Introduction

Can, A. (2005). Introduction. En *A. Can, Principles of Virology Molecular, 4th Edition.* (pp. 11-24). New York: Elsevier Academic Press.

e-uaem
Espacio de Formación Multimodal

Figure 1.2 It is difficult to overestimate the importance of serological techniques in virology. The four assays illustrated by the diagrams in this figure have been used for many years and are of widespread value. (a) The *complement fixation test* works on the basis that complement is sequestered by antigen–antibody complexes. 'Sensitized' antibody-coated red blood cells, known amounts of complement, a virus antigen, and the serum to be tested are added to the wells of a multiwell plate. In the absence of antibodies to the virus antigen, free complement is present which causes lysis of the sensitized red blood cells (haemolysis). If, however, the test serum contains a sufficiently high titre of antivirus antibodies, then no free complement remains and haemolysis does not occur. Titrating the test serum by means of serial dilutions allows a quantitative measurement of the amount of antivirus antibody present to be made. (b) *Immunofluorescence* is performed using derivatized antibodies containing a covalently linked fluorescent molecule that emits a characteristically coloured light when illuminated by light of a different wavelength, such as rhodamine (red) or fluorescein (green). In direct immunofluorescence, the antivirus antibody itself is conjugated to the fluorescent marker, whereas in indirect immunofluorescence a second antibody reactive to the antivirus antibody carries the marker. Immunofluorescence can be used not only to identify virus-infected cells in populations of cells or in tissue sections but also to determine the subcellular localization of particular virus proteins (e.g., in the nucleus or in the cytoplasm). (c) *Enzyme-linked immunosorbent assays* (ELISAs) are a rapid and sensitive means of identifying or quantifying small amounts of virus antigens or antivirus antibodies. Either an antigen (in the case of an ELISA to detect antibodies) or antibody (in the case of an antigen ELISA) is bound to the surface of a multiwell plate. An antibody specific for the test antigen, which has been conjugated with an enzyme molecule (such as alkaline phosphatase or horseradish peroxidase), is then added. As with immunofluorescence, ELISA assays may rely on direct or indirect detection of the test antigen. During a short incubation, a colourless substrate for the enzyme is converted to a coloured product, thus amplifying the signal produced by a very small amount of antigen. The intensity of the product can easily be measured in a specialized spectrophotometer ('plate reader'). ELISA assays can be mechanized and are therefore suitable for routine tests on large numbers of clinical samples. (d) *Western blotting* is used to analyse a specific virus protein from a complex mixture of antigens. Virus antigen-containing preparations (particles, infected cells, or clinical materials) are subjected to electrophoresis on a polyacrylamide gel. Proteins from the gel are then transferred to a nitrocellulose or nylon membrane and immobilized in their relative positions from the gel. Specific antigens are detected by allowing the membrane to react with antibodies directed against the antigen of interest. By using samples containing proteins of known sizes in known amounts, the apparent molecular weight and relative amounts of antigen in the test samples can be determined.

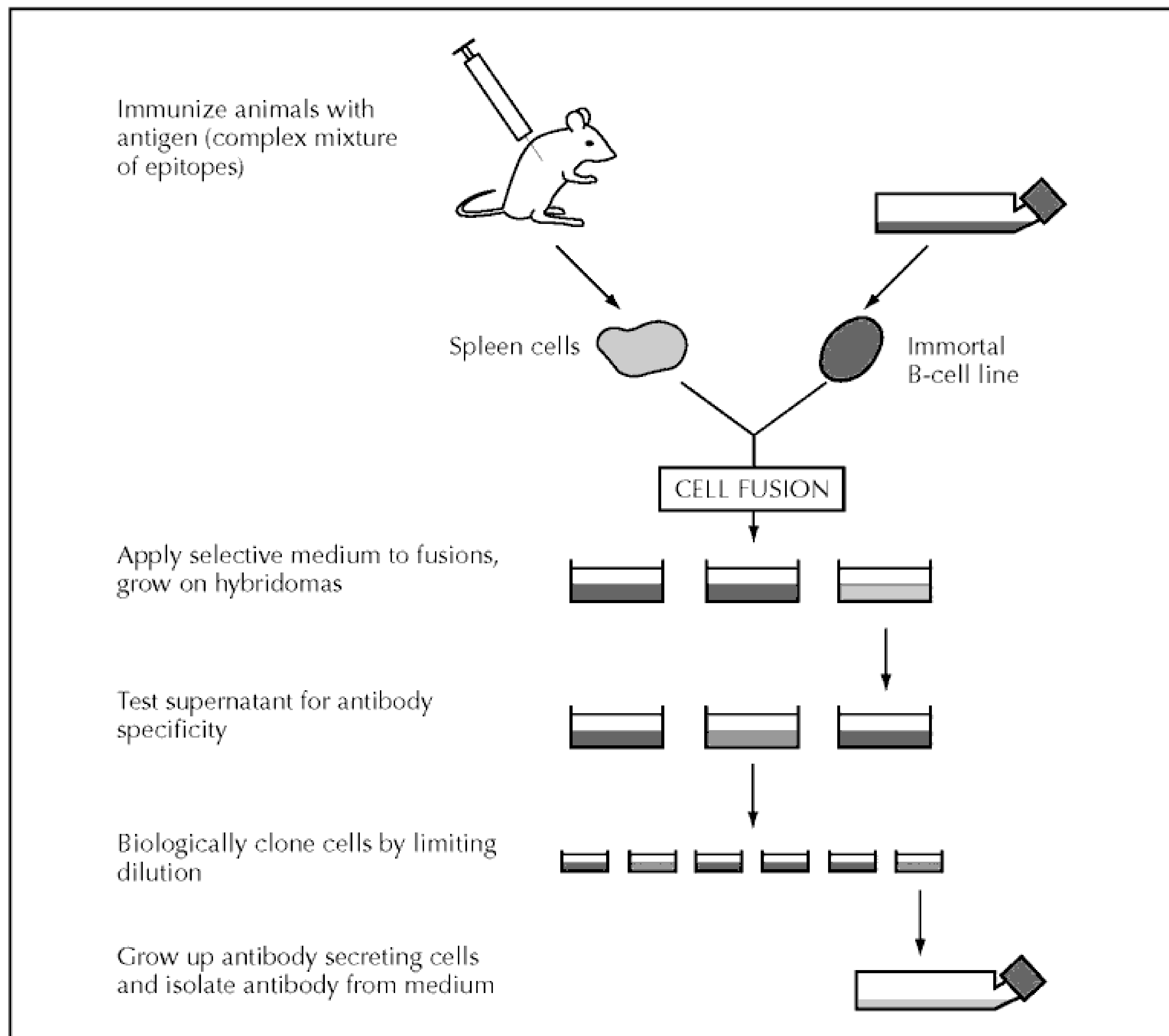


Figure 1.3 Monoclonal antibodies are produced by immunization of an animal with an antigen that usually contains a complex mixture of epitopes. Immature B-cells are later prepared from the spleen of the animal, and these are fused with a myeloma cell line, resulting in the formation of transformed cells continuously secreting antibodies. A small proportion of these will make a single type of antibody (a monoclonal antibody) against the desired epitope. Recently, *in vitro* molecular techniques have been developed to speed up the selection of monoclonal antibodies, although these have not yet replaced the original approach shown here.

Ultrastructural Studies

Ultrastructural studies can be considered under three areas: physical methods, chemical methods, and electron microscopy. Physical measurements of virus particles began in the 1930s with the earliest determinations of their proportions by filtration through colloidal membranes of various pore sizes. Experiments of this

sort led to the first (rather inaccurate) estimates of the size of virus particles. The accuracy of these estimates was improved greatly by studies of the sedimentation properties of viruses in ultracentrifuges in the 1960s (Figure 1.4). Differential centrifugation proved to be of great use in obtaining purified and highly concentrated preparations of many different viruses, free of contamination from host cell components, that can be subjected to chemical analysis. The relative density of virus

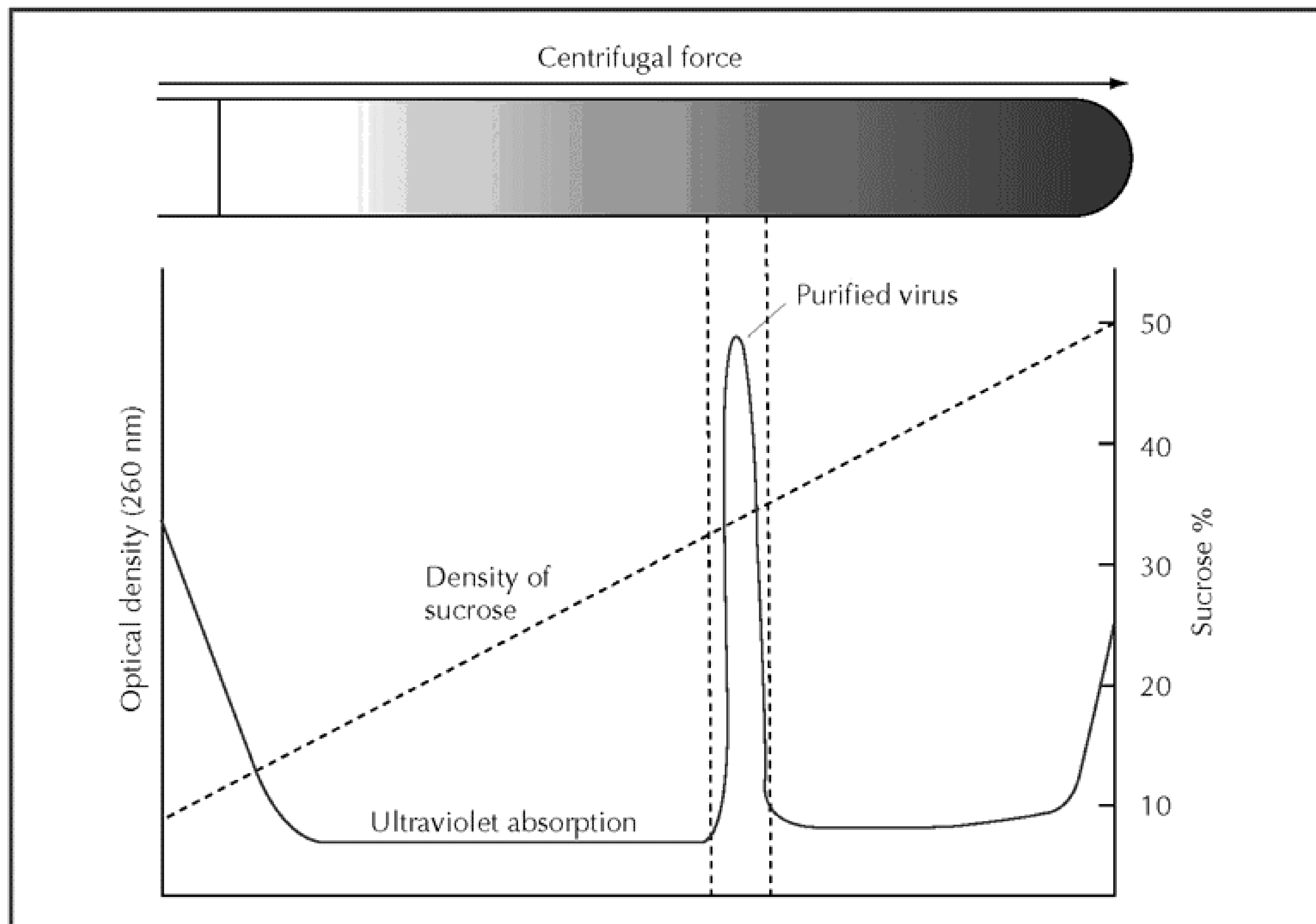
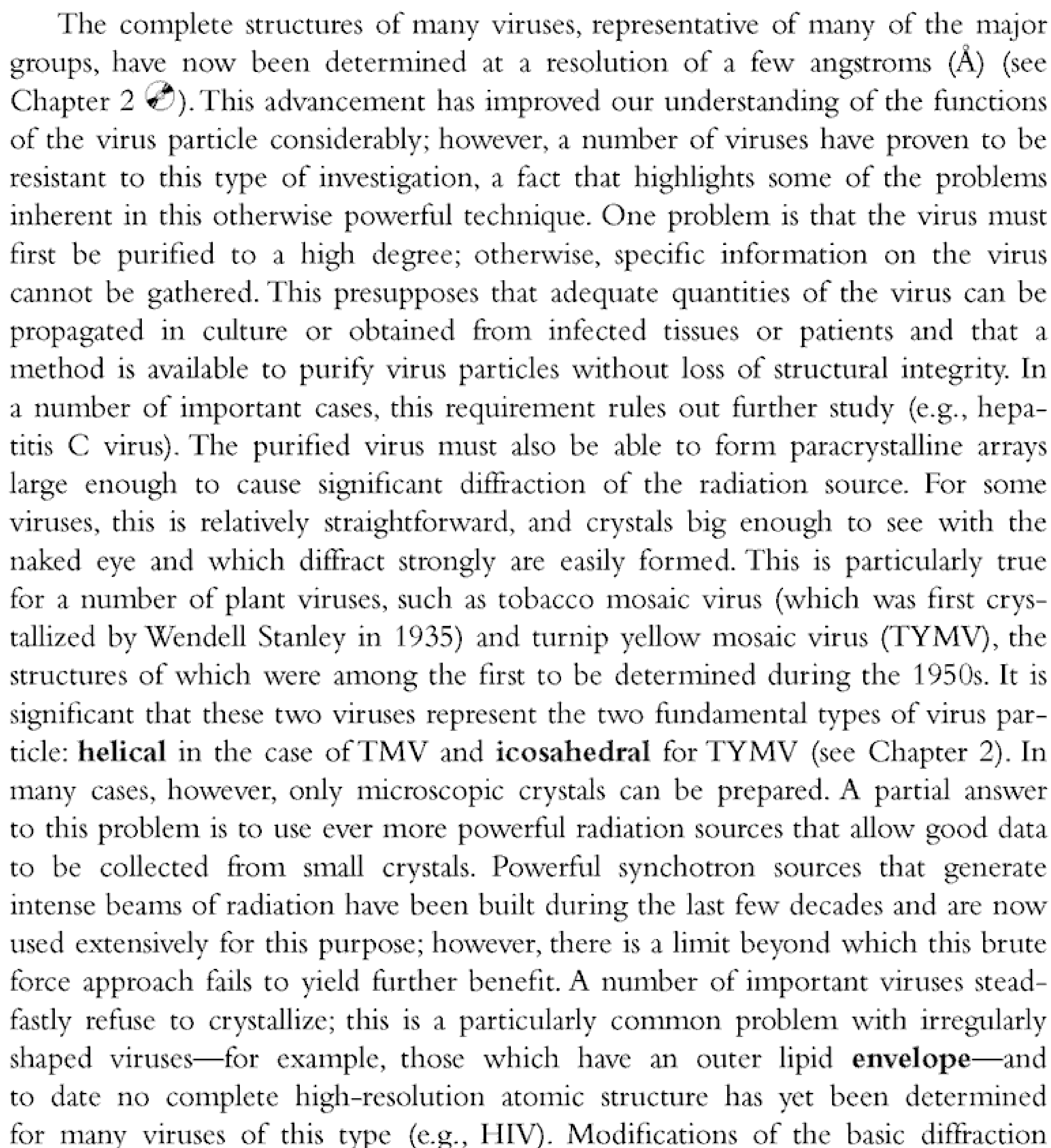


Figure 1.4 A number of different sedimentation techniques can be used to study viruses. In rate-zonal centrifugation (shown here), virus particles are applied to the top of a preformed density gradient, i.e., a sucrose or salt solution of increasing density from the top to the bottom of the tube (top of figure). After a period of time in an ultracentrifuge, the gradient is separated into a number of fractions, which are analysed for the presence of virus particles. In the figure, the nucleic acid of the virus genome is detected by its absorption of ultraviolet light (below). This method can be used both to purify virus particles or nucleic acids or to determine their sedimentation characteristics. In equilibrium or isopycnic centrifugation, the sample is present in a homologous mixture containing a dense salt such as caesium chloride. A density gradient forms in the tube during centrifugation, and the sample forms a band at a position in the tube equivalent to its own density. This method can thus be used to determine the density of virus particles and is commonly used to purify plasmid DNA.

particles, measured in solutions of sucrose or CsCl, is also a characteristic feature, revealing information about the proportions of nucleic acid and protein in the particles.

The physical properties of viruses can be determined by spectroscopy, using either ultraviolet light to examine the nucleic acid content of the particle or visible light to determine its light-scattering properties. Electrophoresis of intact virus particles has yielded some limited information, but electrophoretic analysis of individual **virion** proteins by gel electrophoresis, and particularly also of nucleic acid **genomes** (Chapter 3), has been far more valuable. However, by far the most important method for the elucidation of virus structures has been the use of x-ray diffraction by crystalline forms of purified virus. This technique permits determination of the structure of virions at an atomic level.



The complete structures of many viruses, representative of many of the major groups, have now been determined at a resolution of a few angstroms (Å) (see Chapter 2). This advancement has improved our understanding of the functions of the virus particle considerably; however, a number of viruses have proven to be resistant to this type of investigation, a fact that highlights some of the problems inherent in this otherwise powerful technique. One problem is that the virus must first be purified to a high degree; otherwise, specific information on the virus cannot be gathered. This presupposes that adequate quantities of the virus can be propagated in culture or obtained from infected tissues or patients and that a method is available to purify virus particles without loss of structural integrity. In a number of important cases, this requirement rules out further study (e.g., hepatitis C virus). The purified virus must also be able to form paracrystalline arrays large enough to cause significant diffraction of the radiation source. For some viruses, this is relatively straightforward, and crystals big enough to see with the naked eye and which diffract strongly are easily formed. This is particularly true for a number of plant viruses, such as tobacco mosaic virus (which was first crystallized by Wendell Stanley in 1935) and turnip yellow mosaic virus (TYMV), the structures of which were among the first to be determined during the 1950s. It is significant that these two viruses represent the two fundamental types of virus particle: **helical** in the case of TMV and **icosahedral** for TYMV (see Chapter 2). In many cases, however, only microscopic crystals can be prepared. A partial answer to this problem is to use ever more powerful radiation sources that allow good data to be collected from small crystals. Powerful synchotron sources that generate intense beams of radiation have been built during the last few decades and are now used extensively for this purpose; however, there is a limit beyond which this brute force approach fails to yield further benefit. A number of important viruses steadfastly refuse to crystallize; this is a particularly common problem with irregularly shaped viruses—for example, those which have an outer lipid **envelope**—and to date no complete high-resolution atomic structure has yet been determined for many viruses of this type (e.g., HIV). Modifications of the basic diffraction

technique (such as electron scattering by membrane-associated protein arrays and cryo-electron microscopy) may help to provide more information in the future, but it is unlikely that these variations will solve this problem completely. One further limitation is that some of the largest virus particles, such as poxviruses, contain hundreds of different proteins and are at present too complex to be analysed using these techniques.

Nuclear magnetic resonance (NMR) is increasingly being used to determine the atomic structure of all kinds of molecules, including proteins and nucleic acids. The limitation of this method is that only relatively small molecules can be analysed before the signals obtained become so confusing that they are impossible to decipher with current technology. At present, the upper size limit for this technique restricts its use to molecules with a molecular weight of less than about 30,000 to 40,000, considerably less than even the smallest virus particles. Nevertheless, this method may well prove to be of value in the future, certainly for examining isolated virus proteins if not for intact **virions**.

Chemical investigation can be used to determine not only the overall composition of viruses and the nature of the nucleic acid that comprises the virus **genome** but also the construction of the particle and the way in which individual components relate to each other in the **capsid**. Many classic studies of virus structure have been based on the gradual, stepwise disruption of particles by slow alteration of pH or the gradual addition of protein-denaturing agents such as urea, phenol, or detergents. Under these conditions, valuable information can sometimes be obtained from relatively simple experiments. For example, as urea is gradually added to preparations of purified adenovirus particles, they break down in an ordered, stepwise fashion which releases subvirus protein assemblies, revealing the composition of the particles. In the case of TMV, similar studies of capsid organization have been performed by renaturation of the capsid protein under various conditions (Figure 1.5). In simple terms, the reagents used to denature virus capsids can indicate the basis of the stable interactions between its components. Proteins bound together by electrostatic interactions can be eluted by addition of ionic salts or alteration of pH; those bound by nonionic, hydrophobic interactions can be eluted by reagents such as urea; and proteins that interact with lipid components can be eluted by nonionic detergents or organic solvents.

In addition to revealing fundamental structure, progressive denaturation can also be used to observe alteration or loss of antigenic sites on the surface of particles, and in this way a picture of the physical state of the particle can be developed. Proteins exposed on the surface of viruses can be labelled with various compounds (e.g., iodine) to indicate which parts of the protein are exposed and which are protected inside the particle or by lipid membranes. Cross-linking reagents such as psoralens or newer synthetic reagents with side-arms of specific lengths are used to determine the spatial relationship of proteins and nucleic acids in intact viruses.

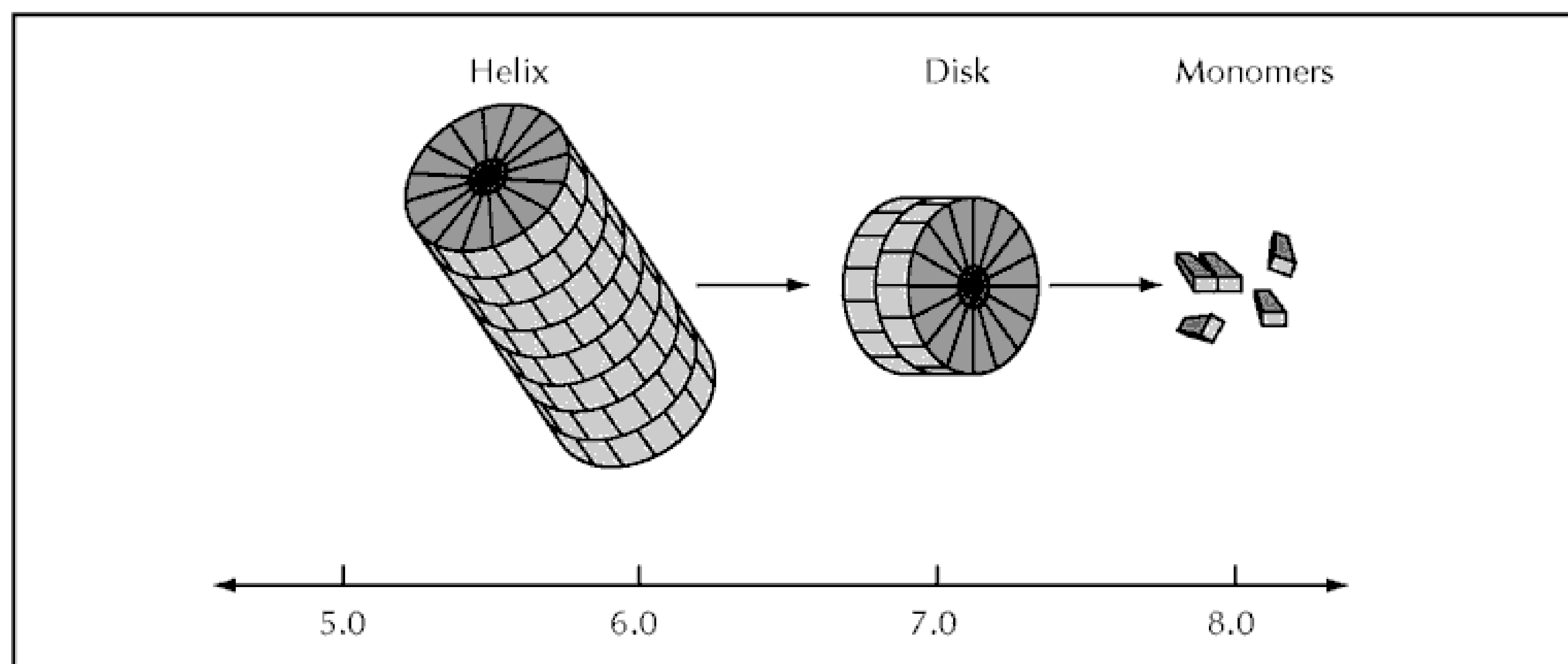


Figure 1.5 The structure and stability of virus particles can be examined by progressive denaturation or renaturation studies. At any particular ionic strength, the purified capsid protein of tobacco mosaic virus (TMV) spontaneously assembles into different structures, dependent on the pH of the solution. At a pH of around 6.0, the particles formed have a helical structure very similar to infectious virus particles. As the pH is increased to about 7.0, disk-like structures are formed. At even higher pH values, individual capsid monomers fail to assemble into more complex structures.

Since the 1930s, electron microscopes have overcome the fundamental limitation of light microscopes: the inability to resolve individual virus particles owing to physical constraints caused by the wavelength of visible light illumination and the optics of the instruments. The first electron micrograph of a virus (TMV) was published in 1939. Over subsequent years, techniques were developed that allowed the direct examination of viruses at magnifications of over 100,000 times. The two fundamental types of electron microscope are the transmission electron microscope (TEM) and the scanning electron microscope (SEM) (Figure 1.6). Although beautiful images with the appearance of three dimensions are produced by the SEM, for practical investigations of virus structure the higher magnifications achievable with the TEM have proved to be of most value. Two fundamental types of information can be obtained by electron microscopy of viruses: the absolute number of virus particles present in any preparation (total count) and the appearance and structure of the **virions** (see below). Electron microscopy can provide a rapid method of virus detection and diagnosis but in itself may give misleading information. Many cellular components (for example, ribosomes) can resemble 'virus-like particles,' particularly in crude preparations. This difficulty can be overcome by using antisera specific for particular virus antigens conjugated to electron-dense markers such as the iron-containing protein ferritin or colloidal gold suspensions. This highly

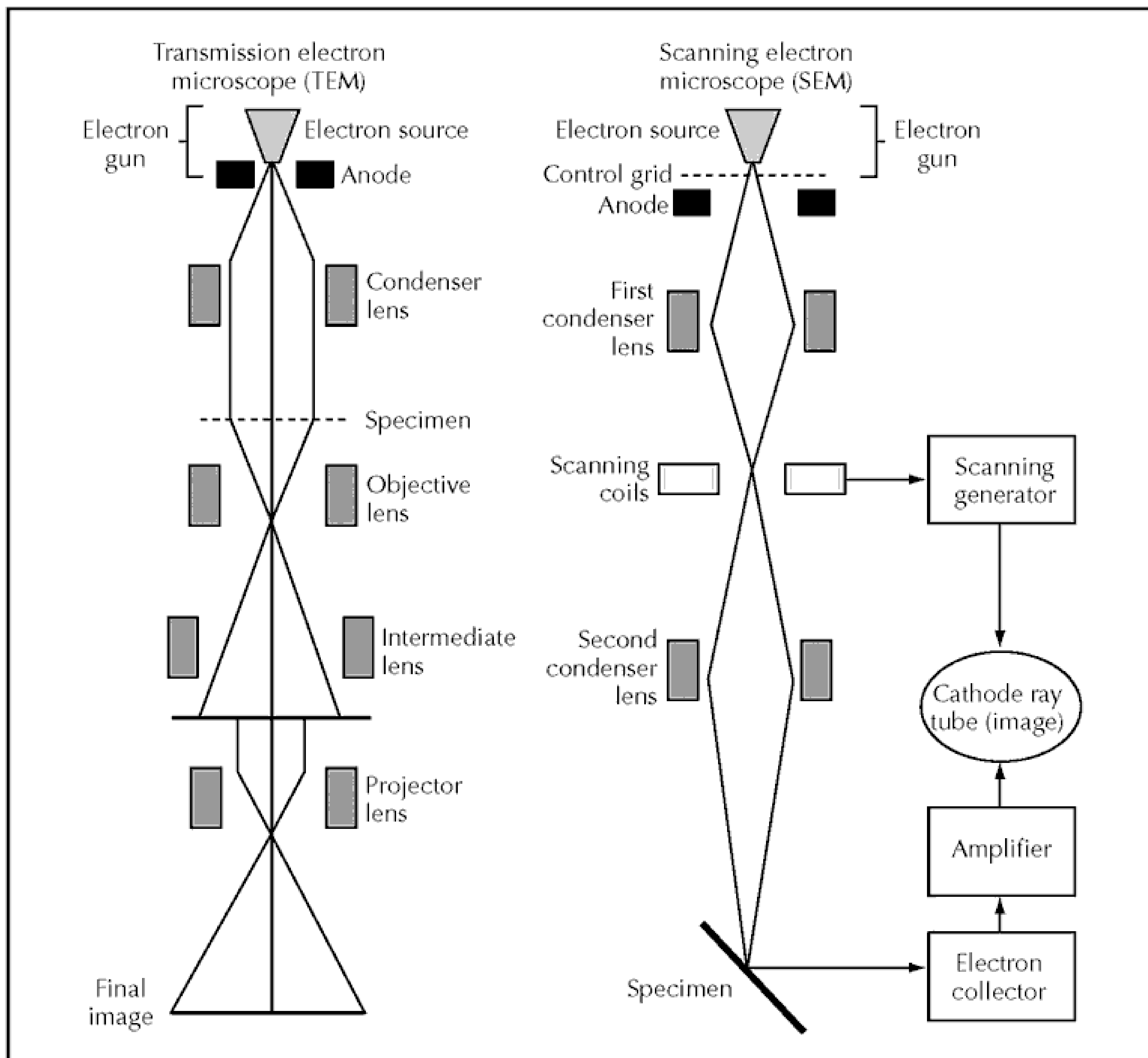


Figure 1.6 Working principles of transmission and scanning electron microscopes.

specific technique, known as immunoelectron microscopy, is gaining ground as a rapid method for diagnosis.

Developments in electron microscopy have allowed investigation of the structure of fragile viruses that cannot be determined by x-ray crystallography. These include cryo-electron microscopy, in which the virus particles are maintained at very low temperatures on cooled specimen stages; examination of particles embedded in vitreous ice, which does not disrupt the particles by the formation of ice crystals; low-irradiation electron microscopy, which reduces the destructive bombardment of the specimen with electrons; and sophisticated image-analysis and image-reconstruction techniques that permit accurate, three-dimensional images to be formed from multiple images that individually would appear as very poor quality. Conventional electron microscopy can resolve structures down to 50 to 70 Å in size (a typical atomic diameter is 2–3 Å; a protein α-helix, 10 Å; a DNA

double helix, 20 Å). Using these newer techniques it is possible to resolve structures of 25 to 30 Å.

In the late 1950s, Sydney Brenner and Robert Horne (among others) developed sophisticated techniques that enabled them to use electron microscopy to reveal many of the fine details of the structure of virus particles. One of the most valuable techniques proved to be the use of electron-dense dyes such as phosphotungstic acid or uranyl acetate to examine virus particles by negative staining. The small metal ions in such dyes are able to penetrate the minute crevices between the protein subunits in a virus **capsid** to reveal the fine structure of the particle. Using such data, Francis Crick and James Watson (1956) were the first to suggest that virus capsids are composed of numerous identical protein subunits arranged either in helical or cubic (**icosahedral**) symmetry. In 1962, Donald Caspar and Aaron Klug extended these observations and elucidated the fundamental principles of symmetry, which allow repeated protomers to form virus capsids, based on the principle of **quasi-equivalence** (see Chapter 2). This combined theoretical and practical approach has resulted in our current understanding of the structure of virus particles.

'Molecular Biology'

All of the above techniques of investigation are themselves 'molecular biology' in the original sense of the term; however, the term 'molecular biology' has taken on the new and different meaning of 'genetic engineering' or 'genetic manipulation.' These techniques for manipulating nucleic acids *in vitro* (that is, outside living cells or organisms) do not comprise a new discipline but are an outgrowth of earlier developments in biochemistry and cell biology over the previous 50 years. This powerful new technology has revolutionized virology and, to a large extent, has shifted the focus of attention away from the virus particle onto the virus **genome**. Again, this book is not the place to discuss in detail the technical aspects of these methods, and readers are referred to one of the many relevant texts, such as those given at the end of this chapter.

Virus infection has long been used to probe the working of 'normal' (i.e., uninfected) cells—for example, to look at macromolecular synthesis. This is true, for example, of the applications of **bacteriophages** in bacterial genetics and in many instances where the study of eukaryotic viruses has revealed fundamental information about the cell biology and genomic organization of higher organisms. In 1970, John Kates first observed that vaccinia virus mRNAs were polyadenylated at their 3′ ends. In the same year, Howard Temin and David Baltimore jointly identified the enzyme reverse transcriptase (RNA-dependent DNA polymerase) in retrovirus-infected cells. This finding shattered the so-called 'central dogma' of biology that there is a one-way flow of information from DNA through RNA into protein and

revealed the plasticity of the **eukaryote** genome. Subsequently, the purification of this enzyme from retrovirus particles permitted cDNA cloning, which greatly accelerated the study of viruses with RNA genomes—a good illustration of the catalytic nature of scientific advances. In 1977, Richard Roberts and, independently, Phillip Sharp recognized that adenovirus mRNAs were spliced to remove intervening sequences, indicating the similarities between virus and cellular genomes.

Initially at least, the effect of this new technology was to shift the emphasis of investigation from proteins to nucleic acids. As the power of the techniques developed, it quickly became possible to determine the nucleotide sequences of entire virus **genomes**, beginning with the smallest **bacteriophages** in the mid-1970s and working up to the largest of all virus genomes, those of the herpesviruses and poxviruses, many of which have now been determined.

This nucleic acid-centred technology, in addition to its ultimate achievement of nucleotide sequencing and the artificial manipulation of virus genomes, also offered significant advances in detection of viruses and virus infections involving nucleic acid hybridization techniques. There are many variants of this basic idea, but, essentially, a hybridization probe, labelled in some fashion to facilitate detection, is allowed to react with a crude mixture of nucleic acids. The specific interaction of the probe sequence with complementary virus-encoded sequences, to which it binds by hydrogen-bond formation between the complementary base pairs, reveals the presence of the virus genetic material (Figure 1.7). This approach has been taken a stage further by the development of various *in vitro* nucleic acid amplification procedures, such as polymerase chain reaction (PCR), which is an even more sensitive technique, capable of detecting just a single molecule of virus nucleic acid (Figure 1.8).

More recently, there has also been renewed interest in virus proteins based on a new biology which is itself dependent on manipulation of nucleic acids *in vitro* and advances in protein detection arising from immunology. Methods for *in vitro* synthesis and expression of proteins from molecularly cloned DNA have advanced rapidly, and many new analytical techniques are now available. Studies of protein–nucleic acid interactions are proving to be particularly valuable in understanding virus structure and gene expression. Advances in electrophoresis have made it possible to study simultaneously all of the proteins in a virus-infected cell, called the **proteome** of the cell (by analogy to the **genome**).

Molecular biologists have one further trick up their sleeves. Because of the repetitive, digitized nature of nucleotide sequences, computers are the ideal means of storing and processing this mass of information. 'Bioinformatics' is a broad term coined in the 1980s to encompass any application of computers to biology. This can imply anything from artificial intelligence and robotics to genome analysis. More specifically, the term applies to computer manipulation of biological sequence data, including protein structural analysis. Bioinformatics permits the inference of

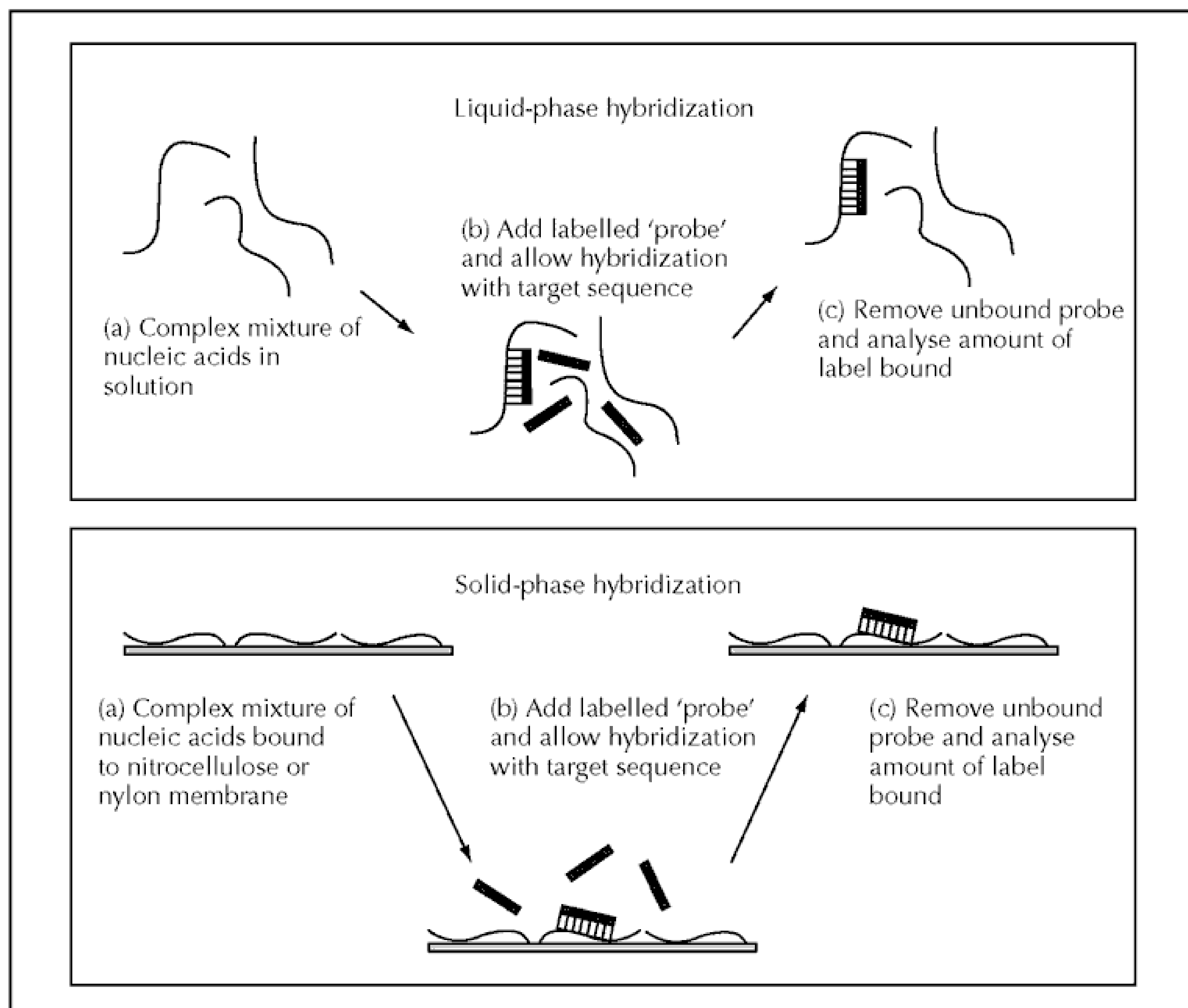


Figure 1.7 Nucleic acid hybridization relies on the specificity of base-pairing which allows a labelled nucleic acid probe to pick out a complementary target sequence from a complex mixture of sequences in the test sample. The label used to identify the probe may be a radioisotope or a nonisotopic label such as an enzyme or chemiluminescent system. Hybridization may be performed with both the probe and test sequences in the liquid phase (top of figure) or with the test sequences bound to a solid phase, usually a nitrocellulose or nylon membrane (below). Both methods may be used to quantify the amount of the test sequence present, but solid-phase hybridization is also used to locate the position of sequences immobilized on the membrane. Plaque and colony hybridization are used to locate recombinant molecules directly from a mixture of bacterial colonies or **bacteriophage** plaques on an agar plate. Northern and Southern blotting are used to detect RNA and DNA, respectively, after transfer of these molecules from gels following separation by electrophoresis (cf., western blotting, Figure 1.2).

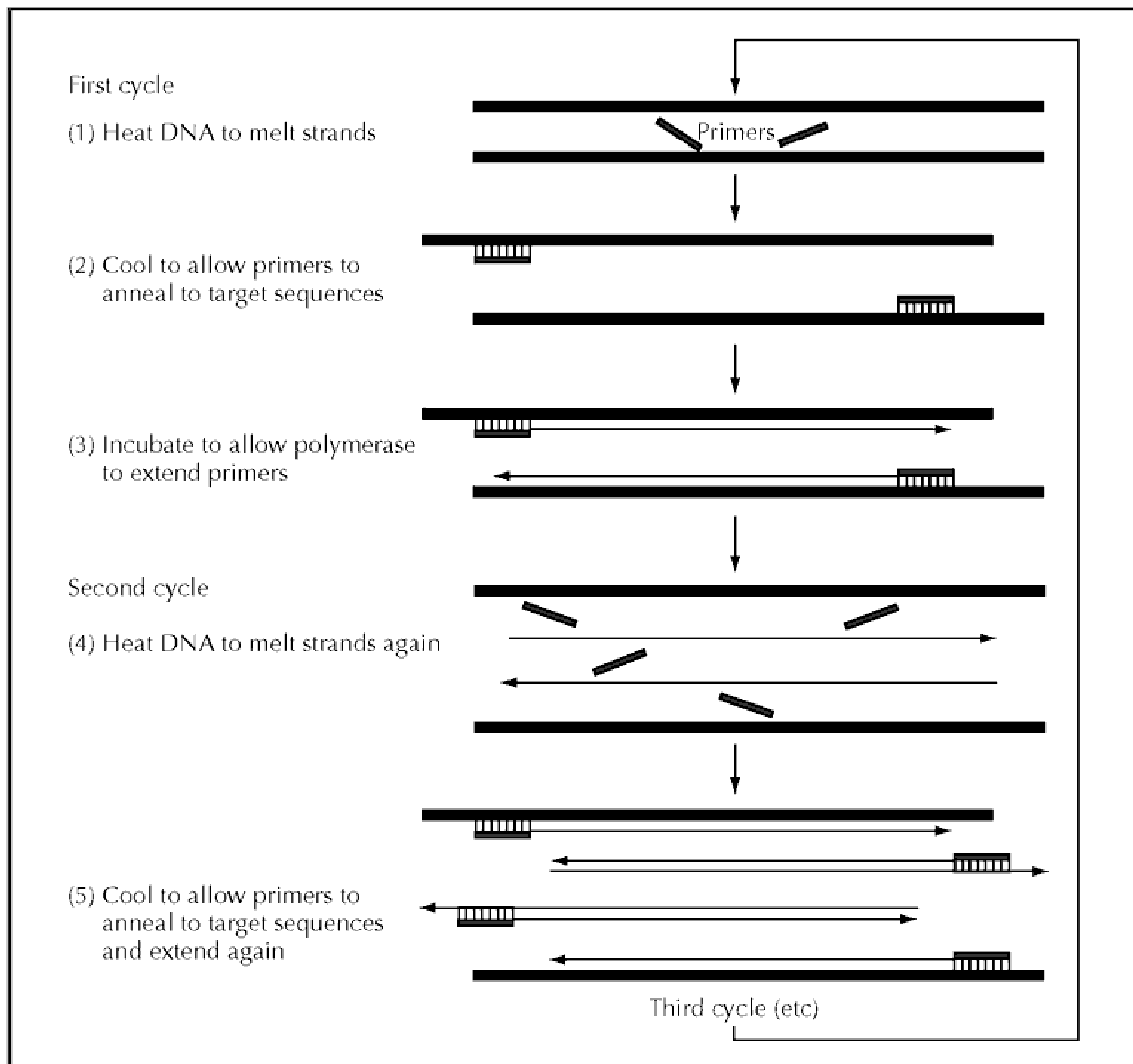


Figure 1.8 Polymerase chain reaction (PCR) relies on the specificity of base-pairing between short synthetic olignucleotide probes and complementary sequences in a complex mixture of nucleic acids to prime DNA synthesis using a thermostable DNA polymerase. Multiple cycles of primer annealing, extension, and thermal denaturation are carried out in an automated process, resulting in a massive amplification (2*n*-fold increase after *n* cycles of amplification) of the target sequence located between the two primers.

function from the linear sequence and is thus central to all areas of modern biology. Due to the flood of new sequence information, computers are being used increasingly to make predictions based on nucleotide sequences (Figure 1.9). These include detecting the presence of open reading frames, the amino acid sequences of the proteins encoded by them, control regions of genes such as **promoters** and splice signals, and the secondary structure of proteins and nucleic acids. However (particularly in the case of RNA), the secondary structure assumed by molecules is almost as important as the primary nucleotide sequence in determining the biological reactions that the molecule may undergo. Caution is needed in interpreting such predicted rather than factual information, and the validity of such predictions should not be accepted without question unless confirmed by biochemical and/or genetic data. However, when the structure of a protein has been determined by x-ray crystallography or NMR, the shape can be accurately modelled and explored in three dimensions on computers (Figure 1.10).

While the **genome** is the nucleic acid comprising the entire genetic information of an organism, by extension 'genomics' is the study of the composition and function of the genetic material of an organism. Virus genomics began with the first complete sequence of a virus genome (bacteriophage ϕX174 in 1977). Vast international databases of nucleotide and protein sequence information have now been compiled, and these can be rapidly accessed by computers to compare newly determined sequences with those whose function may have been studied in great

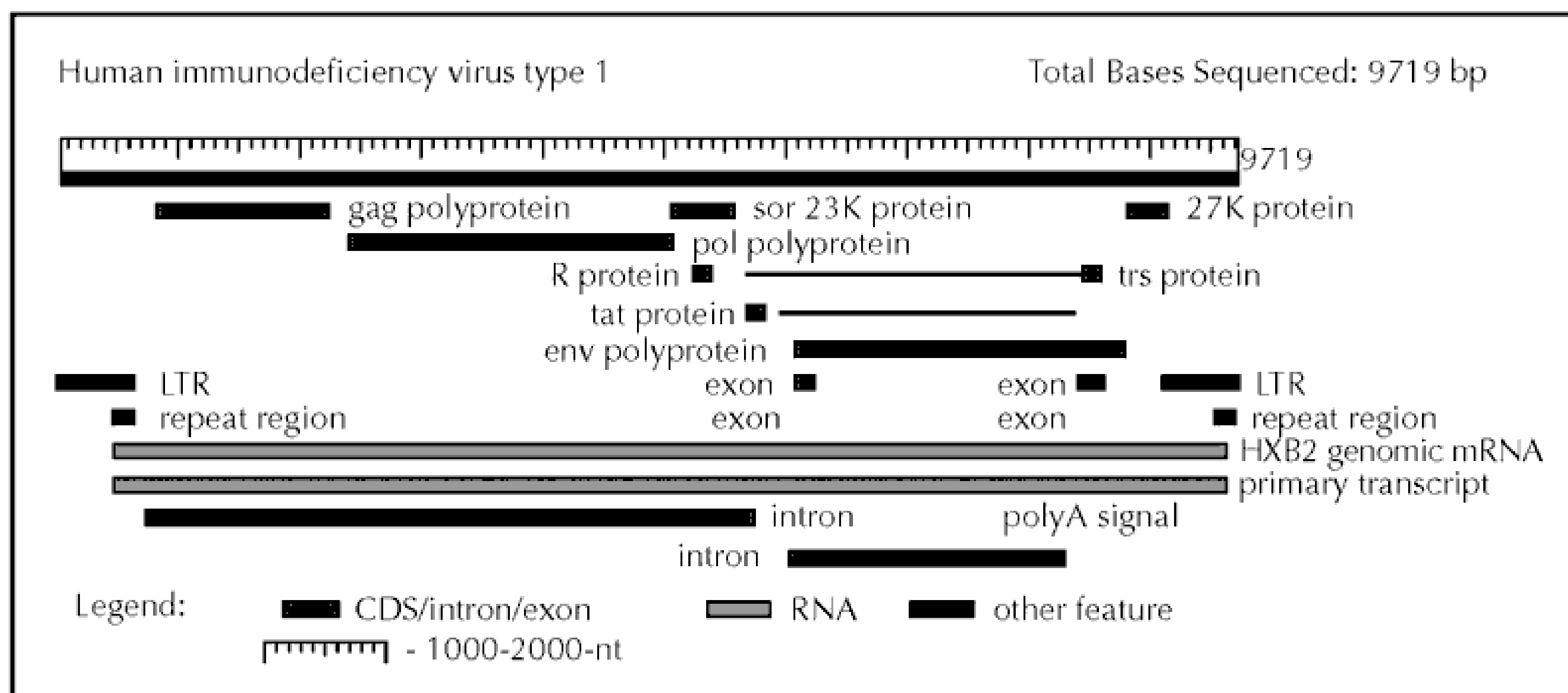


Figure 1.9 An example of the use of a computer to store and process digitized information from a nucleic acid sequence. This figure shows an analysis of all of the open reading frames (**ORFs**) present in an HIV-1 **provirus**. The ORFs present in the three main retrovirus genes, *gag*, *pol*, and *env*, can be seen. This complex analysis took only a few seconds to perform using an ordinary personal computer. Manually, the same task may have taken several days.

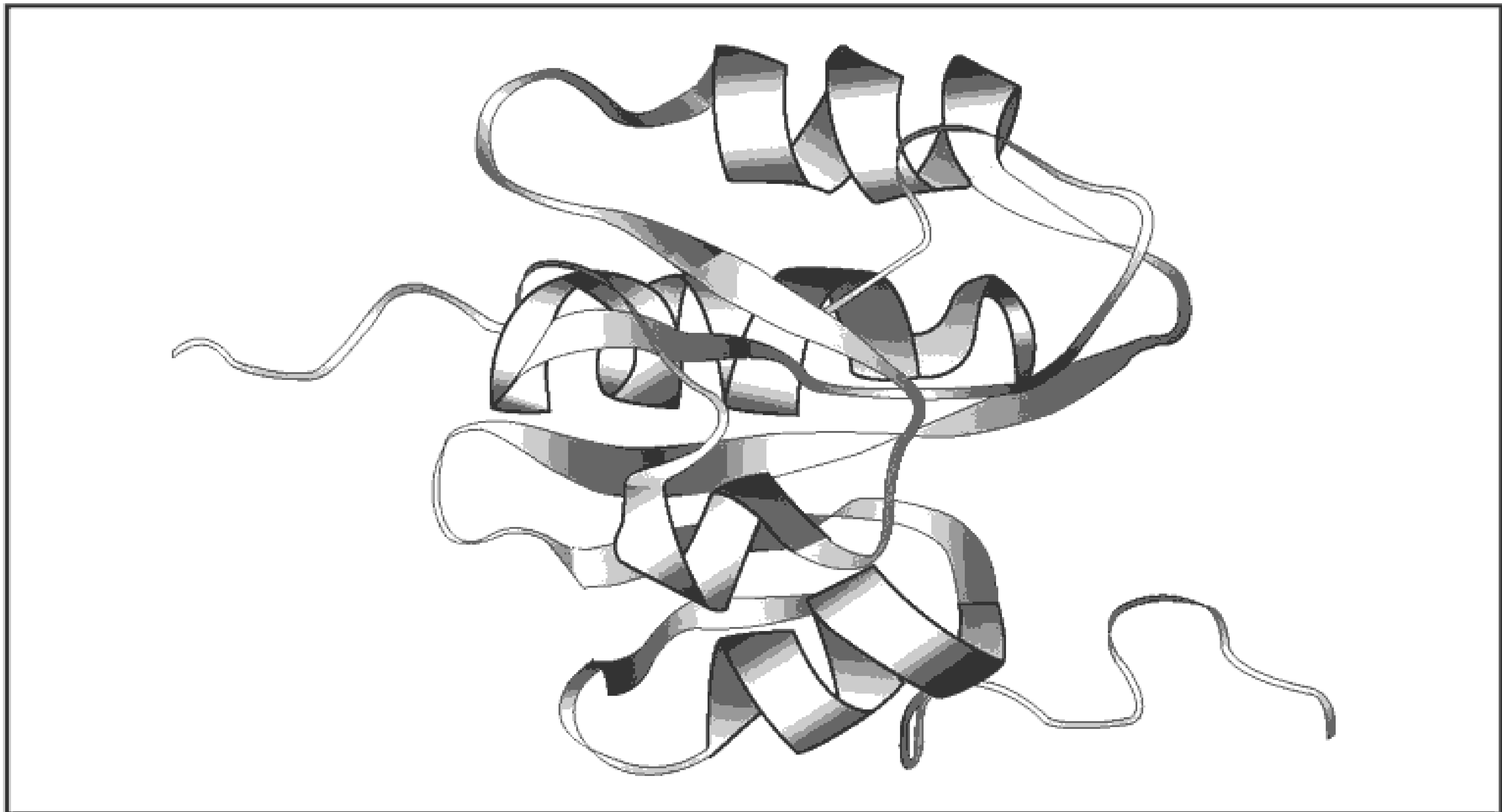
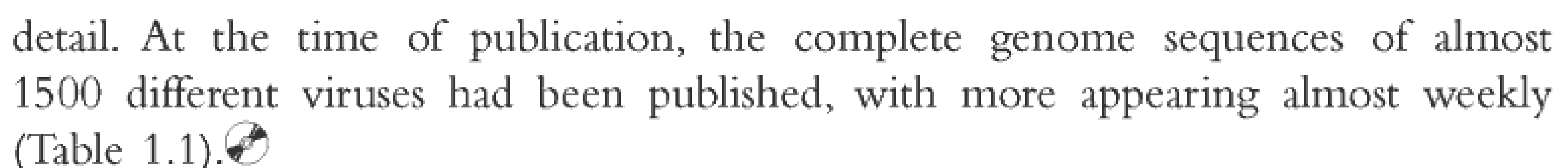


Figure 1.10 Three-dimensional structure of the DNA binding domain of SV40 T-antigen reconstructed from NMR data using a computer.

Table 1.1 Genomic comparison of different organisms

Organism	Number of genes	Percent (%) of genes with known or inferred function
Hepatitis B virus	4	75
SV40	6	100
Herpes simplex virus	80	95
Mimivirus	900	10
Escherichia coli	4,288	60
Yeast	6,600	40
Caenorhabditis elegans	19,000	40
Drosophila	14,000	25
Arabidopsis	25,000	40
Mouse	100,000	10
Human	100,000	10



detail. At the time of publication, the complete genome sequences of almost 1500 different viruses had been published, with more appearing almost weekly (Table 1.1).

Thus we have, in a sense, come full circle in our investigations of viruses—from particles via **genomes** back to proteins again—and have emerged with a far more profound understanding of these organisms; however, the current pace of research in virology tells us that there is still far more that we need to know.

Further Reading

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2003). *Essential Cell Biology*. Garland Science, New York. ISBN 0815334818.

Cann, A.J. (1999). *Virus Culture: A Practical Approach*. Oxford University Press, Oxford. ISBN 0199637148.

Hendrix, R.W. (2003). Bacteriophage genomics. *Current Opinion in Microbiology*, 6: 506–511.

Kuby, J., Goldsby, R., Kindt, T.J., and Osborne, B. (2003). *Immunology*. W.H. Freeman, New York. ISBN 0716749475.

Lesk, A.M. (2002). *Introduction to Bioinformatics*. Oxford University Press, Oxford. ISBN 0199251967.

Lio, P. and Goldman, N. (2004). Phylogenomics and bioinformatics of SARS-CoV. *Trends in Microbiology*, 12: 106–111.

Primrose, S.B., Twyman, R.M., and Old, R.W. (2001). *Principles of Gene Manipulation*. Blackwell Scientific, London. ISBN 0632059540.

Rohwer, F. and Edwards, R. (2002). The Phage Proteomic Tree: a genome-based taxonomy for phage. *Journal of Bacteriology*, 184: 4529–4535.