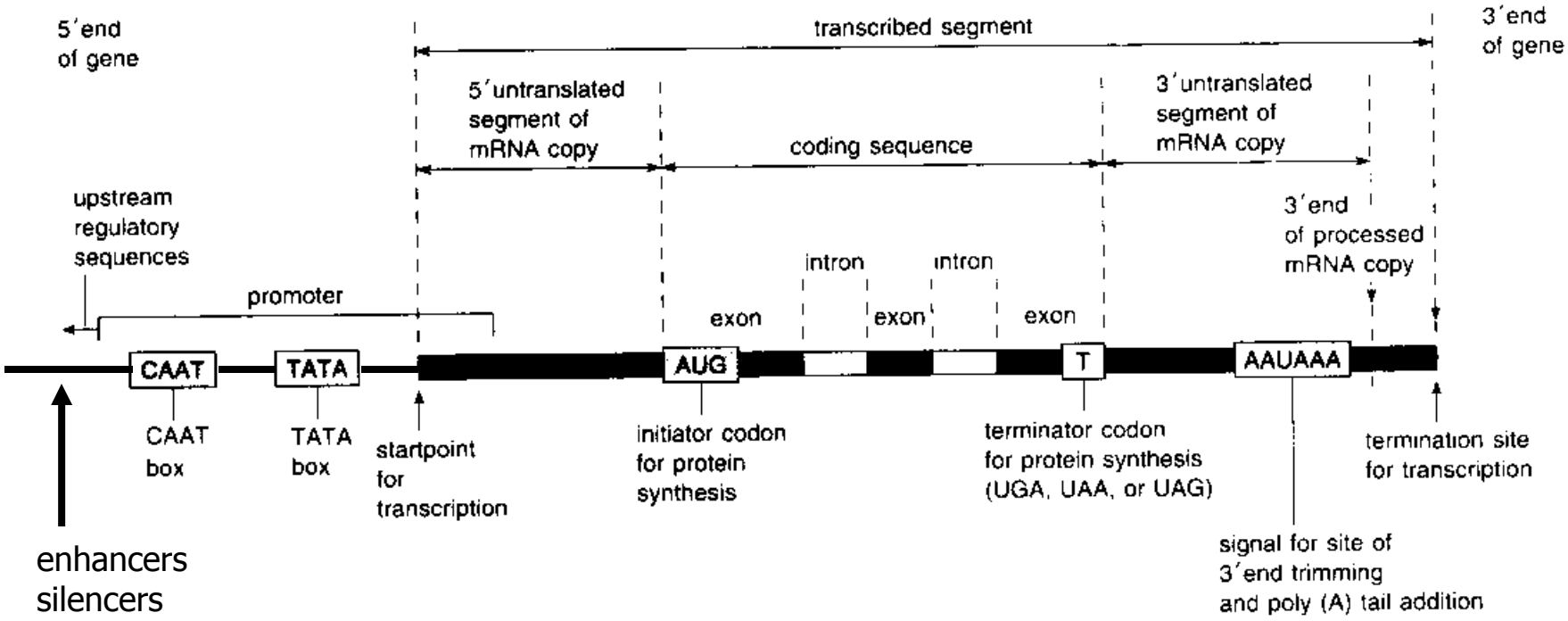
Genomic libraries

Eukaryotic gene organization



Recombinant DNA Libraries

Collection of many clones derived from a single DNA source.

Genomic Libraries:

Many clones, each of which contains a fragment of chromosomal DNA from a particular species.

Complete genomic library: Entire genome is represented in at least one clone.

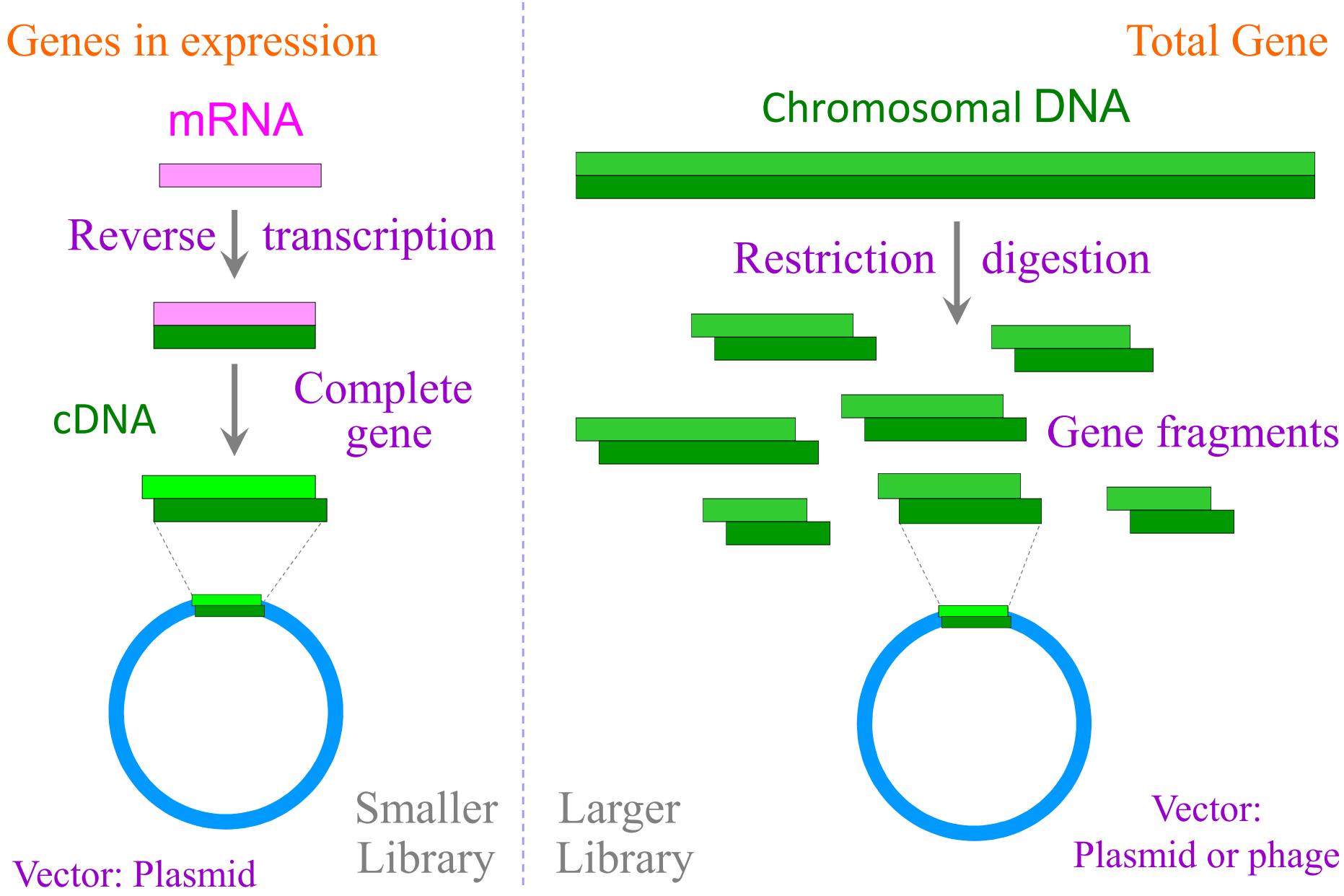
cDNA Libraries:

cDNAs = DNA copies of RNA molecules.

cDNA libraries: Each clone contains DNA copy of an individual mRNA.

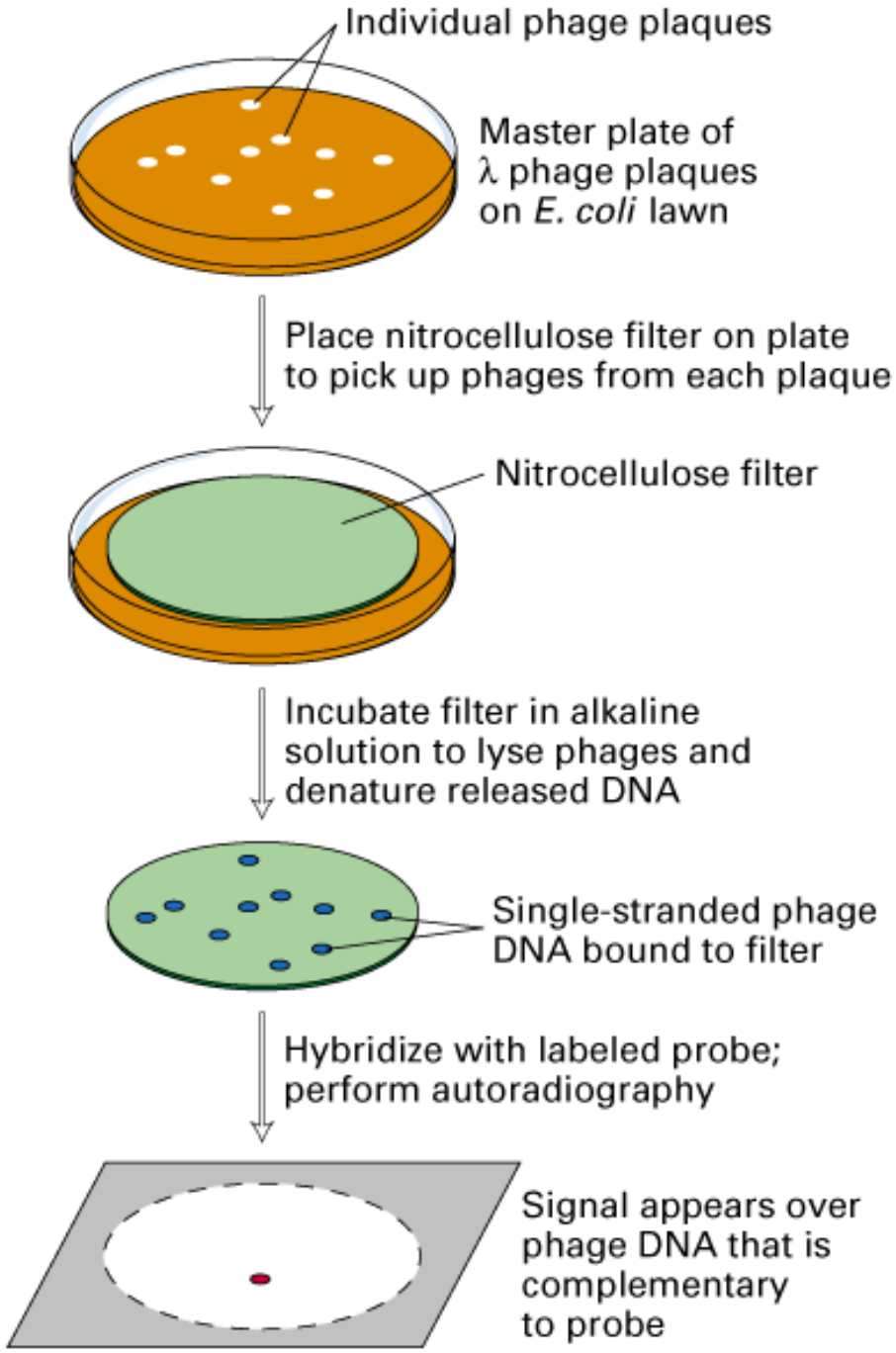
Very useful for studying just the part of a gene that is present in mRNA.

Two Libraries： cDNA Library vs Genomic Library



Screening a genomic library using DNA hybridization to a (radio-)labeled DNA probe

Note: a cDNA is commonly (radio-)labeled and used as a DNA probe to screen a genomic library



Genomic library

- A Gene library is a set of cloned fragments that collectively represent the genes of a particular organism.
- Particular genes can be isolated from DNA libraries
- Both genomic and cDNA libraries can be screened by hybridization

Construction of gene libraries (genomic and cDNA)

A **DNA library or gene library** is a collection of DNA clones, gathered together as a source of DNA for sequencing, gene discovery, or gene function studies.

Depending on DNA source, library can be grouped as **genomic library** or **cDNA library**.

Genomic library are produced when the complete genome of a particular organism is cleaved into thousands of fragments, and all the fragments are cloned by insertion into a cloning vector.

- The first step in preparing a genomic library is partial digestion of the DNA by restriction endonucleases, such that any given sequence will appear in fragments of a range of sizes that are compatible with the cloning vector and ensures that virtually all sequences are represented among the clones in the library.
- Fragments that are too large or too small for cloning are removed by centrifugation or electrophoresis.

- The cloning vector, such as a BAC or YAC plasmid, is cleaved with the same restriction endonuclease and ligated to the genomic DNA fragments.
- The ligated DNA mixture is then used to transform bacterial or yeast cells to produce a library of cell types, each type harboring a different recombinant DNA molecule. Ideally, all the DNA in the genome under study will be represented in the library.

cDNA libraries

- **Complementary DNA (cDNA) libraries are generated by the reverse transcription of mRNA**
  - cDNA is representative of the mRNA population, and therefore reflects mRNA levels and the diversity of splice isoforms in particular tissues

cDNA clones are copies of mRNAs

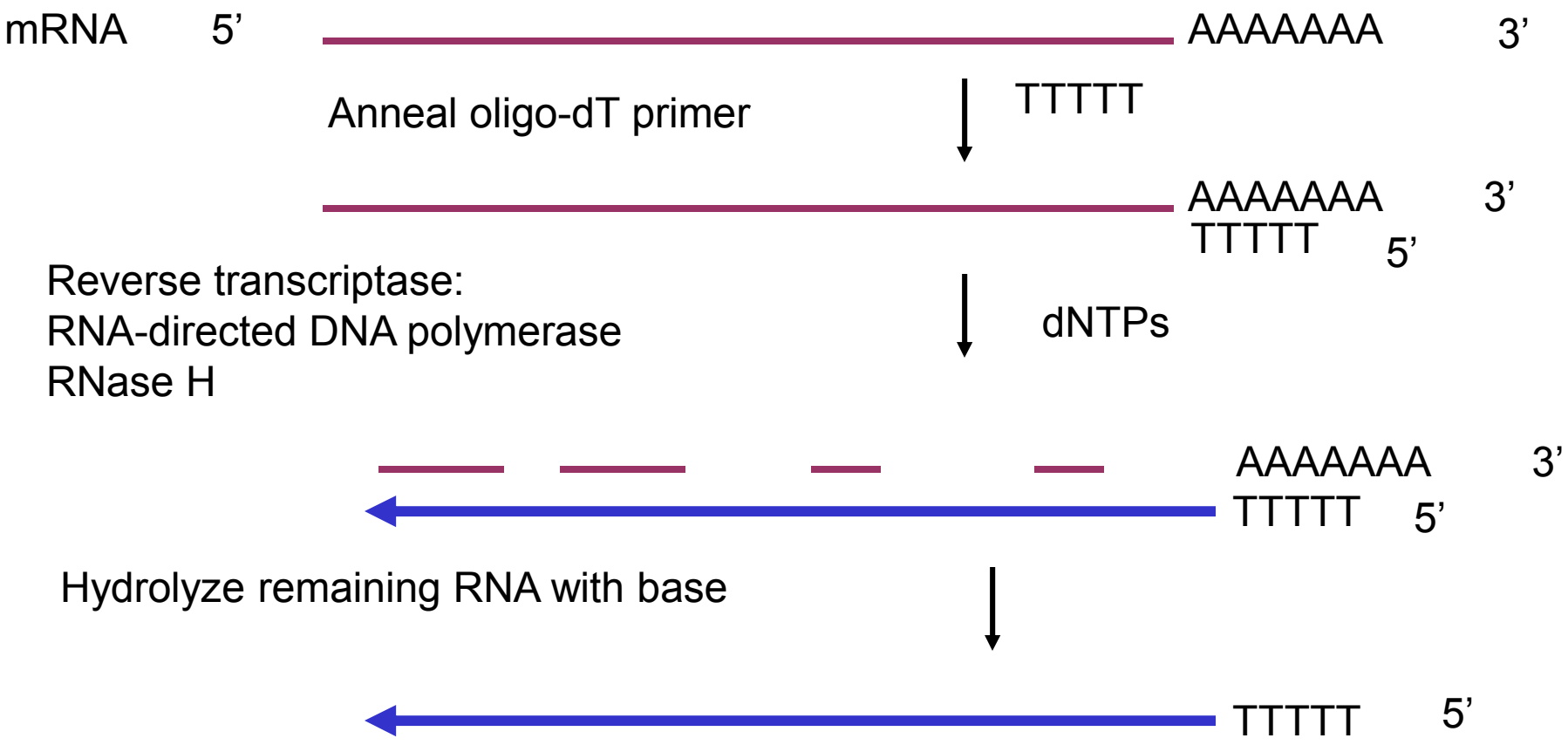
- Much of the genomic DNA is **not** expressed as mRNA
- Many issues about gene function are best addressed by examining the product that they encode.
- The cDNA copies of mRNA contain primarily sequences that encode protein.

Therefore, cDNA clones are useful for many studies of gene function.

Construction of cDNA clones

- Use the enzyme *reverse transcriptase* to copy mRNA into complementary DNA, called cDNA. This is equivalent to the template strand of the duplex DNA.
- Use a DNA polymerase to copy that cDNA.
- Insert the duplex cDNA product into a cloning vector and propagate in a host, e.g. *E. coli.*

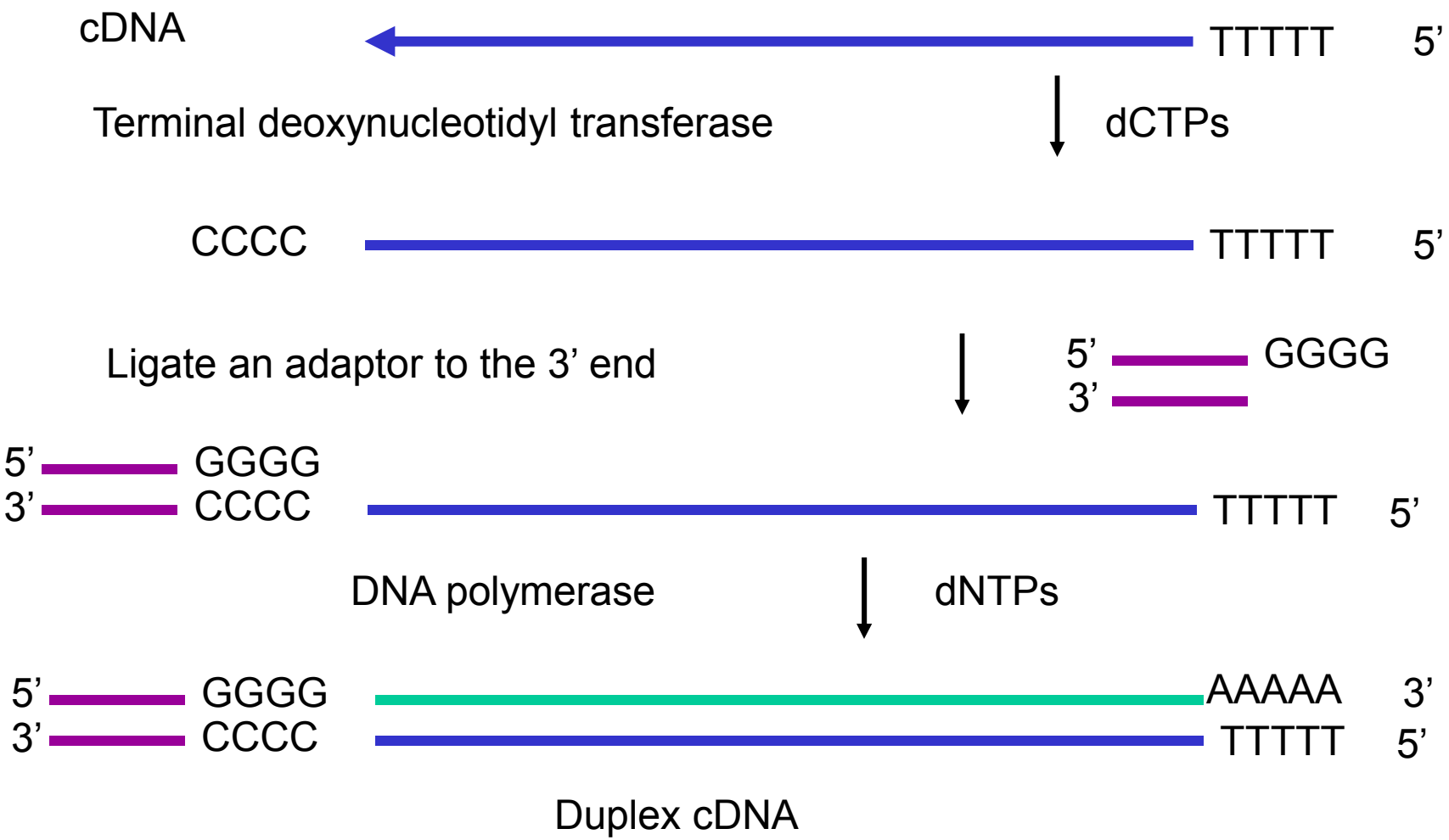
cDNA: first strand synthesis



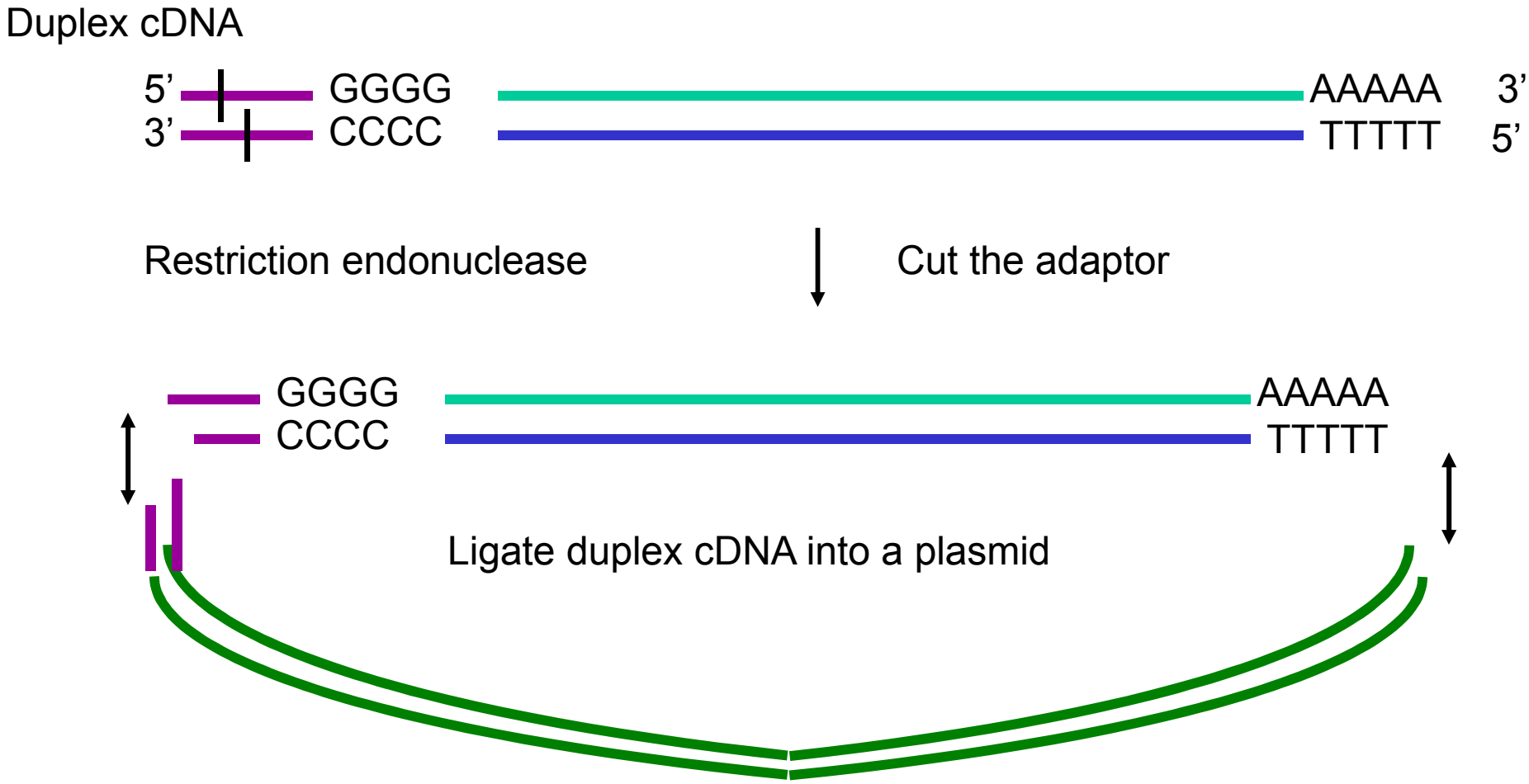
Product is complementary DNA, called **cDNA**. It is equivalent to the template strand of the duplex DNA.

cDNA: second strand synthesis

Problem: How to get a primer for 2nd strand synthesis?



Ligate duplex cDNA into a plasmid



Transform the population of cDNA plasmids into bacteria.
Result is a cDNA library.

cDNA Libraries

cDNA is complementary DNA prepared by reverse-transcribing cellular RNA.

- **Cloned eukaryotic cDNAs are useful because they lack introns** and other non-coding sequences present in the corresponding genomic DNA.
  - Introns are rare in bacteria but occur in most genes of higher eukaryotes and are transcribed by RNA Polymerase as primary transcripts that goes through a series of processing events in the nucleus before appearing in the cytoplasm as mature mRNA in a process called splicing
- **cDNA clones can be expressed in E.coli** since removal of eukaryotic introns by splicing does not occur in bacteria
- To **locate intron/exone boundaries** in eukaryotes

Steps for construction of cDNA library

The synthesis of double-stranded cDNA suitable for insertion into a cloning vector involves three major steps:

- (i**) first-strand DNA synthesis on the mRNA template by reverse transcriptase**
- (ii) **removal of the RNA template**
- (iii**) second strand DNA synthesis using the first DNA strand as a template by DNA-dependent DNA polymerase I.**

All DNA polymerases, whether they use RNA or DNA as the template, require a primer to initiate strand synthesis.

Screening gene libraries (Genomic and cDNA)

- **By Nucleic acid hybridization**

  most commonly used method of library screening because it is rapid, it can be applied to very large numbers of clones and, in the case of cDNA libraries, can be used to identify clones that are not full-length (and therefore cannot be expressed).

- **screening by PCR**

  Is possible if there is sufficient information about the sequence to make suitable primers

- **Screening expression libraries**

  If a DNA library is established **using expression vectors,** each individual **clone can be expressed to yield a polypeptide**.

Nucleic acid hybridization was first developed by **Grunstein and Hogness in 1975 and modified by Hanahan & Meselson 1980**.

- The colonies to be screened are **first replica-plated** on to a nitrocellulose filter disc that has been placed on the surface of an agar plate prior to inoculation
- A reference set of these **colonies on the master plate is retained.**
- The filter bearing the colonies is removed and **treated with alkali** so that the bacterial colonies are lysed and the DNA they contain is denatured.

The **filter is then treated with proteinase K** to remove protein and leave denatured DNA bound to the nitrocellulose, for which it has a high affinity, in the form of a 'DNA print' of the colonies.

The DNA is **fixed** firmly by baking the filter at 80°C. The defining, labelled RNA is hybridized to this DNA and the result of this hybridization is monitored by autoradiography.

A colony whose DNA print gives a positive **autoradiographic** result can then be picked from the reference plate.

Detection of recombinant clones by colony hybridization.

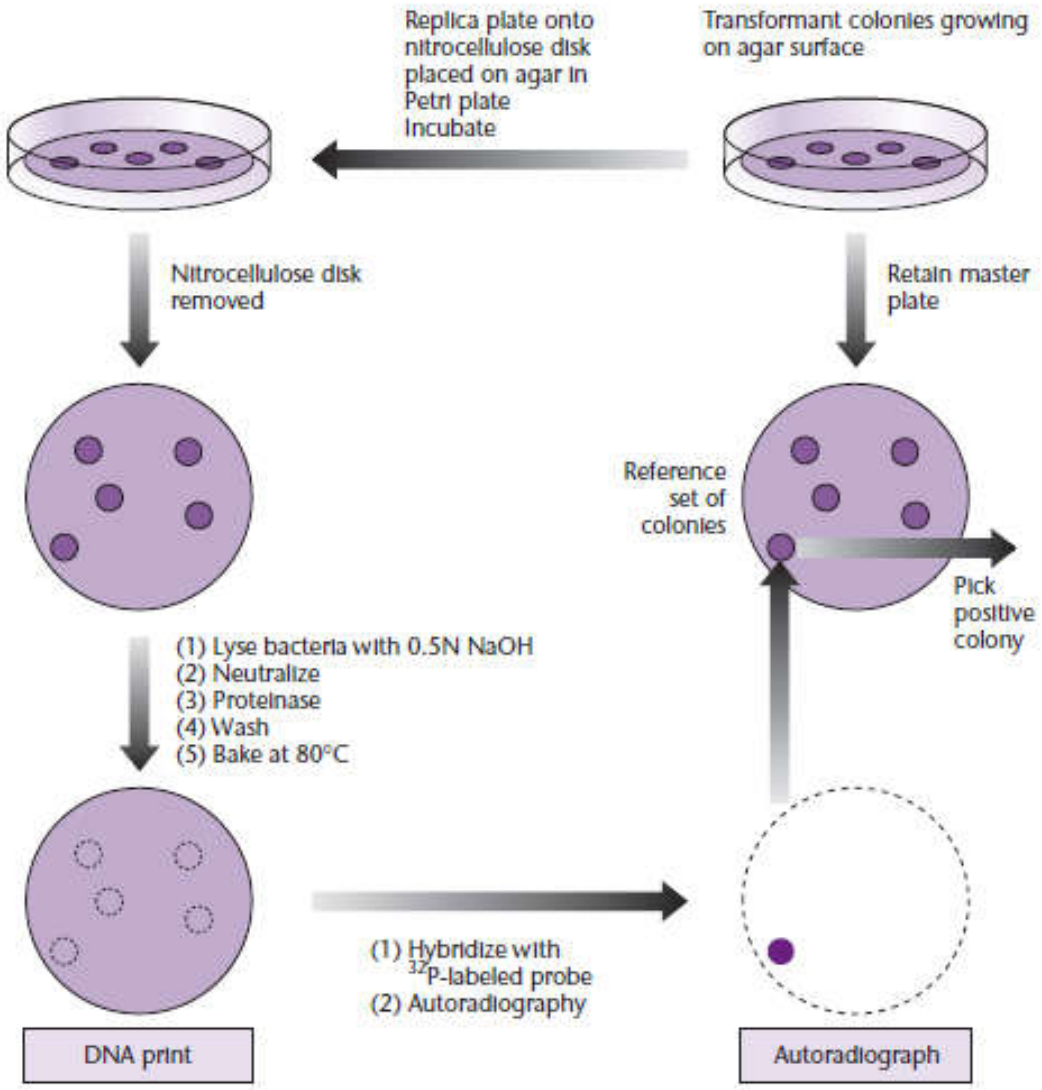


Fig. 6.12 Grunstein–Hogness method for detection of recombinant clones by colony hybridization.

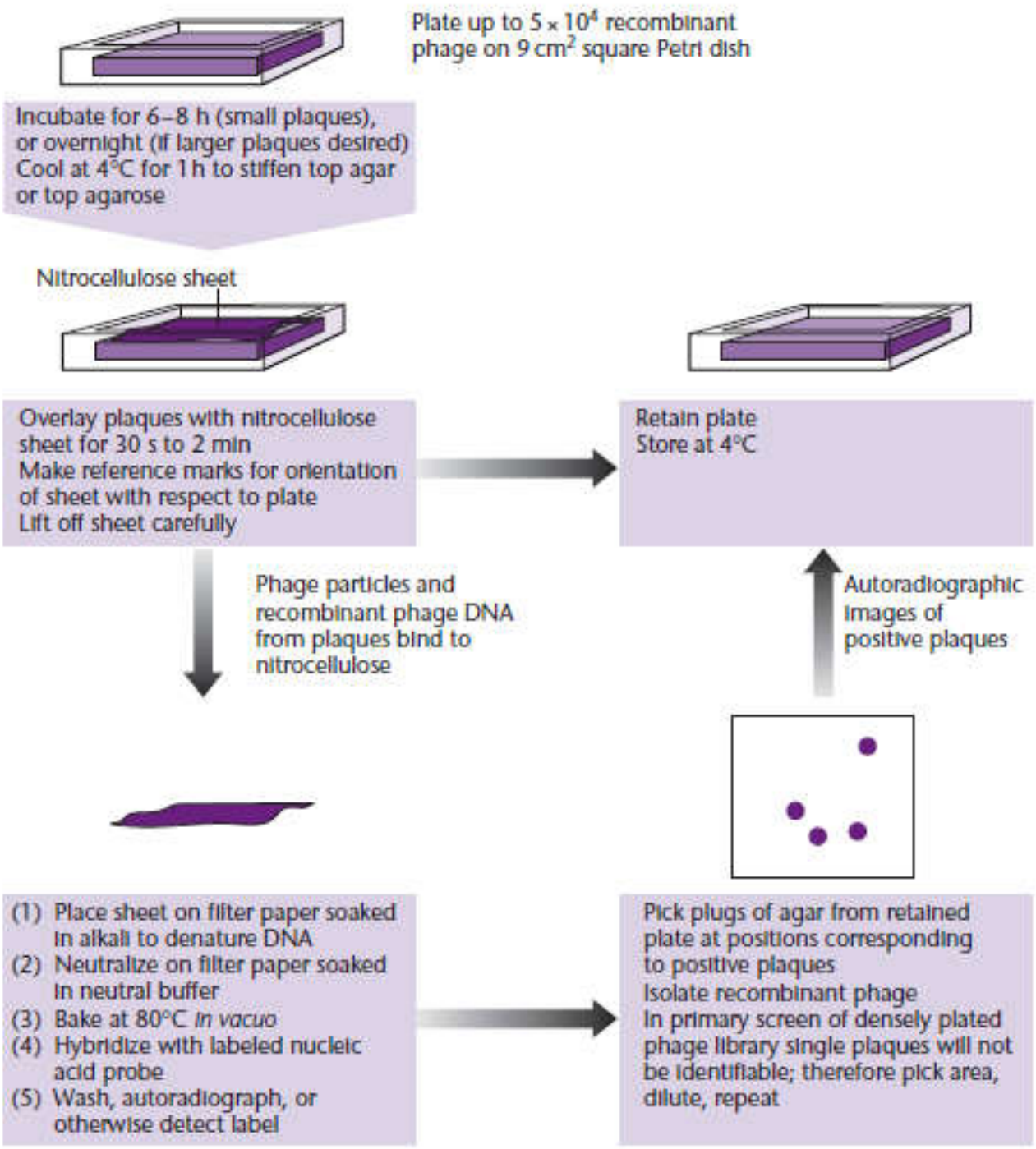


Fig. 6.13 Benton and Davis' plaque-lift procedure.

Screening by gene libraries by PCR

- library screening by PCR is demonstrated by **Takumi and Lodish (1994).**
- Instead of plating the library out on agar, as would be necessary for screening by hybridization, pools of clones are maintained in multiwell plates.

Each well is screened by PCR and positive wells are identified

The clones in each positive well are then diluted into a series in a secondary set of plates and screened again. The process is repeated until wells carrying homogeneous clones corresponding to the gene of interest have been identified.

Screening cDNA Libraries by immunological assay and protein activity

Immunological screening uses specific antibodies to detect expressed gene products

- If cDNA clones are inserted in the correct orientation and reading frame, cDNA sequences cloned in expression vectors can be expressed as *b-galactosidase* fusion proteins or other fusion proteins, and can be detected by immunological screening or screening with other ligands

- Immunological screening involves the use of antibodies that specifically recognize antigenic determinants on the polypeptide synthesized by a target clone.
- This is one of the most versatile expression cloning strategies because it can be applied to any protein for which an antibody is available.
- There is also no need for that protein to be functional. The molecular target for recognition is generally an *epitope, a short* sequence of amino acids that folds into a particular three-dimensional conformation on the surface of the protein.
- Epitopes can fold independently of the rest of the protein and therefore often form even when the polypeptide chain is incomplete, or when expressed as a fusion with another protein.
- Importantly, many epitopes can form under denaturing conditions when the overall conformation of the protein is abnormal.

- The first immunological screening techniques were developed in the late 1970s, when expression libraries were generally constructed using plasmid vectors.
- The method of **Broome & Gilbert (1978)** was widely used at the time.
- This method exploited the facts that antibodies adsorb very strongly to certain types of plastic, such as **polyvinyl**, and that **IgG antibodies can be readily labeled with 125I** by iodination *in vitro.*
- *As usual, transformed cells were plated* out on Petri dishes and allowed to form colonies. In order to release the antigen from positive clones, the colonies were lysed, e.g. using chloroform vapor or by spraying with an aerosol of virulent phage (a replica plate is required because this procedure kills the bacteria).

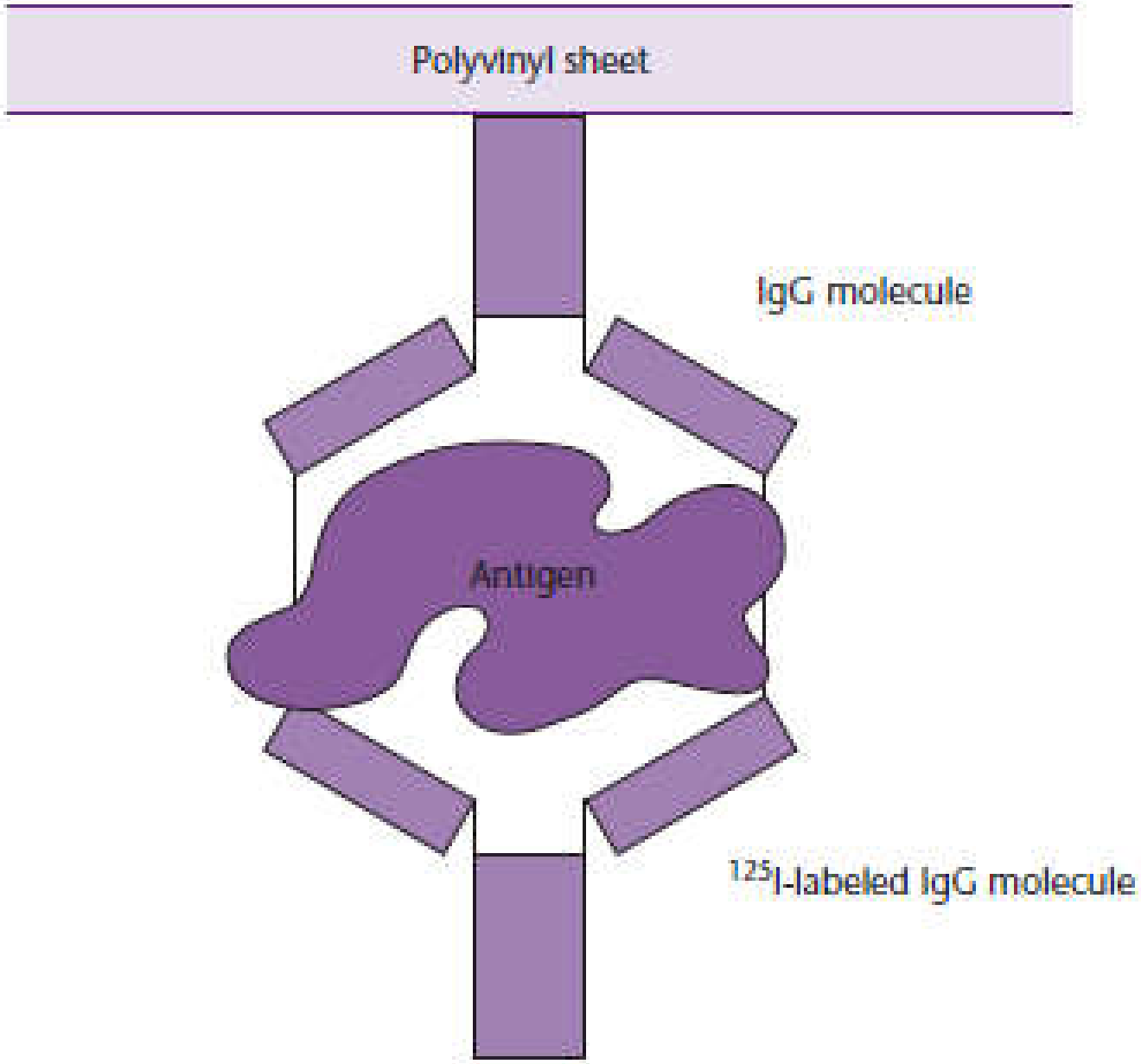


Fig. 6.14 Antigen–antibody complex formation in the immunochemical detection method of Broome and Gilbert.

- The original detection method using iodinated antibodies has been superseded by more convenient methods using non-isotopic labels, which are also more sensitive and have a lower background of nonspecific signal.
- Generally, these involve the use of unlabeled primary antibodies directed against the polypeptide of interest, which are in turn recognized by secondary antibodies carrying an enzymatic label.
- As well as eliminating the need for isotopes, such methods also incorporate an amplification step, since two or more secondary antibodies bind to the primary antibody.
- Typically, the secondary antibody recognizes the species-specific constant region of the primary antibody, and is conjugated to either horseradish peroxidase (de Wet *et al. 1984) or alkaline phosphatase* (Mierendorf *et al. 1987), each of which can* in turn be detected using a simple colorimetric assay

carried out directly on the nitrocellulose filter.

- Polyclonal antibodies, which recognize many different epitopes, provide a very sensitive probe for immunological screening, although they may also crossreact to proteins in the expression host.
- Monoclonal antibodies and cloned antibody fragments can also be used, although the sensitivity of such reagents is reduced because only a single epitope is recognized.