that desired sequence from a complex mixture. There are two main ways of labelling gene probes; traditionally it has been carried out using radioactive labels, but gaining in popularity are non-radioactive labels. Perhaps the most often used radioactive label is phosphorus-32 ( $^{32}\text{P}$ ), although for certain techniques sulfur-35 ( $^{35}\text{S}$ ) and tritium ( $^3\text{H}$ ) are used. These may be detected by the process of autoradiography, where the labelled probe molecule, bound to sample DNA, located for example on a nylon membrane, is placed in contact with an X-ray-sensitive film. Following exposure, the film is developed and fixed just as a black and white negative and reveals the precise location of the labelled probe and therefore the DNA to which it has hybridised.

Non-radioactive labels are increasingly being used to label DNA gene probes. Until recently, radioactive labels were more sensitive than their non-radioactive counterparts. However, recent developments have led to similar sensitivities, which, when combined with their improved safety, have led to their greater acceptance.

The labelling systems are termed either direct or indirect. Direct labelling allows an enzyme reporter such as alkaline phosphatase to be coupled directly to the DNA. Although this may alter the characteristics of the DNA gene probe, they offer the advantage of rapid analysis since no intermediate steps are needed. However, indirect labelling is at present more popular. This relies on the incorporation of a nucleotide which has a label attached. At present, three of the main labels in use are biotin, fluorescein and digoxigenin. These molecules are covalently linked to nucleotides using a carbon spacer arm of 7, 14 or 21 atoms. Specific binding proteins may then be used as a bridge between the nucleotide and a reporter protein such as an enzyme. For example, biotin incorporated into a DNA fragment is recognised with a very high affinity by the protein streptavidin. This may be either coupled or conjugated to a reporter enzyme molecule such as alkaline phosphatase. This is able to convert a colourless substrate, *p*-nitrophenol phosphate (PNPP), into a yellow compound, *p*-nitrophenol (PNP), and also offers a means of signal amplification. Alternatively labels such as digoxigenin incorporated into DNA sequences may be detected by monoclonal antibodies, again conjugated to reporter molecules including alkaline phosphatase. Thus, rather than the detection system relying on autoradiography, which is necessary for radiolabels, a series of reactions resulting in either a colour or a light or chemiluminescent reaction takes place. This has important practical implications since autoradiography may take 1–3 days, whereas colour and chemiluminescent reactions take minutes.

### 1.7.1 End Labelling of DNA Molecules

The simplest form of labelling DNA is by 5' or 3' end labelling; 5' end labelling involves a phosphate transfer or exchange reaction where the 5' phosphate of the DNA to be used as the probe is removed and in its place a labelled phosphate, usually  $^{32}\text{P}$ , is added. This is usually carried out by using two enzymes; the first, alkaline phosphatase, is used to remove the existing phosphate group from the DNA. Following removal of the released phosphate from the DNA, a second enzyme, polynucleotide kinase, is added, which catalyses the transfer of a phosphate group ( $^{32}\text{P}$ -labelled) to the 5' end of the DNA. The newly labelled probe



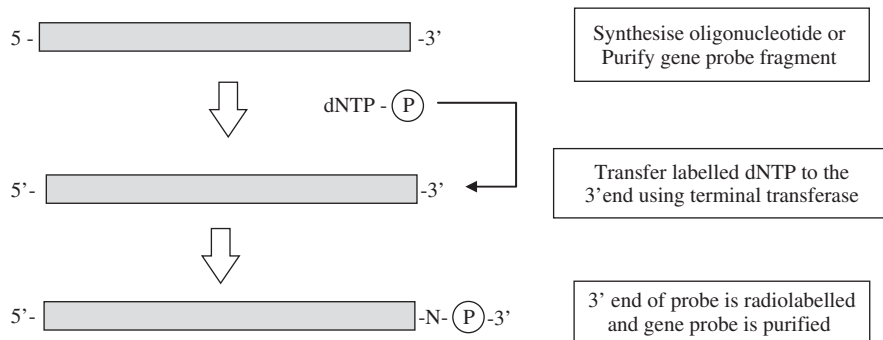
**Figure 1.7** The steps involved in the production of a 5'-labelled gene probe.

is then purified, usually by chromatography through a Sephadex column, and may be used directly (Figure 1.7).

Using the other end of the DNA molecule, the 3' end, is slightly less complex. Here a new dNTP which is labelled (*e.g.* [ $^{32}$ P]adATP or biotin-labelled dNTP) is added to the 3' end of the DNA by the enzyme terminal transferase. Although this is a simpler reaction, a potential problem exists because a new nucleotide is added to the existing sequence and so the complete sequence of the DNA is altered, which may affect its hybridisation to its target sequence. End labelling methods also suffer from the fact that only one label is added to the DNA so they are of a lower activity in comparison with methods that incorporate label along the length of the DNA (Figure 1.8).

### 1.7.2 Random Primer Labelling

In random primer labelling the DNA to be labelled is first denatured and then placed under renaturing conditions in the presence of a mixture of many different random sequences of hexamers or hexanucleotides. These hexamers will, by chance, bind to the DNA sample wherever they encounter a complementary sequence and so the DNA will rapidly acquire an approximately random sprinkling of hexanucleotides annealed to it. Each of the hexamers can act as a primer for the synthesis of a fresh strand of DNA catalysed by DNA polymerase since it has an exposed 3'-hydroxyl group. The Klenow fragment of DNA polymerase is used for random primer labelling because it lacks a 5'-3' exonuclease



**Figure 1.8** The steps involved in the production of 3'-labelled gene probe.

activity. This is prepared by cleavage of DNA polymerase with subtilisin, giving a large enzyme fragment which has no 5' to 3' exonuclease activity, but which still acts as a 5' to 3' polymerase. Thus, when the Klenow enzyme is mixed with the annealed DNA sample in the presence of dNTPs, including at least one which is labelled, many short stretches of labelled DNA will be generated (Figure 1.9). In a similar way to random primer labelling, the polymerase chain reaction may also be used to incorporate radioactive or non-radioactive labels.

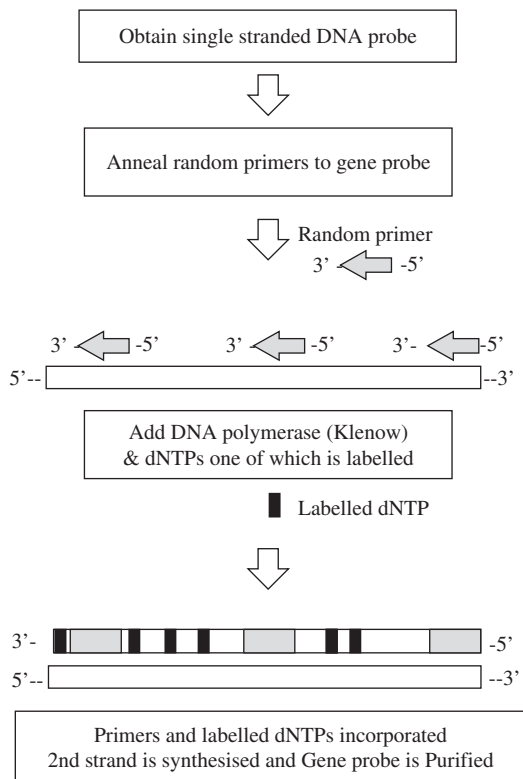
### 1.7.3 Nick Translation

A traditional method of labelling DNA is by the process of nick translation. Low concentrations of DNase I are used to make occasional single-strand nicks in the double-stranded DNA that is to be used as the gene probe. DNA polymerase then fills in the nicks, using an appropriate deoxyribonucleoside triphosphate (dNTP), at the same time making a new nick to the 3' side of the previous one. In this way, the nick is translated along the DNA. If labelled dNTPs are added to the reaction mixture, they will be used to fill in the nicks and so the DNA can be labelled to a very high specific activity (Figure 1.10).

## 1.8 THE POLYMERASE CHAIN REACTION

There have been a number of key developments in molecular biology techniques. However, one that has had the most impact in recent years has been the polymerase chain reaction (PCR). One of the reasons for the adoption of the PCR is the elegant simplicity of the reaction and relative ease of the practical manipulation steps. Frequently this is one of the first techniques to be used when analysing DNA and RNA and in

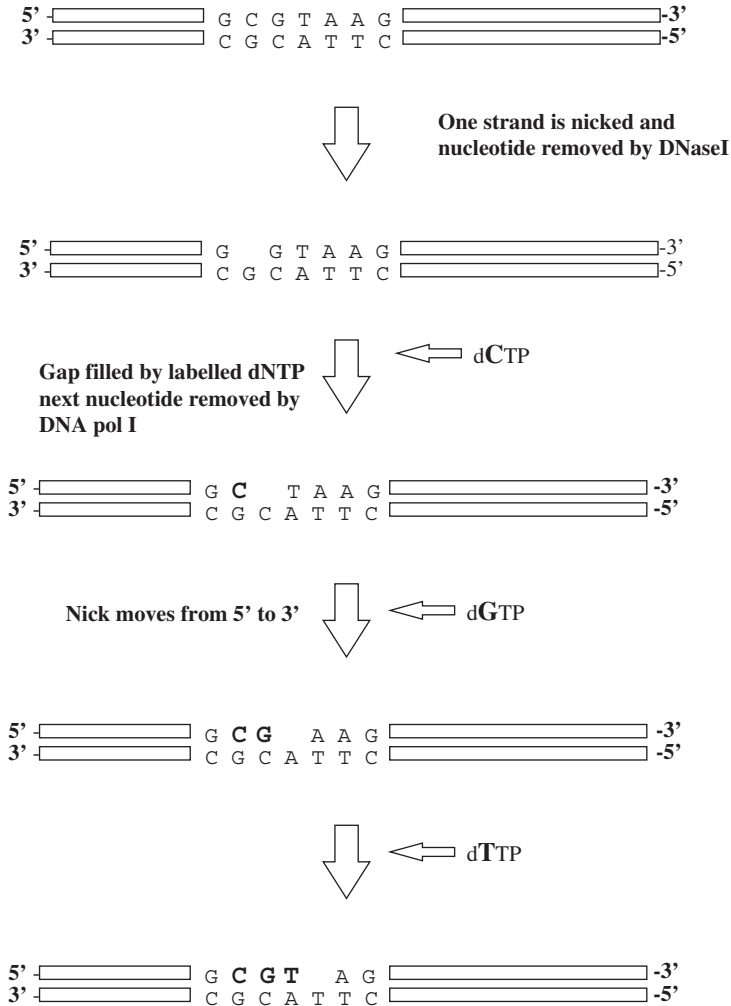




**Figure 1.9** The steps involved in the production of a gene probe produced by the random hexamer method.

its quantitation it has opened up the analysis of cellular and molecular processes to those outside the field of molecular biology.

The PCR is used to amplify a precise fragment of DNA from a complex mixture of starting material, usually termed the template DNA, and in many cases requires little DNA purification. It does require the knowledge of some DNA sequence information that flanks the fragment of DNA to be amplified (target DNA). From this information, two oligonucleotide primers may be chemically synthesised, each complementary to a stretch of DNA to the 3' side of the target DNA, one oligonucleotide for each of the two DNA strands. The result is an amplification of a specific DNA fragment which obviates the need for more time-consuming cloning procedures. The technique of the PCR is described in detail in Chapter 4. Further developments in molecular biology and biotechnology have allowed numerous genomes to be analysed and genes identified. It is not surprising that this has been aided



**Figure 1.10** The steps involved in the production of a gene probe by the nick translation method.

by the developments in the area of bioinformatics and whole genome analysis.<sup>13</sup> DNA databases and other nucleic acid sequence and protein analysis software may all be accessed over the internet given the relevant software and authority (Table 1.3). This is now relatively straightforward with web browsers that provide a user-friendly graphical interface for sequence manipulation. Consequently, the new expanding and exciting areas of bioscience research are those that analyse genome and DNA sequence databases, (genomics) and also their protein counterparts

**Table 1.3** Nucleic acid and protein database resources available over the Internet.

<i>Database or Resource</i>		<i>URL (uniform resource locator)</i>
General DNA Sequence Databases		
EMBL	European genetic database	<a href="http://www.ebi.ac.uk">http://www.ebi.ac.uk</a>
GenBank	US genetic database	<a href="http://ncbi.nlm.nih.gov">http://ncbi.nlm.nih.gov</a>
DDBJ	Japanese genetic database	<a href="http://ddbj.nig.ac.jp">http://ddbj.nig.ac.jp</a>
Protein Sequence Databases		
Swiss-Prot	European protein sequence database	<a href="http://expasy.hcuge.ch/sprot/sprot-top.html">http://expasy.hcuge.ch/sprot/sprot-top.html</a>
TREMBL	European protein sequence database	<a href="http://www.ebi.ac.uk/pub/databases/trembl">http://www.ebi.ac.uk/pub/databases/trembl</a>
PIR	US protein information resource	<a href="http://www-nbrf.gerogetown.edu/pir">http://www-nbrf.gerogetown.edu/pir</a>
Protein Structure Databases		
PDB	Brookhaven protein database	<a href="http://www.pdb.bnl.gov">http://www.pdb.bnl.gov</a>
NRL-3D	Protein structure database	<a href="http://www.gdb.org/Dan/proteins/nrl3d.html">http://www.gdb.org/Dan/proteins/nrl3d.html</a>
Genome Project Databases		
	Human Mapping Database Johns Hopkins, USA	<a href="http://gdbwww.gdb.org">http://gdbwww.gdb.org</a>
	dbEST (cDNA and partial sequences)	<a href="http://www.ncbi.nih.gov">http://www.ncbi.nih.gov</a>
	Genethon Genetic maps based on repeat markers	<a href="http://www.genethon.fr">http://www.genethon.fr</a>
	Whitehead Institute (YAC and physical maps)	<a href="http://www-genome.wi.mit.edu">http://www-genome.wi.mit.edu</a>

(proteomics). This is sometimes referred to as *in silico* research, and there is no doubt that for basic and biotechnological research it is as important to have Internet and database access as it is to have equipment and reagents for laboratory molecular biology.<sup>14</sup>

## REFERENCES

1. H. O. Smith and K. W. Wilcox, *J. Mol. Biol.*, 1970, **51**, 379.
2. K. Wilson and J. M. Walker, *Principles and Techniques of Practical Biochemistry*, Cambridge University Press, Cambridge, 2005.
3. R. E. Farrell, *RNA Methodologies. A Laboratory Guide for Isolation and Characterisation*, Academic Press, New York, 2005.
4. R. Rapley and D. L. Manning, *RNA Isolation and Characterisation Protocols*, Humana Press, Totowa, NJ, 1998.
5. P. Jones, *Gel Electrophoresis of Nucleic Acids*, Wiley, Chichester, 1995.

6. G. J. S. Jenkins, in *Molecular Biomethods Handbook*, 2nd edn, ed. R. Rapley and J. M. Walker, Humana Press, Totowa, NJ, 2008, pp. 17–27.
7. L. Chen, A. Manz and P. J. Day, *Lab Chip*, 2007, **7**, 1413.
8. E. M. Southern, *J. Mol. Biol.*, 1975, **98**, 503.
9. J. C. Alwine, D. J. Kemp and G. R. Stark, *Proc. Natl. Acad. Sci. USA*, 1977, **74**, 5350.
10. M. Aquino de Muro, in *Molecular Biomethods Handbook*, 2nd edn, ed. R. Rapley and J. M. Walker, Humana Press, Totowa, NJ, 2008, pp. 41–54.
11. G. J. Hannon, *RNAi. A Guide to Gene Silencing*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2003.
12. T. P. McCreery and T. R. Barrette, in *Molecular Biomethods Handbook*, ed. R. Rapley and J. M. Walker, Humana Press, Totowa, NJ, 1998, pp. 73–76.
13. A. D. Baxevanis and B. F. Ouelette, *Bioinformatics. A Practical Guide to the Analysis of Genes and Proteins*, Wiley, Chichester, 2004.
14. D. W. Mount, *Bioinformatics Sequence and Genome Analysis*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2004.