Part II

Determination of the Structure and Physical Properties of Viruses

Chapter 3
Conventional Electron Microscopy, Cryo-Electron Microscopy and Cryo-Electron Tomography of Viruses

José R. Castón

Abstract Electron microscopy (EM) techniques have been crucial for understanding the structure of biological specimens such as cells, tissues and macromolecular assemblies. Viruses and related viral assemblies are ideal targets for structural studies that help to define essential biological functions. Whereas conventional EM methods use chemical fixation, dehydration, and staining of the specimens, cryo-electron microscopy (cryo-EM) preserves the native hydrated state. Combined with image processing and three-dimensional reconstruction techniques, cryo-EM provides 3D maps of these macromolecular complexes from projection images, at subnanometer to near-atomic resolutions. Cryo-EM is also a major technique in structural biology for dynamic studies of functional complexes, which are often unstable, flexible, scarce or transient in their native environments. As a tool, cryo-EM complements high-resolution techniques such as X-ray diffraction and NMR spectroscopy; these synergistic hybrid approaches provide important new information. Three-dimensional cryo-electron tomography goes further, and allows the study of viruses not only in their physiological state, but also in their natural environment in the cell, thereby bridging structural studies at the molecular and cellular levels.

Keywords Capsid • Cryo-electron microscopy • Cryo-electron tomography • Electron microscopy • Fourier transform • Image processing • Resolution • Three-dimensional reconstruction • Viral macromolecular assembly • Virus

---

J.R. Castón (✉)
Department of Macromolecular Structure, Centro Nacional de Biotecnología (CSIC),
c/Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain
e-mail: jrcaston@cnb.csic.es

M.G. Mateu (ed.), *Structure and Physics of Viruses: An Integrated Textbook*,
Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_3,
© Springer Science+Business Media Dordrecht 2013

Abbreviations

2D	Two-dimensional, two dimensions
3D	Three-dimensional, three dimensions
3DR	Three-dimensional reconstruction
CCD	Charge-coupled device
cryo-EM	Cryo-electron microscopy
cryo-ET	Cryo-electron tomography
CTF	Contrast transfer function
EM	Electron microscopy, electron microscope
FEG	Field emission gun
FSC	Fourier shell correlation
FT	Fourier transform
SEM	Scanning electron microscopy
SSE	Secondary structure element
TEM	Transmission electron microscopy

3.1 Introduction

Microscopy is probably the technique with the greatest influence in the history of biology. Optical (or light) microscopy was introduced to biology in the seventeenth century by Antoni van Leeuwenhoek, who described bacteria and protozoa using a microscope able to enlarge image size approximately 300 times. Light microscopy uses visible light and a system of lenses to magnify images, providing intuitive information on morphology and general structure of the biological model under study. The end of the nineteenth and beginning of the twentieth centuries saw the discovery of electrons and X rays; both provided new alternatives for the observation of matter and stimulated the development of the electron microscope and the X-ray diffractometer. In 1931, Ernst Ruska and colleagues invented the electron microscope (Ruska obtained the Nobel Prize in Physics in 1986). One of the early micrographs taken in the 1930s was of bacterial viruses or bacteriophages, called “tiny hostile bacteria”.

It took 300 years to improve the simple light microscope, but less than 40 years to perfect the electron microscope. With electron microscopy (EM), scientists were provided with a powerful method that extended the range of observable structures far beyond the limits imposed by the physics of visible light. Our understanding of cell structure and tissues is based in numerous EM contributions. As in many other fields in biology, viruses have had a decisive role in the development of EM, as they constitute simplified tools with which to standardize preparation methods in biological applications. In turn, EM has made outstanding contributions to understand virus structure and its relationship with biological function.

Biological EM might be considered a descriptive technique in its origins. The questions addressed by EM have evolved together with our concept of the molecular reality of the cell [1]. The experience accumulated over many decades in the multidisciplinary field of biology has shown that each molecular event is more than the sum of its parts. This is reflected, for example, in the regulation and integration of physiological signals, currently termed “systems biology”. Genomics has provided a complete list of the macromolecules in a cell, which led to structural genomics, which in turn provides the basic principles of the structure of proteins, of DNA and of RNA. The basic cellular entities that carry out fundamental biological processes are nonetheless multiprotein assemblies [2]. These macromolecular complexes (some of which may contain nucleic acids) respond to factors that determine their association, disassembly or signal transmission, leading finally to specific molecular events.

To decipher biological processes, the organization of macromolecules and components of macromolecular complexes must be established. X-ray crystallography (Chap. 4) and nuclear magnetic resonance (NMR) spectroscopy (Chap. 5) have provided the atomic structures of the modules or subunits of these molecular machines. Given the difficulties in producing sufficient amounts of these complexes and in their crystallization, the number of structures of large macromolecular complexes that have been determined at near-atomic resolution (by X-ray crystallography) is limited compared to the number of protein structures available. EM acts as a link between structural cell biology and high-resolution structural molecular biology. Over time, descriptive EM has given way to a quantitative approach [3]. Three-dimensional (3D) density maps can be synthesized from two-dimensional (2D) images, as electron micrographs are 2D projections of the 3D (multi-) macromolecular assembly under study [4]. EM images nonetheless have deficiencies, as the electron microscope is not a perfect instrument. These problems can be reduced by subsequent image processing of the micrographs. This progress is reflected in the relatively large number of biological structures that have been resolved thus far to resolutions in the 3.5–4.5-Å range.

EM facilitates the understanding of an atomic structure in its biological context (at lower resolution), particularly when macromolecular assemblies and the relationships of their components are involved. The frontiers between NMR spectroscopy, X-ray crystallography and EM are becoming diffuse, as they form a smooth continuum for visualizing macromolecular structures [5] (see also Chaps. 4, 5 and 7). The emerging relationship among these approaches is a complementary one that can overcome the limitations of each method alone; for instance, atomic models can be fitted accurately into EM-derived maps of larger assemblies [6] (see Chap. 7). Viruses are excellent models of nanomachines; their capsids and other viral protein assemblies have provided some of the most striking images using this hybrid approach [7]. In addition to structure and assembly, the form of EM termed cryo-electron microscopy (cryo-EM) also allows dynamic studies of complexes, which are often unstable, scarce or transient, in their native environment.

A structural description of a macromolecular assembly in its different functional states facilitates a mechanistic understanding of the corresponding processes [8].

In this chapter, we describe the principles of EM, with examples of its applications to the study of some viruses. Most sections in this chapter outline the basis of 3D cryo-EM (cryo-EM plus image processing), illustrating its potential in the marriage with X-ray crystallography (see also Chap. 7). The powerful technique of 3D cryo-electron tomography (cryo-ET), which allows the study of viruses in their natural environment in the cell, is also described. The examples mentioned may give the reader a perspective as to how this approach can be applied to the study of large viral complexes and allow determination of their overall shape, the location of subunits and domains within them, and identification of conformational changes that accompany intermediates in their functional cycle. Moreover, even today in the age of molecular diagnosis, EM remains a mainstay in detecting the causative organism in outbreaks of new and unusual diseases; EM does not require specific reagents to be able to recognize a pathogenic agent [9].

3.2 Transmission Electron Microscopy of Viruses

3.2.1 Basic Concepts and General Experimental Design

The physical basis of EM is the dual particle-wave nature of the electron and its charge. Electrons can be accelerated by electric fields reaching high velocities and accordingly much shorter associated wavelengths than visible light photons, leading to much improved resolution compared with an optical microscope. Electrons are provided by an emission source, such as a heated tungsten or lanthanum hexaboride filament or, alternatively, by a field emission gun (FEG). A FEG generates a much brighter electron beam with better spatial coherence than a thermal ionic electron source. Electrons are accelerated by high voltage, and focused by electromagnetic lenses (condenser lenses) to make a collimated (highly parallel) beam that interacts with the specimen. The interior of the microscope column must be under a high vacuum system to reduce the frequency of electron collision with gas atoms to negligible levels in the path of the electrons; this system also prevents degradation of the coherence and monochromaticity of the electron beam (Fig. 3.1). The sample is inserted in the column through an airlock.

There are two basic types of electron microscopes, the transmission electron microscope (TEM) and the scanning electron microscope (SEM). TEM projects electrons through a very thin specimen; the beam interacts with the specimen and the transmitted electrons are focused by electromagnetic lenses (objective lens and projector lenses) onto a detector to produce a magnified 2D image of the specimen [10]. The electron beam covers a wide area of the specimen. Detectors are a fluorescent screen, a photographic film or a charge-coupled device (CCD). In addition, direct electron detectors (without scintillators to convert the electrons

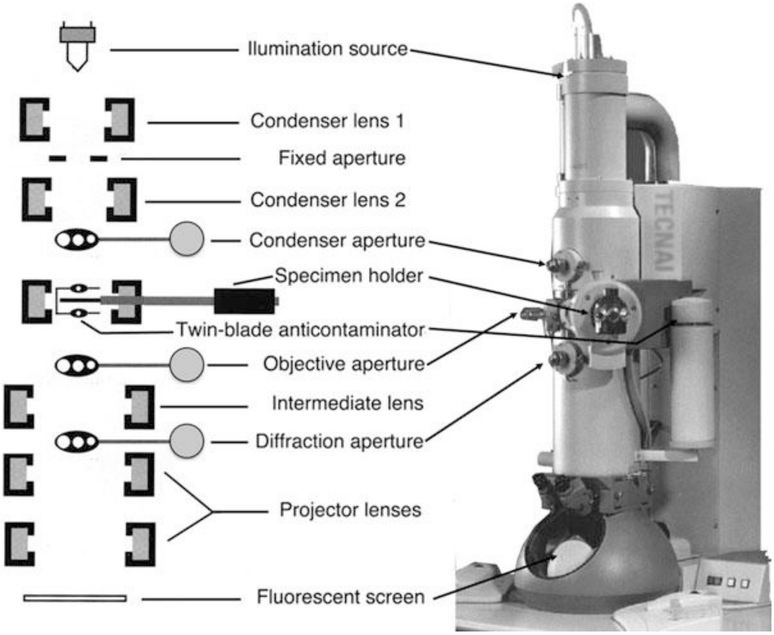


Fig. 3.1 Major components of a transmission electron microscope. Diagram of major components in the column of a modern transmission (cryo-)electron microscope

into photons) are being explored, as they would be superior recording devices for EM images. The brightness of a particular area of the image is proportional to the number of electrons that are transmitted through the specimen.

SEM produces an image that gives the impression of three dimensions. In this microscope, the electron beam is focused to a fine probe and scans the surface of the specimen, point by point, to generate secondary (reflected) electrons from the specimen; these are then detected by a sensor. The image is produced over time as the entire specimen is scanned. There are also scanning transmission electron microscopes (STEM) for analytical use; a scanning beam of electrons penetrates thin samples and provides quantitative determinations of macromolecular masses, and indicates the presence and distribution of the atomic elements in the specimen. The majority of macromolecular (and viral) EM studies are based on TEM.

EM specimens must be solid and thin. Although water is by far the foremost component of living entities, it has long been neglected in EM, as it evaporates in the vacuum of the microscope. Years of research have established many useful procedures for observing biological samples with acceptable structural preservation although their water has been removed (see below). Full preservation of native conformation after complete dehydration nonetheless remains unresolved. Cryo-EM techniques, which began to emerge in the 1980s, consider water as the natural intrinsic component that it actually is. Specimens remain fully hydrated, immobilized in vitreous water (or vitreous ice). Water is cooled very rapidly in such a way it is solidified without ordering, as it is in hexagonal or cubic ices; the resulting vitreous water is in an amorphous solid state. The specimen is maintained at very low temperature (−170 °C or lower) and the evaporation rate becomes negligible, preserving the hydrated state of the specimen in the microscope [11].

The amount of electron scattering is dependent on specimen density and overall thickness. The specimen must be thin enough to prevent image degradation due to the numerous elastic and inelastic scattering interactions of electrons passing through the specimen [12]. Electrons that interact with the potential field of atomic nuclei usually undergo elastic collisions and are deflected with no loss of energy. They emerge with their wavelength unchanged, but experience an associated phase shift. Elastically scattered and unscattered electrons contribute to image formation. Electrons that interact with the outer electrons of atomic nuclei undergo inelastic collisions and transfer some of their energy to the specimen atoms; this process contributes to image blurring, and is responsible for radiation damage to the specimen. Due to the nonconductive nature of the specimen, these inelastic interactions lead to the formation of many highly reactive ions and free radicals; this results in the breakage of covalent bonds, causing molecular structures to degrade rapidly during observation. Image blurring can be caused by specimen movement, which can include mechanical drift, vibration and/or image thermal drift. These adverse effects can be minimized using imaging strategies in which the electron dose is low [13].

A constant problem with biological specimens is their low contrast, defined as perceptible variation in the intensity between different regions. Two components contribute to contrast, amplitude and phase contrast. Amplitude contrast is due to the corpuscular nature of electrons, and requires use of an objective lens aperture that traps the electrons scattered at large angles (inelastic scattering or high angle elastic scattering). Biological samples scatter electrons weakly (at small angles), and artificial methods are used to enhance scattering properties, such as the introduction of heavy atoms (for example, negative staining). Regions of the specimen that give rise to these trapped electrons appear dark in the image. This mechanism is negligible in unstained specimens. In theory, smaller objective apertures increase the amplitude contrast, but are more prone to astigmatism problems if the aperture is displaced from the center of the beam axis.

Phase contrast is due to the wave-like nature of electrons, and arises from the constructive or destructive interference between the elastically scattered electrons and the unscattered, transmitted electrons. This is the dominant contrast mechanism in cryo-EM of vitrified (unstained) specimens. Elastically scattered electrons experience a phase shift (or difference in path length) as they emerge from the specimen, with no alterations in their amplitude, and this phase shift can be modified. Spherical aberration (inherent to the objective lens of the microscope) and focus setting (directly manipulated by the microscopist) are factors that generate phase contrast. No phase contrast would be seen in an aberration-free microscope when the specimen is imaged exactly in focus (see below).

3.2.2 Sample Preparation Techniques: Negative Staining, Metal Shadowing, Embedding and Ultramicrotomy

As important as the development of the electron microscope itself is the development of adequate sample preparation techniques to be used with it. In the late 1950s, detailed images of viruses were obtained by embedding them in a high density (electron-opaque), highly soluble stain, which replicates specimen structure. Heavy metal salts such as phosphotungstic acid, uranyl acetate, and ammonium molybdate have been used successfully as negative stains. Stains stabilize the sample in the harsh conditions necessary for observation, including the hostile atmosphere (high vacuum) and continuous electron irradiation (energy transfer processes). As a general rule, the difficulties associated with biological samples are solved by following four steps: (1) fixation (chemical or physical), (2) dehydration or drying, (3) the introduction of a support (plastic and amorphous carbon layers, resin, metal shadow or vitreous ice), and (4) increase in contrast (heavy, electron-dense elements).

For optimal results, a firm, structureless substrate is essential for specimen support; plastic films such as collodion or formvar stabilized with a thin layer of carbon are appropriate. These films are used as a support in a specimen grid, the electron microscope analog of the glass slide used in light microscopy. EM grids are fine mesh supports (~3 mm in diameter), on which a plastic/carbon film is deposited, and the biological material is adhered for transport and viewing in the electron microscope. Copper is normally used in the fabrication of grids, although other more inert metals such as gold or nickel may be needed.

There are several negative staining procedures; the flotation method is probably the most popular. A plastic/carbon-coated grid is floated on a droplet containing the specimen for 1–3 min to permit adsorption of the specimen. The grids are previously glow-discharged to render the carbon film more hydrophilic and increase specimen adhesion. The grid is then passed through two to three droplets of water, and transferred onto a drop of negative stain for 30–60 s, blotted with filter paper, and air-dried. In principle, the low pH of the negative stain solution, the adsorption of the specimen to the carbon layer, the embedding in a heavy-metal salt cast, as well as the dehydration method can all be expected to severely disrupt the native conformation of the specimen. The technique nonetheless preserves quite well the structure of wide range of specimens, and constitutes a indispensable initial step for higher resolution studies by cryo-EM [14, 15], as resolution is limited by the granularity of the stain to ~20 Å.

The negative staining technique is a very quick and easy method for assessing the homogeneity and quality of many particulate solutions such as purified virus preparations, and requires a minimum of experience and equipment. Limitations are due mainly to structural preservation; stain depicts only the surface features of the specimen and any cavities the stain penetrates, but little or no information is derived about the partially- or unstained internal features. The method is important in structural virology for fast but unrefined morphological identification. Enveloped viruses are very susceptible to disruption by negative stains and/or air-drying.

Metal shadowing techniques also use heavy metals to enhance the contrast of the low intrinsic electron scattering of biological samples. Shadowing experiments reveal topographic features of specimens by spraying particles of vaporized metal to form a thin coating on whole pre-dried specimens. Fixing and drying of the specimen should preserve the details to be studied; the simplest method is to allow the sample to air-dry, although this technique can result in excessive shrinkage, cracking and/or collapse of fine structures. The freeze-drying method freezes the specimen rapidly while it is still in the aqueous phase. The frozen specimen is transferred to a chamber that maintains a temperature below −80 °C under vacuum, and the solid ice gradually sublimates; preservation is usually better using this process. The freeze-fracture technique replicates fractured surfaces of frozen specimens obtained with a sharp knife; it is especially suited for studies of membrane structure. A variant of this approach is freeze-etching, in which sublimation (etching of ice) is introduced after fracturing.

Platinum or tungsten/tantalum are usually melted and evaporated from a heated electrode, and a second coat of carbon is often used to improve stability of the replica coating. Specimen shadowing can be either unidirectional or rotary. In unidirectional shadowing, the specimen is immobile and is shadowed at an average angle of ~45° in a high vacuum evaporator. This method can be used to measure the height of a specimen from the length of its shadow, and these micrographs are analyzed to obtain topographic maps. In rotary shadowing, the specimen is rotated on a motorized stage and is shadowed from all directions. Metal replicas are released from the underlying biological material by careful chemical treatments such as a hypochlorite solution, and the floating replica is washed thoroughly, carefully fished on fine grids, dried, and then viewed in the microscope.

Embedding and ultramicrotomy of biological samples, including isolated virus particles or a virus-infected tissue or cells, is an elaborate procedure. Specimens are immersed in a matrix that is solidified, then cut into thin slices or sections (50–100 nm thick). The specimen is first chemically fixed to preserve its structure; osmium tetroxide was the first fixative used for EM, but laboratories today use many variations on the fixation process, including potassium permanganate and the standard glutaraldehyde-osmiun tetroxide protocol. The sample is next dehydrated, which consists of the gradual replacement of water with an organic solvent that acts as a "transition solvent" between the aqueous environment of the specimen and the hydrophobic plastic embedding medium; ethanol and acetone are common dehydrating agents. The sample is infiltrated in a resin; in this process, dehydrants are gradually replaced by (viscous) resin monomers such as epoxy mixtures (often called Epon). The resin specimens are transferred into molds or capsules, and placed in an oven in which the epoxy components polymerize to form a solid. There are many types of embedding media with special uses, such as Lowicryl and LR White (acrylic media) used for immunocytochemistry. Finally, the embedded sample is thinly sectioned with a precision (diamond or glass) knife on an

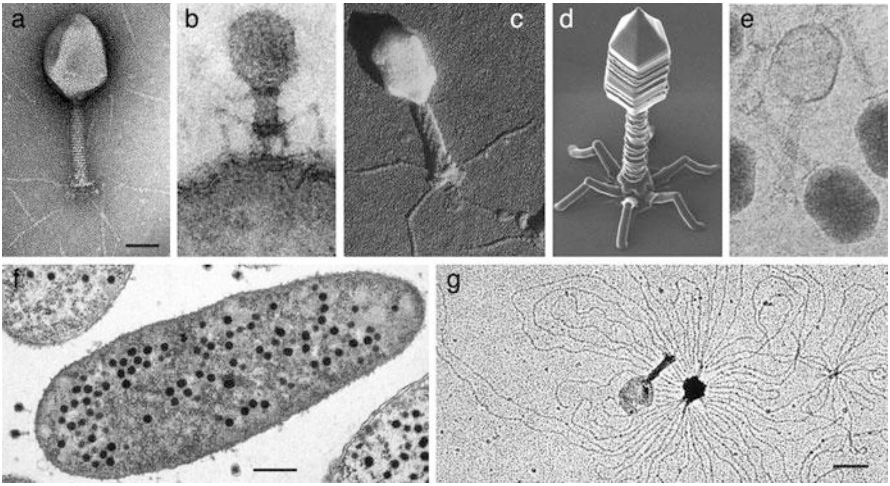


Fig. 3.2 Analysis of bacteriophage T4 by conventional EM and cryo-EM techniques. (**a**) T4 phage negatively stained with 2 % uranyl acetate; darker areas represent local concentrations of stain. The general morphology of the T4 phage, including head, tail, baseplate and fibers can easily be distinguished. The 40.5 Å tail spacing is seen as stacked disks. Note that phage heads are collapsed to different extents, giving the impression of structural heterogeneity due to phage collapse. Bar, 50 nm. (**b**) Thin section showing phage T4 injecting its DNA genome through the bacterial wall; fibers adhere to the outer wall surface and the tail is contracted (Dr. Michel Wurtz, University of Basel, Switzerland; micrograph by J. v.d. Broek). (**c**) Freeze-dried, metal-shadowed phage T4. General morphology is seen as with negative staining, and additional structural details can be inferred: head and tail shadows indicate different heights for these structures; the helical pattern in the tail region is evident in these conditions. (**d**) SEM image of an artificial nano T4 phage fabricated from carbon by focused-ion-beam-induced chemical vapor deposition on a Si surface, about 10× larger than the real virus (micrograph by Reo Kometani and Shinji Matsui, University of Hyogo, Japan). (**e**) Vitrified phage T4; full (*dark*) and empty (*clear*) heads are seen. At this defocus, tail spacing of 40.5 Å is enhanced. (**f**) Thin section through a phage T4-infected *E. coli* bacterium (Dr. Michel Wurtz; micrograph by M. Maeder). Bar, 0.5 μm. (**g**) Air-dried and shadowed broken (by osmotic shock) phage T4 that has released its packaged DNA (Dr. Michel Wurtz; micrograph by J. Meyer). The partially extended DNA indicates high packaging density within the head. Bar, 100 nm

ultramicrotome, sections are picked up on a naked copper grid, and stained with heavy metal salts to enhance contrast.

Bacteriophage T4, of the *Myoviridae* family, is an excellent model that illustrates the complementarity of the information that can be obtained by a variety of conventional TEM procedures (Fig. 3.2). A double-stranded (ds) DNA tailed virus that infects *Escherichia coli*, phage T4 is characterized by a large elongated (prolate icosahedral) capsid (head) (1,150 Å long, 850 Å wide), a tail made of two concentric protein cylinders (1,000 Å long, 210 Å wide), and a complex base plate (460 Å diameter) with six long tail fibers attached to it (1,450 Å long). It is approximately 200 nm long and 80–100 nm wide. The tail is surrounded by a contractile sheath, with helical symmetry (with a pitch of 40.5 Å), which contracts during infection of the bacterium. dsDNA is densely packaged, ~24 Å interstrand spacings, inside the head [16].

3.2.3 Sample Preparation Techniques: Freeze-Substitution, High Pressure Freezing, and Cryo-Ultramicrotomy

The high vacuum required in the electron microscope column does not allow visualization of living cells in liquid water but, as discussed above, specimen ultrastructure is altered as soon as water is removed. The approaches described thus far increase the stability of dehydrated biological samples by introducing chemical fixatives such as a negative stain agent or osmium tetroxide. As we will see below, the closest approach to the native state is the vitrification of the biological material. This approach is restricted to "thin" structures (viruses are ideal samples), which can be cryo-fixed on a specimen grid by plunge-freezing (rapid immersion in a cryogenic liquid; see below). Thick samples (100–300 μm) can be cryo-fixed by high-pressure freezing.

Freeze-substitution is a low-temperature dehydration method; it can be considered a hybrid technique that bridges the gap between cryo-EM and conventional EM. To ensure good structural preservation, biological samples are cryo-fixed, for example, by high-pressure freezing (see below). They are then freeze-substituted, which consists of dehydration at −90 °C by replacing ice with an organic solvent based on acetone, ethanol or methanol and, if desired, chemical fixatives. After dehydration (or substitution), the temperature is raised to −40 °C to −30 °C and the sample is resin-embedded. Cryo-fixation by high-pressure freezing provides excellent preservation of cellular and molecular architecture for samples between 100 μm and 300 μm thick. This approach has provided realistic views of the crowded eukaryotic cytoplasm, in which protein concentration is ~300 mg/ml.

High-pressure freezing is an approach that permits optimal preservation of biological ultrastructure by cryo-fixation, provided the biological material is not destroyed by water crystallization. This requires high cooling rates (higher than 10,000 K/s) at atmospheric pressure (~1 bar). The native structure is well preserved by conventional freezing (or vitrification) of specimens up to 0.5 μm. Thicker specimens (up to 30 μm) can be frozen without visible ice crystal damage using antifreezing (cryoprotectant) agents, but chemical fixatives are needed; this implies deviation of the native structure and, therefore, loss of high resolution details. Thicker samples (~200 μm) can be adequately frozen under high pressure (~2,000 bars), in which physical properties of water are changed (high pressure is, in fact, a physical cryoprotectant that prevents ice nucleation and crystal growth and increases the depth of vitrification).

Material embedded in vitreous ice is ideal for cutting directly into ultrathin cryo-sections (with a cryo-ultramicrotome using diamond knives) and can be observed in a cryo-electron microscope at liquid nitrogen temperatures, although cutting-induced deformations and artifacts are a serious problem. This approach is known as cryo-EM of vitreous sections (CEMOVIS). A three-dimensional reconstruction of the specimen in the cryo-section (200 nm thick) can be generated by computerized cryo-ET. CEMOVIS is the best method for preserving whole cells and tissues in their native state, and relatively high resolution is achievable [17].

3.3 Cryo-Electron Microscopy of Viruses

3.3.1 Basic Concepts and General Experimental Design

In cryo-EM of viruses (and other types of biological macromolecular assemblies), the specimen is not dehydrated; instead, it is included in a thin film of water (~0.2 μm or less) that is vitrified [18]. An aliquot of a virus suspension is placed on a grid coated with a holey carbon support film. Excess sample solution is blotted away with filter paper, leaving a very thin film (up to several thousand Å) of virus suspension on the grid, which is rapidly plunged into a cryogen (such as ethane slush) at liquid nitrogen temperature. Freezing is so rapid that water molecules are immobilized in an amorphous state, so-called vitreous or glass-like ice, avoiding physical damage of the specimen by ice crystal formation (Fig. 3.2e). This method provides a natural water-like environment for the virus particle, with no staining or fixing artifacts, and the native structure is preserved with structural integrity near atomic resolution. To avoid devitrification of the sample, the grid bearing the virus sample is maintained at liquid nitrogen temperatures while it is introduced in the microscope with a specimen cryoholder, and for several hours for the recording of micrographs. Alternatively, the grids can be stored indefinitely in liquid nitrogen. Heat transfer between the cryoholder and the microscope can be a serious problem and cause the specimen to vibrate (a result of liquid nitrogen boiling in the cryoholder dewar). Mechanical drift can be another obstacle when the cryoholder is not stably inserted in the column (the specimen must not move by more than a few Å during a 0.5–1 s exposure).

At approximately −170 °C, ice does not sublimate at a significant rate, even in the high vacuum of the electron microscope. Transfer steps must be very rapid to minimize contamination by deposition of hoarfrost (from water vapor), and the cryoholder and the interior of the microscope column are equipped with a cryoshielding device and a twin-blade anti-contaminator, respectively, which closely sandwich the grid. Holey carbon film grids are used as is, or can be glow-discharged, washed with acetone vapor to enhance hydrophilicity, or replaced by continuous non-perforated carbon film grids.

The concentration of the virus suspension is critical for obtaining a thin film with even specimen distribution after blotting; negative staining analysis is used to determine an appropriate range of particle concentration. Another limiting factor is that of the solutes used during virus purification by ultracentrifugation in gradients such as glycerol, sucrose and cesium chloride. These chemicals will bubble during observation and must be removed by floating the grid on a drop of distilled water (routine for preparing negatively stained samples) or dialysis of the selected virus fractions.

The cryo-negative staining method preserves biological samples in the frozen hydrated state, incorporating negative staining to improve the signal-to-noise ratio of the images. As a contrast agent, ammonium molybdate is normally used (at physiological pH) when this procedure is used for conventional cryo-EM; uranyl formate or acetate is used in the carbon-sandwich method, in which the sample is trapped between two carbon films and is frozen in liquid nitrogen. Unlike air-dried negative staining at room temperature, in the carbon-sandwich method, sample hydration is maintained in all steps (preserving its structural integrity) and specimens are observed at liquid nitrogen temperature [19].

3.3.2 Magnification, Calibration and Minimal Electron Dose

The definition of absolute magnification is necessary, and real magnifications should be calibrated periodically to avoid uncertainty in image processing. The hysteresis of magnetic lenses and changes in specimen height in the objective lens cause small changes in nominal magnification that must be considered in high-resolution studies. Comparisons of three-dimensional reconstruction (3DR) of a virus whose dimensions have been calculated by X-ray analysis is a precise approach to determining the absolute scale; another approach is to mix the test virus with a bacteriophage T4 solution, using the 40.5 Å axial spacing of the bacteriophage tail sheath as an internal magnification standard.

Vitrified specimens are highly sensitive to electron irradiation, and the act of imaging with electrons could eventually destroy the specimen. After insertion of the specimen stage into the microscope, the cryogrid is searched at low magnification (<2,000–3,000×) and irradiation levels. An experienced cryomicroscopist is able to select optimal areas by visual inspection; dark areas in the grid are too thick for electron penetration, while bright areas are probably too thin for correct virus particle embedding or can even be dried. Most viruses can be seen directly at low magnification as dark spots, allowing assessment of particle distribution and concentration. These particles are destroyed at high magnification, but can be observed directly for a few seconds as a "bubbling" effect. These sacrificed specimen areas are a convenient form of direct evidence of vitrified specimen quality, assuming that intact adjacent areas in the grid have similar particle concentration and ice thickness.

The strategy for image acquisition is based in this approach, and is done following a relatively "blind" (but not random) procedure, in the sense that the area selected for recording will be seen (and selected or discarded) once the micrograph is developed. Cryo-microscopes are equipped with a low dose system to limit exposure to electrons. The system switches among three states, searching

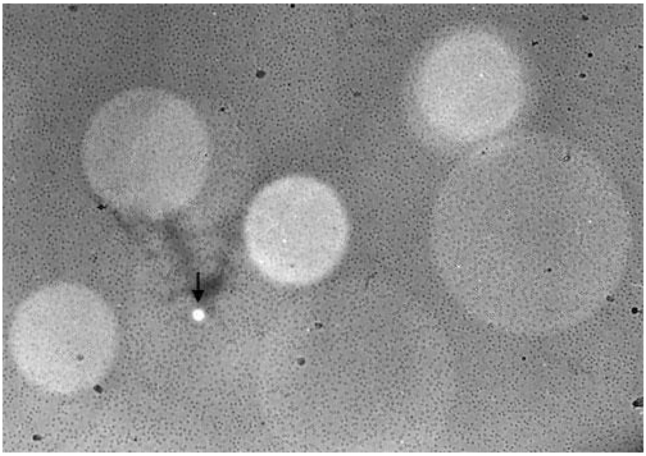


Fig. 3.3 Radiation damage in vitrified bacteriophage T4. In this low magnification field, the smaller circular areas have been sacrificed during focusing; the larger circular areas correspond to recorded regions (the electron dose needed to acquire an image causes no visible effect, but the microscopist exposed the recorded area for longer times to obtain this general view). The bubbling effect can be seen in the particles of some of these exposed areas. The bright area (*arrow*) is literally a hole (made with a focused beam) in the thin vitrified layer that contains the viral particles in suspension

for promising specimen areas at low magnification, focusing at high magnification, and recording at the same high magnification (40,000–60,000×). All of the electrons that pass through the specimen are used to form its image. The image is focused in an adjacent sacrificed area by observing the Fresnel fringes that appear at the edge of radiation-induced bubbles (that first form within the specimen). In this step, astigmatism is corrected and other adjustments are made. Images are recorded with electron doses between 5 $e^-/\text{Å}^2$ and 20 $e^-/\text{Å}^2$, which produces micrographs of sufficient optical density and signal for subsequent image processing steps (Fig. 3.3).

Imaging on a CCD camera exposes a much smaller sample area, and thus minimizes possible beam-induced specimen movement or charging, similar to the "spot-scan" approach. Inelastic scattering by the sample leads to radiolysis of macromolecules and the embedding medium (vitrified water). The radiolysis products require more space than the original molecules, leading to increase of internal pressure in the sample and ultimately to breakdown (bubbling) of the sample matrix. The pressure increase and the bubbling cause movement in the exposed area of the sample and blurring in the image, thereby attenuating the signal at higher resolution. Inelastic scattering also leads to ejection of electrons from the sample and accumulation of positive charge (sample movement and charging must be minimized).

3.3.3 Imaging Conditions: Contrast Transfer Function

Another drawback in cryo-EM is related to imaging conditions. Images are noisy, because there is no staining agent that enhances the scattering. The amplitude contrast of unstained biological specimens is thus quite low (~10 % or less), and phase contrast is the dominant contrast mechanism, which requires defocusing the microscope objective lens (*i.e.*, images must be deliberately out of focus to enhance contrast). Extensive computer processing is then needed to correct the images and recover structural details.

In a perfect electron microscope, the 2D image recorded on film would be a realistic representation of the projection of the 3D scattering matter in the specimen. As in any other optical system, the image of an object is modulated by the lens characteristics such that each point in the image is the result of a convolution of the corresponding point of the object and a mathematical function, known as the contrast transfer function (CTF). In the case of the electron microscope, the CTF depends on the spherical aberration of the objective lens, the effects of defocusing and inelastic electron scatter, and the partial coherence of the beam. This fundamental concept implies that all TEM images are altered by the CTF and must be corrected for accurate interpretation of the structure of a biological specimen. The CTF is a characteristic function of each microscope (*i.e.*, the spherical aberration coefficient of the objective lens) and the imaging settings (defocus level of the objective lens, beam coherence and accelerating voltage).

By choosing different defocus settings (focal series or pairs are usually taken), specific structural features (or frequencies in the Fourier space) of the image can be accentuated selectively at the expense of others. In appropriate imaging conditions, a contrast-enhancing effect can thus be obtained. CTF-modulated information implies the existence of frequency ranges in which there are no data (zeros in the CTF), regions in which the information has reversed contrast (between the first and the second zero, between the third and fourth zero, and so on), and a fall-off of useful information at higher frequencies (the envelope function), dependent mainly on beam coherence. In the same way that observing a sample by different conventional EM techniques provides complementary data, recording several images at different focal settings to enhance distinct specimen features is a “must-do” (Fig. 3.4).

Cryo-EM micrograph quality can be analyzed by inspection of the CTF zeros or Thon rings in the optical or computed diffraction patterns of images (the sum of the Fourier transform amplitudes is termed the average power spectrum). An image free of aberrations has the signal in the power spectrum as a series of concentric circular rings (zeros or Thon rings) that extends outward from the center of the Fourier transform; the greater the extent of rings, the higher the resolution in the image. Images with drift or astigmatism are easily detected and excluded from further structural analysis (Fig. 3.5).

As good contrast is required for the cryo-micrograph, a compromise should be reached among the factors that modulate contrast. The use of a small objective

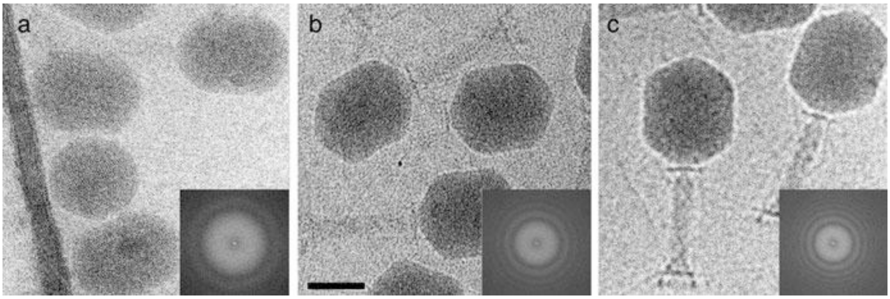


Fig. 3.4 Unstained, frozen, hydrated specimen of phage T4 in vitreous ice, prepared on holey carbon film. Focus series of phage T4 under imaging conditions that enhance dominant spatial frequencies corresponding to (**a**) interstrand spacings between double-stranded DNA filaments (~26 Å) within the head in a relatively close-to-focus image, (**b**) tail sheath spacing (40.5 Å) in a regular defocused image, and (**c**) general morphology in a highly defocused image. Note that detail of the tail is almost lost in the close-to-focus image. Insets show FT (Fourier transform) for each cryomicrograph; the further from focus, the closer the Thon ring is to the FT center. Bar, 50 nm

aperture increases image contrast, but might introduce astigmatism, and lower acceleration voltages increase image contrast, but cause more radiation damage to the specimen. Higher defocus means greater contrast, but also more rapid dampening of the envelope function and many zero transitions of the CTF; CTF of the envelope function is reduced using a brighter beam, such as that generated by FEG. Image contrast can also be improved by additional instrumentation. Energy filters remove inelastically scattered electrons, responsible for much of the noise in micrographs, thus improving the signal-to-noise ratio of the images.

3.4 Cryo-EM Image Processing and Three-Dimensional Reconstruction

Electron microscopes have a large depth of focus compared to virus size. This means that all characteristics at different heights in the virus are in focus simultaneously, and each image recorded in cryo-EM conditions is a 2D projection of 3D scattering density in the specimen. Superposition of all structural information from different levels results in a garbled image. In a single projected view, a specific feature can be derived from any level in the specimen. To reconstruct the native 3D structure, many images must be recorded to obtain different views of the object; viruses tend to be randomly oriented in the vitreous specimen layer [20]. Alternative views can be obtained by tilting the specimen in the microscope. A 10° tilt, which can be accommodated by the anticontaminator, is usually sufficient to give an adequate sampling of orientations. This approach is necessary, for example, to identify the absolute handedness of icosahedral viruses using cryo-EM (see Chap. 2).

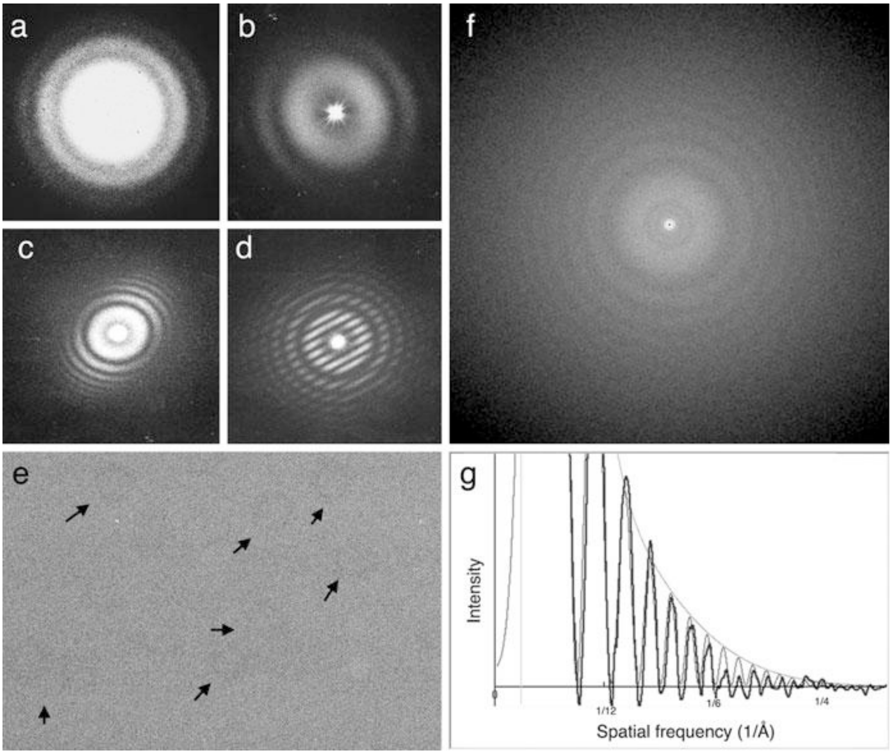


Fig. 3.5 Power spectrum (equivalent to diffraction pattern) analysis of cryomicrographs with distinct aberrations. (**a**) An aberration-free image shows circular concentric Thon rings (zeros). The exact underfocus can be calculated from the first zero position. Signal near the center of the FT corresponds to lower spatial frequencies (low resolution details); signal extending outward from the center corresponds to high spatial frequencies. The last Thon ring indicates the maximum detectable frequency (*i.e.*, resolution) that could be rescued from the 2D projection image. (**b**) Astigmatic migrographs are detected because Thon rings appear elliptical. This aberration can originate from the illumination source and is corrected by the objective lens (an objective aperture that is not well centered can also cause astigmatism). (**c**) Sample drift occurs when resolution is lost preferentially in one direction of the power spectrum. Drift can be thermal (temperature instability in the refrigeration system of the objective lens) or mechanical (stage movement during exposure). Vibrations (acoustic or physical due to liquid nitrogen bubbling in the cryo-stage dewar) limit resolution in all directions. (**d**) Double exposure to determine drift speed. Taking two exposures on the same micrograph, spaced 15–30 s apart, produces a pattern with these typical bands; narrow bands mean larger drift movements and vice versa. (**e**) An excellent image of a rabbit hemorrhagic disease virus solution taken in a FEG 200 kV microscope, in which contrast is very low (*arrows* indicate some viral particles). There is no doubt as to the quality of this micrograph, since the average (**f**) FT calculated from many boxed particle images shows Thon rings to ~4 Å resolution. (**g**) Average power spectrum curve (*thick line*) is used to fit theoretical contrast transfer function (CTF) curves (*thin line*). Underfocus values of 1 μm; first zero of the contrast transfer function at 16 Å

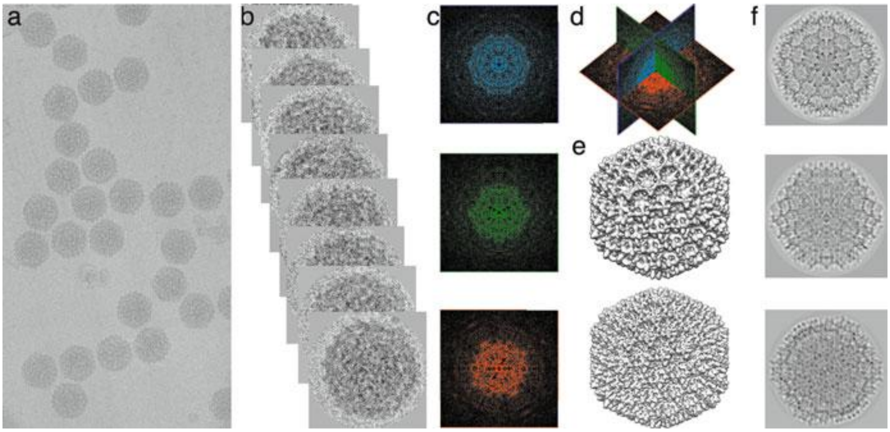


Fig. 3.6 The central section theorem. (**a**) A digitized micrograph containing unstained vitrified viral particles of infectious bursal disease virus (IBDV, a double-stranded RNA virus that infects birds, with a diameter of 650–700 Å). (**b**) Particles are individually selected (*boxed*); (**c**) their 2D FT are calculated from the 2D projections representing different views of the same icosahedral structure. Particle orientation is determined (for example, within 1°) following the projection matching or common lines methods, and the 2D FT are combined in the correct orientation in (**d**) the 3D FT of the virus. (**e**) Accuracy of the initial low resolution 3D map is improved (refined) as the model incorporates more particles and orientation parameters are defined at higher precision in the asymmetric unit. (**f**) Three reprojections from the 3D map are shown

In 1968, DeRosier and Klug found that a 3DR could be recovered from micrographs with 2D images [21], using a mathematical device termed the Fourier transform (FT). Sir Aaron Klug received the Nobel Prize in chemistry in 1982. The FT separates different components of the image into all 2D spatial frequency components (or constituent waves). The FT of each image corresponds to a central plane of the 3D FT of the object density; this is the central section theorem. Distinct projected views fill in different central sections of the transform; the larger the number of views included in the 3D FT, the better the sampling and the 3D object density recovered in Fourier synthesis (Fig. 3.6); many computer programs are available for Fourier analysis and synthesis. In addition to this approach, FT are used for general quality control (evaluating the CTF of each micrograph), correcting the defocus effects (in high resolution studies), and determining particle orientation.

The FT of an image (termed reciprocal or frequency space) is an equivalent, reversible representation of the image (real or Cartesian space). FT provides a mathematical description of the diffraction pattern when imaging an object with a lens (optical or electromagnetic). In the diffraction pattern (the object's FT) produced in the back focal plane of the lens, waves are recombined (positive and negative interference) to form an image. This concept of FT to give a diffraction pattern, followed by Fourier synthesis to produce an image, is central to image processing. In the real-space representation, the parameters that define an image are Cartesian coordinates (x, y, z) and an associated density value. In the FT representation, these parameters are reciprocal coordinates, an associated amplitude (strength) and an associated phase (which specifies position relative to origin of the wave representing spatial frequency). FT also allows filtering according to spatial frequency. The highest spatial frequencies usually have a large fraction of noise; by eliminating these frequencies in Fourier space and using the inverse FT to convert back to real space, the resulting image is less noisy [22]. The same principles used for 3D image reconstruction are applied for structure determination in X-ray crystallography (Chap. 4).

3.4.1 Digitization

The micrograph must first be scanned with a high precision film scanner to convert it to an array of numbers that represent the darkness on the film; it is then in a form suitable for computer processing. The Shannon sampling theorem states that scanning step size to scan a micrograph should be at least twice as fine as the finest detail to be analyzed (highest frequency or resolution desired), termed the Nyquist sampling. That is, if the resolution desired is 10 Å, 1 pixel should correspond to 5 Å; a 50,000× micrograph scanned with a 7 μm step size corresponds to 1.4 Å/pixel in the specimen; in practice, an image is typically scanned at a scan size of one-fifth to one-third the desired resolution. Another factor to be considered is the limited dynamic range of the cryo-electron micrographs; the scanner should be adjusted to capture this range optimally without image degradation and loss of resolution. This step is omitted if the cryo-electron microscope is equipped with a CCD and digital images are available directly. CCD cameras offer other advantages, such as easy handling, high signal linearity as a function of the number of incoming electrons, and a large dynamic range.

3.4.2 Image Screening

Image quality can be assessed visually by an experienced microscopist to exclude micrographs or CCD frames with high ice contamination, astigmatism, and beam-induced specimen movement. Quantitative analysis is performed by inspection of the CTF rings by calculating the average power spectrum of the micrograph (Fig. 3.4). The CTF is determined by fitting simulated CTF to the observed Thon rings in the calculated power spectrum of the image, which is generally easier for highly defocused images. A CCD camera provides instant feedback for image quality, as FT is calculated dynamically during image acquisition. Once the best micrographs are selected, particles are individually picked (boxed) by manual or automatic methods.

3.4.3 Determination and Refinement of Particle Orientation and Origin

After selection of many thousands of single-particle images (in the best cases), the center of symmetry of each image is established using cross-correlation with a reference image or with the same image rotated 180°. To calculate the 3DR of the object, it must then be determined whether a projection is the front, side or top view, or somewhere in between. Particle orientation determination then begins, that is, the viewing direction relative to the icosahedral symmetry axes. Particle orientations are specified relative to the smallest repeating unit of the icosahedron, termed the asymmetric unit.

The first step of this process requires an initial 3D model as a reference. Many icosahedral virus density maps are available from the Electron Microscopy Data Bank (EMDB: http://www.ebi.ac.uk/pdbe/emdb). The map must be scaled to match the dataset under study, and then low-pass filtered (~35-40 Å resolution) to reduce model bias. A more sophisticated approach involves computing the common lines of a small number of virus images, from which an initial model can be calculated. Due to the inherent difficulties of processing low contrast close-to-focus images, it is often useful to collect focal series with a range of defocus values to facilitate orientation determination. The most time-consuming step when calculating a 3D virus map is the refinement (improvement) of the orientation and center parameters. This analysis involves comparing the particle images with 2D reprojections from a 3D model; projection matching and common lines are the methods most commonly used to determine these parameters [23].

In the projection matching method, the model is projected computationally in all possible directions to generate a set of reference images that are used to match each image with one projection corresponding to the appropriate orientation. The initial low resolution model is updated at the end of each refinement cycle during this iterative process until no further improvement is obtained. Projection matching can be carried out in real or Fourier space.

In the common line method, the central section theorem provides the mathematical foundation for orientation determination of projection images from the same object. The FT of two projection images with different orientations are planes in the 3D FT of the object (a virus) that intersect one another. This intersection line is termed common line and has the same values in each plane for amplitudes and phases (Fig. 3.6).

For an icosahedral virus, a single image with a defined orientation in the asymmetric unit has associated 59 symmetry-related orientations. Each of the 60 equivalent orientations defines the corresponding planes in the 3D FT. The intersection line between any of these planes is known as a self-common line, and there are a maximum of 37 pairs of self-common lines in the FT of a 2D projection image. To estimate individual particle orientation parameters, an exhaustive search is made for all possible orientations within the asymmetric unit. This method does not require an initial model, but the useful information is limited to the self-common lines.

Cross-common lines arise from the application of icosahedral symmetry between a particle and another particle (with different orientations); there are 60 pairs of cross-common lines between any two images. The orientation angles are initially estimated by assuming an orientation and comparing the agreement of the phases between the common line pairs. All the information in the 2D FT is therefore useful for estimating particle orientation.

3.4.4 Three-Dimensional Reconstruction in Fourier Space and CTF Correction

Considering the central section theorem, and using direct Fourier inversion, a 3D map can be calculated from 2D images [24]. In this method, CTF correction must be applied to the individual particle images before constructing the 3D FT from each 2D FT. It is essential to ensure that sufficient particle orientations have been provided for adequate sampling of Fourier space, since interpolation errors can limit the resolution. This step normally requires intensive memory and computational resources. Although there are other methods for 3D reconstruction of icosahedral viruses, they lie outside our scope here.

The CTF that enables visualization of unstained specimens must be corrected (deconvoluted) to retrieve the high-resolution signal and obtain a 3DR that reliably represents the structure of the specimen. The CTF reverses, removes and attenuates data in frequency ranges in the FT of the image. Precise determination of the Thon ring pattern (CTF zeros or points of minimum contrast) in the power spectrum allows reversing the phases in the appropriate regions of the FT. Multiple images with complementary defocus must be used to fill in the missing data near the zeros. This method is reliable when FEG of high-voltage (200–300 kV) data are used.

3.4.5 Resolution

The resolution of the 3D map depends on the total number of particles, distribution of views within the asymmetric unit, accuracy of center and orientation parameters, the quality of the data (*i.e.*, the radius to which reliable data extend in the FT), and many other factors such as the CTF correction. The resolution in 3DR is less objectively determined than in X-ray crystallography, for which resolution has come to mean the highest spatial frequency contributing meaningfully to the map [25].

Assessment of the resolution is performed statistically, by splitting the particle set in half and computing two independent maps from each half-set. The two maps (often referred to as even and odd maps) are then correlated as a function of spatial frequency to determine the extent to which the fine structural details have been determined reliably (termed FSC method, Fourier Shell Correlation).

Amplitude scaling is used to restore high-resolution features, using X-ray solution scattering profiles as a reference, or by determining an effective temperature factor (B-factor) (see below). This process of map sharpening scales specific spatial frequencies (resolutions) above others.

Although image variability is a limiting resolution factor, it can also indicate multiple conformations. Conformational states of a macromolecular complex are difficult to classify, as variability must be differentiated from viewing geometry. If these (normally evasive) conformations are assigned to a time-ordered sequence, the resulting density maps can be used to make elegant time-resolved movies of dynamic processes such as virus particle maturation (Chap. 13). As noted above, cryo-EM is the best approach for trapping and visualizing transient structures with a lifetime of a millisecond or longer, as the specimen is rapidly vitrified. Buffer conditions can also be adjusted to extend the lifetime of a transient state; this approach has been used in virus maturation analysis for systems such as herpes simplex virus and several dsDNA bacteriophages (reviewed in [26], an references therein) (see Chap. 13).

3.4.6 Visualization and Structure Representation

Surface rendering of large volume data is very demanding computationally, and can be a rate-limiting step in structural interpretation. For this reason, structural units from the macromolecular complex are normally segmented. Segmentation of individual components permits non-icosahedral averaging of structurally similar components, to enhance signal-to-noise ratio and further improve the resolution of the averaged subunits. For example, the $T = 13$ capsid of rotavirus allows the nonicosahedral averaging of 13 subunits of VP6 molecules of the asymmetric unit.

In some cases, the structural asymmetric unit boundaries can be established by contouring the map at different levels, based on its compactness and contacts with neighboring densities; homologous models and/or biochemical evidence are generally used. At subnanometer resolutions in the 6–9 Å range, secondary structure elements (SSE) are discernible. The α-helices appear as cylinders of density with diameters of 5–6 Å, and β-sheets as flat densities or planks. An example at this resolution of the final structural model of a spherical virus, the 40 nm-diameter fungal virus *Penicillium chrysogenum* virus (PcV), is shown in Fig. 3.7.

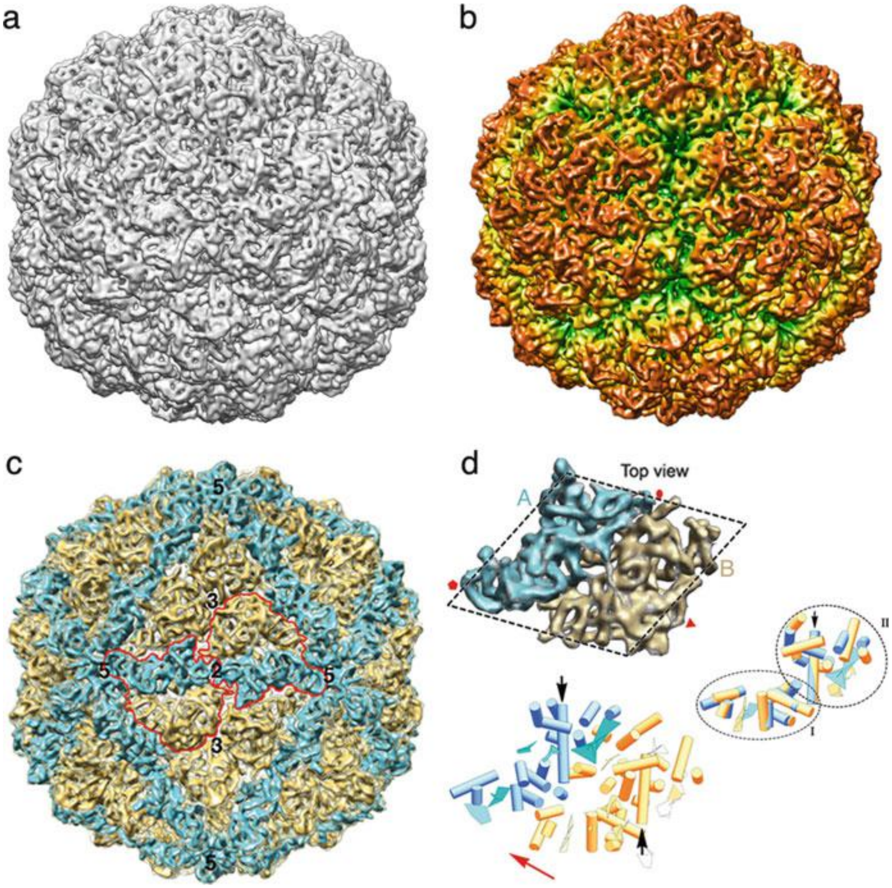


Fig. 3.7 Three-dimensional cryo-EM of the fungal virus Penicillium chrysogenum virus (PcV) at 8.0 Å resolution (based on a FSC threshold of 0.5). (**a**) Surface-shaded representation of the outer surface of the PcV T = 1 (diameter is 400 Å), viewed along a two-fold axis of icosahedral symmetry. Bar, 100 Å. (**b**) As in (a), with the radially color-coded outer surface, which emphasizes the most prominent features as 12 outward-protruding pentamers. (**c**) As in (a); here the asymmetric unit is *cyan* and *yellow*. The asymmetric unit is formed by two ellipsoid-like structures with roughly similar morphology. Boundaries for two asymmetric units are outlined in *red*. Icosahedral symmetry axes are numbered. (**d**) Segmentation of the asymmetric unit, the PcV capsid protein monomer. The *dashed line* highlights the rhomboidal shape, and protein halves A (*cyan*) and B (*yellow*) are indicated (*top left*). At this resolution, secondary structural elements (SSE) of the PcV capsid protein are identified. Cylinders, α-helices; planks, β-sheets. *Black arrows* indicate the ~37 Å-long α-helices of both PcV capsid protein halves. When one of the halves (treated as a rigid body) is superimposed on the other by an ~45-Å translation (*red arrow* indicates translation direction), the relative spatial locations of 13 α-helices and two planar regions are close and require only minor local adjustment. Preserved SSE between both PcV capsid protein halves consist of two domains, domains I and II. This subnanometer resolution analysis indicates that the two PcV capsid protein halves have a comparable structural signature, suggesting ancestral gene duplication (Reproduced/adapted with permission from American Society for Microbiology, ([27], doi:10.1128/JVI.00432-10))

3.4.7 Hybrid Methods: Combining High-Resolution Structures with Cryo-EM Maps

As for most research areas, hybrid approaches in structural virology involve multi-disciplinary studies that include a broad set of techniques [28, 29] (see Chap. 7). Here we will refer briefly to the hybridization of atomic structures obtained by X-ray crystallography (see Chap. 4) and by NMR (see Chap. 5) in conjunction with cryo-EM 3DR [30–32]. A more extensive description of combined approaches is included in Chap. 7.

Cryo-EM structural models yield significant information about macromolecular complexes such as viruses (at low to medium resolution), or the folding of individual components (at subnanometer resolution). If more detail is required in the structural model, and higher resolution structural models of individual subunits in the complex are available, the high-resolution structures can be fit ("docked") to the cryo-EM electron density map to obtain a "pseudo-atomic" model of the complex. Fitting tools use rotational and translational searches of a given model within the density map as described in Chap. 7. Such fitting can reveal, for example, interaction interfaces between individual components in atomic detail, providing structural insight into how individual intersubunit interactions might affect complex function. This crucial information cannot, in general, be inferred from the structures of the individual components, and 3D crystals of large functional complexes tend to be very difficult to obtain.

Many studies of large viruses or of the dynamic character of particles that constitute intermediates of virus assembly and function, which are difficult to crystallize, have followed this hybrid approach. Another example is the study of neutralizing monoclonal antibodies and their antigen-binding (Fab) fragments or cell receptors bound to the surface of a virus (see Chaps. 7 and 15). To map the Fab or cell receptor footprint on the virus surface, the antibody-virus complex is studied by cryo-EM and combined with the crystallographic structure of both; this approach was used successfully, for example, for different picornaviruses (rhinovirus, poliovirus and foot-and-mouth disease virus). It is nonetheless possible that the structure of a component will differ when assembled with other components. There are numerous algorithms available for rigid body as well as flexible fitting of components into the larger structure. A quantitative measure of fit accuracy is determined by global and local cross-correlation coefficients.

As the resolution of the cryo-EM map improves, discrepancies in some parts of the individual subunits are evident once fitted into the cryo-EM map; they can be adjusted optimally by permitting flexible fitting of specific domains and/or SSE. Structures in crystals might not always represent conformations related to in situ molecular interactions in physiological conditions. This technique has also been used to validate homology-derived models [33]. Flexible docking can be achieved manually using molecular graphics, but the most-used methods are based on molecular dynamic simulations or on normal mode analysis (Fig. 3.8) (see Chap. 19 for a general description of these theoretical approaches).

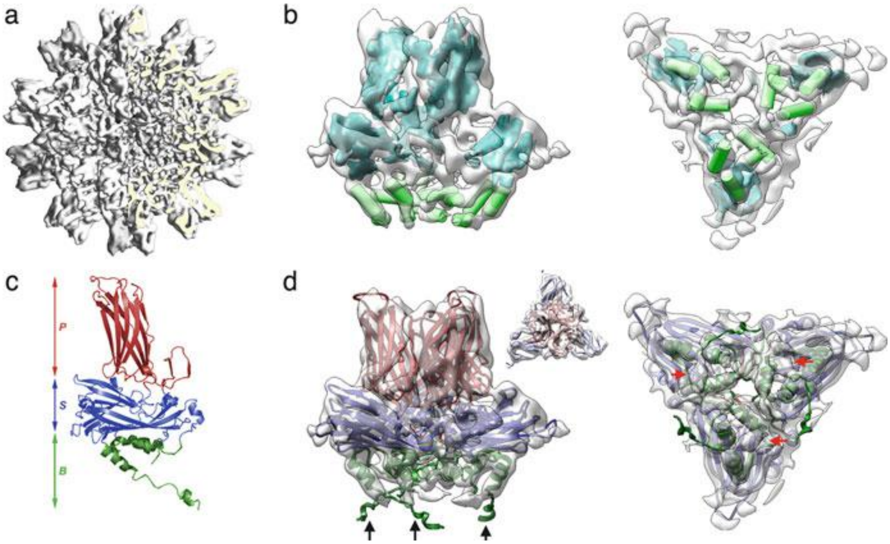


Fig. 3.8 Three-dimensional cryo-EM of T = 1 subviral particles (SVP, ~23 nm in diameter) of IBDV at 7.2-Å resolution. (**a**) Surface-shaded representation of the complete SVP (a quarter has been computationally removed). The most prominent features are 20 VP2 trimeric protrusions of the capsid protein. (**b**) Identification of SSE: green cylinders, α-helices; blue planks, β-sheets. (**c**) The 2.6-Å VP2 X-ray model (PDB entry 2GSY). The VP2 subunit is folded into three domains, termed the projection (*P*), shell (*S*), and base (*B*) domains. Domains S and P are β-barrels. The B domain consists of four N- and C-terminal α-helices that line the inner capsid surface. (**d**) The VP2 trimer X-ray fitted into the corresponding density in the cryo-EM map. The concordance of the two maps is clear in the matching of P and S domain β-barrels. *Arrows* indicate the cut directions to show P and S domain fits (*inset*). Whereas three α-helices fit remarkably well in the cryo-EM structure, a C-terminal α-helix has no corresponding density in the cryo-EM map (*black arrows*), which shows a discernible vacant rod-shaped density adjacent to the C-terminal α-helix (*red arrows*). The proposed movement involves allowed rigid body rotations of the C-α angles in the loop preceding this α-helix. The conformational flexibility of this C-terminal α-helix is critical for IBDV capsid assembly (Reproduced/adapted with permission from American Society for Microbiology ([34], doi:10.1128/JVI.00077-07])

Cryo-EM can also be used as an aid for the atomic resolution of viruses and other biomolecular complexes by X-ray crystallography. Diffraction of crystals of large complexes (such as viruses) requires various phasing methods (see Chap. 4). EM maps are used to solve the phase problem when no phasing by molecular replacement or anomalous scattering is feasible, and heavy atom derivatives do not scatter sufficiently to alter structural factors. A phase extension approach based on molecular envelopes obtained by 3D cryo-EM (or even by conventional negative staining EM) can be used for these crystals (see also Chaps. 4 and 7). If the number of complexes and their positions in the unit cell are known, the envelope specifies those portions of the unit cell whose density is at a constant level. This provides a "solvent flattening" constraint, which with the additional constraint of noncrystallographic symmetry, is imposed iteratively, starting with a low EM resolution model; the model is progressively improved by introducing diffraction data to higher and higher resolution. This approach has worked well with a number of icosahedral viruses.

3.5 Near-Atomic Resolution of Virus Structures by Cryo-EM

High-resolution 3D cryo-EM is becoming an important structural tool for determining detailed 3D structures of large biological complexes, such as viruses, that are often too large or too heterogeneous to be studied by "conventional" high-resolution methods like X-ray crystallography (Chap. 4) and NMR spectroscopy (Chap. 5) [25, 35–37]. Several recent studies with icosahedral viruses have pushed the limit of single-particle cryo-EM reconstruction to near-atomic resolution (~3.6–4.5 Å). At this resolution, many detailed structural features can be resolved, including deep grooves and pitches of helices, interstrand distance of ~4.4 Å in β-sheets (except for hydrogen-bonded regions), densities for loops and voluminous amino acid side chains, and the zigzag pattern of Cα atoms separated by ~3.8 Å [38]. These structural data can be used as constraints to trace amino acid backbones (Cα model building) using modeling tools. When the X-ray model is available, comparison of the X-ray and cryo-EM structures of the same virus reveals excellent agreement to the level of amino acid side chains.

Several factors have contributed synergistically to this improvement in resolution:

1. *Sample handling*. Structural homogeneity is more important than purity. It is desirable to avoid possible structural damage caused by density gradient centrifugation, since structural heterogeneity (disorder) can be a limiting factor. Icosahedral viruses are regular, more rigid structures than lipid-containing viruses. Dynamics states (*i.e.*, conformational flexibility) also limit resolution.
2. *Cryo-EM imaging*. Improvements in cryo-EM instrumentation such as more stable electron beams and sample holders have reduced specimen drift, a limiting factor in achieving atomic resolution. The high voltage cryo-microscope (FEG, 300 kV) must be optimally aligned so that the electron beam is parallel, with minimal beam tilt. Only those images with visible CTF rings up to 5 Å in their power spectra should be selected (the resolution of the 3DR usually extends beyond the last detectable of Thon ring).
3. *Computational methods for processing noisy cryo-EM images*. Structure refinement is carried out as a two-step, iterative procedure based on projection matching, orientation-origin parameter determination and 3DR by gradually pushing toward higher resolution. High-resolution images are often recorded under close-to-focus conditions, and many thousands of these low-dose phase-contrast images must be used to enhance the high-resolution features. This enormous number of particle images is also related to the fall-off in the Fourier amplitudes at high spatial frequencies (equivalent to high B factor in X-ray

crystallography). Resolution is evaluated by the FSC criterion or by directly evaluating the structural features resolved in the reconstructed map. As in X-ray structures, some regions can have lower resolution than others, probably due to local intrinsic flexibility.

4. *Atomic modeling tools optimized for use with cryo-EM-derived density maps*. Map quality of some protein subunits can be improved by averaging quasi-equivalent subunits, *i.e*., viral capsids with a triangulation (T) number higher than 1 (provided that these subunits are found in similar conformations) (see Chap. 2). An empirical B factor can be estimated by a trial-and-error method that evaluates features such as continuity of backbone and side-chain densities; this can improve high-resolution features and prevent over-sharpening. At this resolution, individual structural components are segmented from the entire capsid and used for model building. At near-atomic resolution (~4 Å), backbone tracing is difficult and error-prone, as some densities are branched and side-chains are limited. Homologous structures obtained by sequence-based comparative modeling can also be used. There are modeling methods adapted to the characteristics of cryo-EM density maps, and modeling tools used in protein crystallography can be adjusted to verify models derived from cryo-EM maps. Evaluation of the atomic models follows the same restrictions used in X-ray crystallography (consistency with protein sequence and reasonable Ramachandran plot, among others). In any case, uncertainties are unavoidable in interpreting maps at 3.5–4.5 Å. Resolved cryo-EM maps are thus a combination of full-atom residue models (SSE-rich regions) and Cα models (regions with less-defined densities).

Near-atomic resolution of virus structures that address important biological questions has been reported for a still-limited number of viruses (reviewed in [24, 25, 35], and references therein). Cytoplasmic polyhedrosis virus was the first example (at 3.88 Å resolution) of a complete *de novo* chain trace from cryo-EM maps. Other examples include the double- and triple-layered rotaviruses, which were unambiguously equivalent to the X-ray model (Fig. 3.9); dsDNA bacteriophage ε15, which is related to bacteriophage HK97; P22, which was resolved in two morphogenetic states; bovine papillomavirus (at 3.6 Å resolution), a dsDNA eukaryotic in which key interactions were described in particle assembly; human adenovirus, which shows how the cement proteins tie together the map structural proteins in the icosahedral capsid; and dsRNA aquareovirus (at 3.3 Å resolution), which showed a priming mechanism for cell entry.

3.6 Reconstructing Viruses Without Imposing Symmetry

Many icosahedral symmetry-based virions present features/structures that do not obey icosahedral symmetry. Examples include the packaging complex, the tail or the core in the capsid interior of dsDNA tailed bacteriophages [40]. These

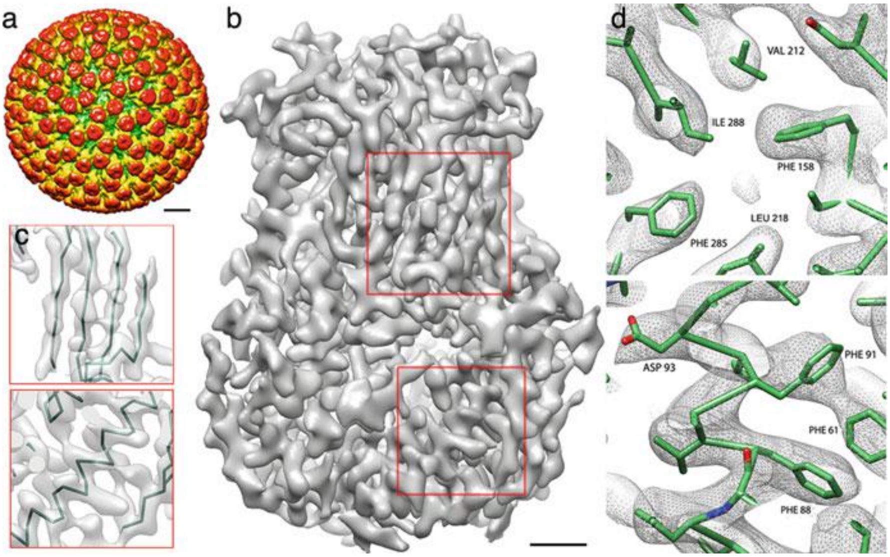


Fig. 3.9 High resolution rotavirus VP6 structure generated by 3D single-particle cryo-EM. A total of 18,125 images of rotavirus double-layered particles (DLP) collected at 300 kV were downloaded from the Grigorieff lab web page at Brandeis University (http://emlab.rose2.brandeis.edu/), and processed using Xmipp software. (**a**) Rotavirus DLP-filtered at ~25 Å. Bar, 100 Å. (**b**) VP6 trimer after 13-fold averaging at ~4 Å resolution. At this resolution, α-helices, β-sheets and densities for loops and bulky amino acid side chains are seen. Bar, 10 Å. (**c**) Close-up of selected VP6 regions included in the squares in (**b**) with the corresponding X-ray structure (PDB 1QHD). Individual β-strands (*top*) and deep grooves and pitches of α-helices (*bottom*) are seen clearly. (**d**) Densities in selected VP6 areas, shown with the atomic model of VP6; bulky side-chain densities match the X-ray structure. These structures are similar to those published by Grigorieff's group [39]

non-icosahedral components can have their own symmetry, which is obscured by the symmetry mismatch with the icosahedral capsid shell (see Chaps. 12 and 17). A map without icosahedral symmetry requires at least 60 times as much data to achieve a similar resolution. Elaborate approaches during the data processing step have enabled structural resolution of complex viruses as a single entity without imposing symmetry, using projection-matching techniques to orient the particles. The asymmetric orientation of a particle is initially determined by a time-consuming search of all possible orientations.

In the case of the T4 phage, the capsid and the tail were reconstructed separately by cryo-EM, after which the atomic models of their component proteins were fitted to determine how the individual parts contribute to the function of this complex molecular machine [41, 42] (see Chap. 11).

3.7 Reconstructing Viruses with Helical Symmetry

The capsids of many plant viruses and bacteriophages have helical symmetry [43]. In addition, the genome of some enveloped animal viruses is surrounded by a helical symmetry-based nucleocapsid; these include filoviruses (Ebola and Marburg viruses), rhabdoviruses (rabies and vesicular stomatitis viruses (VSV)), paramyxoviruses (measles and mumps), orthomyxoviruses (influenza virus) and retroviruses (human immunodeficiency virus, HIV) (see Chap. 2). In certain non-physiological conditions, some capsid proteins of icosahedral viruses may assemble into helical structures, revealing an intrinsic structural polymorphism. Helical viruses and capsids do not crystallize, and many such structures have been determined using X-ray fiber diffraction methods at high resolution. Cryo-EM recently supplied structures at ~10 Å resolution for the VSV virion trunk [44] and 3.3 Å for tobacco mosaic virus (TMV) [45], which allowed the construction of a *de novo* atomic model of TMV in a state that is biologically relevant to its assembly/disassembly.

The redundancy intrinsic to helical assemblies renders them optimal for cryo-EM averaging, since a single image provides all the projection images needed to reconstruct the 3D structure. Helical parameters are determined by measuring the pitch of the helix from layer lines in the FT of the tubular structure. A 2D classification is usually made if structural variability or heterogeneity is detected (or suspected). The 3DR is computed using the iterative helical real space reconstruction (IHRSR) method, a single-particle type approach that does not require long high-ordered filaments. The IHRSR algorithm uses the helical symmetry as a constraint to impose it on asymmetric reconstructions [46].

3.8 Cryo-Electron Tomography of Viruses

3.8.1 Basic Concepts and General Experimental Design

A preparation of icosahedral viruses observed by cryo-EM normally provides multiple views with random orientations of identical particles (without considering structural heterogeneity). Many enveloped viruses are pleomorphic, however; examples include orthomixo-, paramyxo-, retro- and poxviruses (see Chap. 11). Herpesviruses and other enveloped viruses are intermediate forms, with an icosahedral capsid surrounded by a pleomorphic structure made of protein and lipid layers (see Chap. 2). The structural variability of all these viruses means that each particle has a different shape, size or conformation, and 3DR by averaging cryo-EM images as described above is not possible. However, a reconstructed 3D image can be produced by obtaining multiple views of a single viral particle. To do this, the

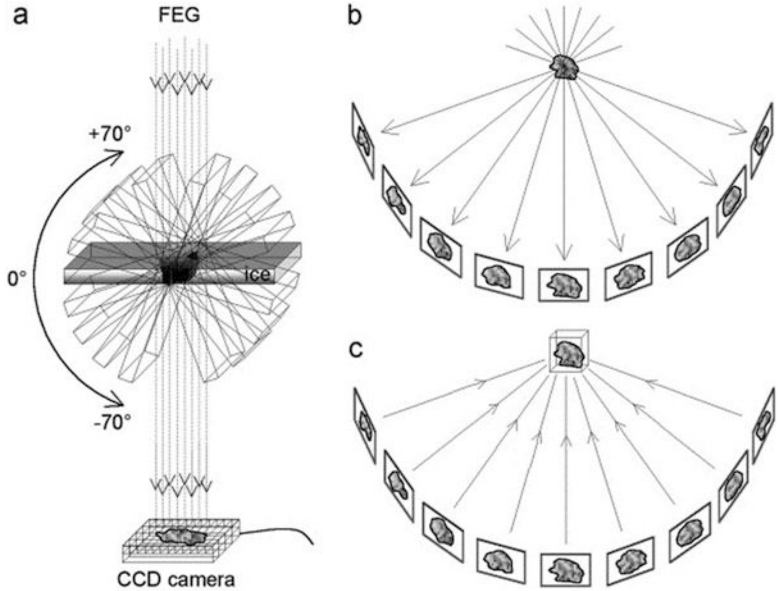


Fig. 3.10 Electron tomography. (**a**) 2D projection images are acquired on a CCD camera while the same specimen area embedded in a vitreous thin layer is tilted incrementally (~1–2°, tilt range ± 70°). (**b**) Diagram illustrating the images projected by a specimen at successive tilt angles. (**c**) After mutual alignment of all these 2D projection images, a density map (tomogram) is reconstructed by a weighted backprojection procedure (Adapted from [47]. Courtesy of W. Baumeister)

specimen is tilted around an axis perpendicular to the electron beam inside the microscope. This type of cryo-EM is termed cryo-ET. Tilting of the specimen provides images of the same field of view (typically more than 100), for example, over a range of ±70° at steps of 1° or 2° (limited by the design of specimen holders and because specimen thickness increases very rapidly at high tilt angles) (Fig. 3.10).

The 3D structure, or tomogram, is calculated as a back-projection in real space; assembly of correctly aligned projections is reverse-projected into 3D space. Cryo-ET has allowed the study of viruses not only in their native physiological state, but also in their natural cell environment (see Chap. 14), thereby providing a bridge between structural studies at the molecular and cellular levels (reviewed in [23, 48, 49], and references therein).

To obtain a detailed, undistorted reconstruction, the tilt series must cover as wide an angular range as possible in as many increments as possible. Tomogram resolution is limited because the range of angles sampled is incomplete and there are no data in some parts of the 3D FT (the missing wedge); in addition, data are missing in the spaces between sampled planes. At the same time, the electron dose must be kept as low as possible to avoid radiation damage. In principle, a total dose (~100 $e^-/Å^2$) is fractioned over the entire tilt series (~1 $e^-/Å^2$ per image), but the resulting projections are very noisy, making their alignment and CTF correction unreliable. This dose represents less than 5 % of the exposure normally used for single-particle cryo-EM. To increase contrast, images are taken at high defocus values (at least 3–10 μm; compared to 0.5–3 μm for cryo-EM), and resolution is somewhat lower than the first zero of the CTF (in the 4 nm range), as restoration of the true projection is difficult with these low signal-to-noise ratio micrographs. These limitations can be overcome with a double-tilt series; two single-tilt series of the objects are recorded, during which the specimen is rotated by 90° around the beam direction after the first series. The resolution problem is also reduced, if applicable, by image averaging of selected regions (homogeneous subcomponents such as glycoproteins or capsids) of the polymorphic (heterogeneous) viruses.

A cryo-ET setup consists of a high voltage cryo-electron microscope equipped with a FEG, a CCD camera and possibly an energy filter. Automation of the acquisition process is essential for acquiring tomographic data sets in low-dose conditions. High voltage electrons can penetrate thicker samples, which enables imaging of cells or organelles. The interaction of electrons with thick (>200 nm) specimens produces many inelastic scattering events, leading to strong blurring. Energy filters allow imaging of thick specimens and are indispensable for tomography of whole ice-embedded cells or other objects of 0.5–1 μm thickness. Cryo-sectioning is a solution for very thick specimens frozen at high pressure. Without sectioning, only the peripheral areas of flat cells are accessible to cryo-ET.

3DR of an object from its projections involves two steps; first, the micrographs must be aligned to a common coordinate system, and second, the aligned micrographs are merged into a tomogram. Fiducial markers added to the specimen (typically colloidal gold), which appear as high-density dots in the images, are commonly used for alignment, although this can also be done without markers. The most common reconstruction method in cryo-ET is the weighted back-projection, to form a 3DR. Because of uneven sampling in the Fourier space, the low frequencies are artificially enhanced, which requires weighting of projections prior to back-projection to obtain a reliable reconstruction.

The interpretation of tomograms at the ultrastructural level requires decomposition of a tomogram into its structural components, *e.g*., the segmentation of defined components such as membranes and glycoproteins. Manual assignment of features is commonly used. In addition, to increase the signal-to-noise ratio, so-called denoising algorithms have been developed for analysis and three-dimensional visualization.

Cryo-ET maps of viral complexes at 20–40 Å resolution can be obtained by 3D averaging of different subtomograms from the same particle. Of course, single-particle cryo-EM provides higher resolution maps with an unbiased starting model. Cryo-ET can provide reliable starting models that are then refined by single-particle analysis. In addition to 3D alignments, processing subvolumes can require classification analysis.

3.8.2 Structure of Pleomorphic Viruses and Capture of Viral Life Cycles Using Cryo-ET

Cryo-ET has been applied to the visualization of the 3D structures of non-isometric virus particles (reviewed in [50, 51], and references therein); for example, it has been used to study the architecture of influenza virus, herpesvirus, immunodeficiency viruses and vaccinia virus, among other enveloped viruses (Fig. 3.11a).

Influenza virus has two types of protein spikes emerging from its envelope, hemagglutinin (HA) and neuraminidase (NA); there are about 400 densely-packed spikes (85 % HA, 15 % NA) and NA spikes tend to cluster; in addition, HA molecules show conformational changes after fusion with liposomes. Cryo-ET of intact virions of herpes simplex virus 1 (HSV-1) has provided new information about the pleomorphic structures crucial for infection of host cells: tegument (an amorphous protein layer that includes host structures such as actin filaments) and an envelope coated with glycoprotein spikes. The icosahedral $T = 16$ capsid enclosing the viral DNA adopts an eccentric position within the approximately spherical virus. The capsid images were extracted from tomograms, averaged, and compared to the nontegumented capsid, revealing capsid-tegument contacts. The entry portal (a specialized DNA packaging vertex) in different herpesvirus capsids was also visualized in one of the 12 five-fold vertices. An equivalent cryo-ET study to detect asymmetric components inside icosahedral viruses was performed with poliovirus, a ~30 nm icosahedral virus that lacks an envelope; isolate particles were imaged releasing their genome, and RNA exited through channels near the two-fold axes. Cryo-ET of the vaccinia virus identified a complex shell consisting of different layers, in which the inner core membrane incorporates pore-like formations. Envelope glycoproteins (env, gp120) of simian and human immunodeficiency viruses (SIV and HIV) mediate binding to the receptor CD4 to initiate infection; analysis by cryo-ET showed that env trimers change conformation in order to interact with the cell receptor.

Beyond the study of purified viruses, cryo-ET extends its scope to different stages of the viral life cycle and cell pathogenesis (cell-associated state) [57, 58]. Until recently, this information was derived mainly from EM of thin sections of plastic-embedded virus-infected cells; the specimen is chemically fixed, dehydrated in organic solvents, embedded in a resin or plastic, and sectioned. The slices are stained to increase contrast. For cryo-ET, the slices are usually 200–500 nm thick. This differs from the thin (20–100 nm) sections usually used in serial section reconstruction, in which single images of individual thin sections are merged computationally to obtain a 3DR, limiting the resolution in the z-axis to the slice thickness. Specimen thickness is a limiting factor for the preservation of native ultrastructure (unless the high pressure freezing approach is used). Cryo-ET of cells is restricted to unsectioned thin prokaryotic cells and to peripheral thin regions of eukaryotic cells that are flattened, and thus transparent to the electron beam. Analysis of larger cells requires technically challenging cryosectioning.

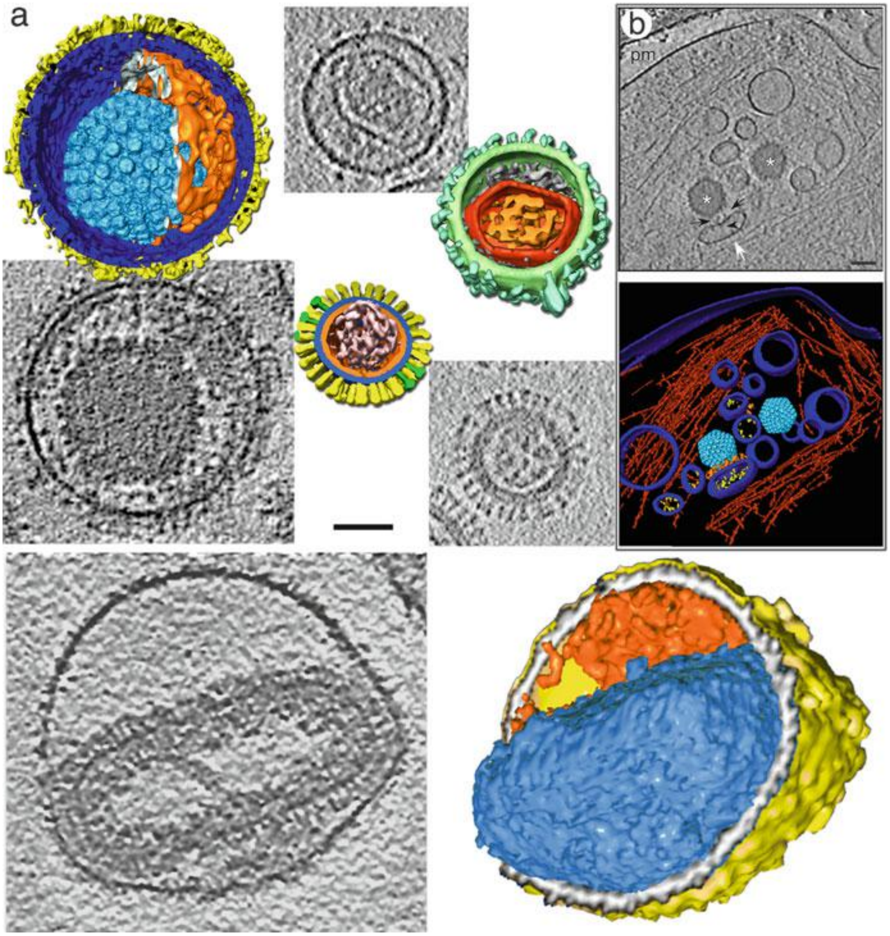


Fig. 3.11 Cryo-ET of pleomorphic viruses and viral life cycles. (**a**) Tomographic reconstructions of HSV-1 (*left*), Rous sarcoma virus (RSV, *right top*), influenza A virus (*right bottom*) and vaccinia virus (VV, *bottom*). Surface-shaded representations of the tomograms are cut to see the virion interior, and a slice of each tomogram is shown. HSV-1 has a T = 16 capsid (*light blue*) buried in the proteinaceous tegument (*orange*), surrounded by a membrane (*dark blue*) studded with glycoprotein spikes (*yellow*) (Adapted from [52]). The RSV virion has a polyhedral capsid (*red*) with an internal density that corresponds to the viral genome (*orange*), surrounded by the matrix protein layer (*grey*); the outermost membrane with the Env spikes is *green* (Adapted from [53]). The genome of influenza A virus is organized as ribonucleoprotein complexes (*pink*); the matrix protein layer (*orange*) is resolved from the lipid bilayer (*blue*); the glycoproteins HA (*yellow*) and NA (*green*) are distinguishable (Adapted from [54]; all three figures courtesy of A.C. Steven). VV mature virion comprises an internal core (*blue*) and the lateral body (*orange*) surrounded by an envelope (Adapted from [55]; courtesy of J.L. Carrascosa). Bar, 50 nm. (**b**) Secondary envelopment of HSV-1 capsids at axon terminals. Slice of the tomogram (*top*) of an intact axon terminal (*asterisks*, capsids; *white arrow*, enveloping vesicle; *arrowhead*, glycoproteins; *black arrows*, tegument; *pm*, plasma membrane). Bar, 100 nm. Surface-shaded representation of the tomogram (*bottom*); color code as in (a); actin is in *red* (Adapted from [56], doi:10.1371/journal.ppat.1002406)

Cryo-ET of the ε15 bacteriophage infecting its bacterial host has allowed visualization of structural changes in the tail machinery, including a tubular density that spans the periplasmic space of the bacterial wall that is the conduit for viral DNA entry into the cell. HSV-1 virions entering synaptosomes (subcellular structures) were caught in action in 3D using cryo-ET; the viral envelope fuses directly with the host membrane, followed by detegumentation. Cryo-ET of frozen-hydrated neurons showed that most egressing capsids were transported independently of the viral envelope; this indicates that progeny herpes viruses are transported along axons as subassemblies and not as complete virions within transport vesicles. Secondary envelopment takes place in axon terminals (Fig. 3.11b). The life cycle of murine gammaherpesvirus, including viral attachment, entry, assembly in virus-induced nuclear inclusion bodies and egress, has been also tracked using cryo-ET. Cryo-ET studies of intact plunge-frozen human cells showed that, during HIV assembly in the native budding sites, the polyprotein Gag lattice is indistinguishable from that of the released immature virion; loss of control of proteolytic maturation led to the formation of aberrant particles.

3.9 Understanding Viruses: Some Major Contributions of Electron Microscopy and Tomography

The first icosahedral virus reconstructions from EM images were made in 1970 from just six negatively stained TBSV (tomato bushy stunt virus) particles at 28 Å resolution, and two of human papillomaviruses at 60 Å [59]. Based on numerous structural virology studies, outstanding progress has been made in the last 40 years. EM imaging has become an indispensable tool in structural virology, and determination of structure is the key to understanding function.

Single-particle cryo-EM has allowed routine analysis of many viral macromolecular complexes in native environments at subnanometer resolutions, in some instances to near-atomic resolution. Cryo-EM is applicable to capsids, genome packaging motors, bacteriophage tails, and other asymmetric supramolecular assemblies (for example, viral polymerase complexes and filamentous ribonucleoprotein complexes). Cryo-EM is best suited to the study of viral complexes too flexible or too scarce to be crystallized for X-ray diffraction studies, or too large for NMR spectroscopy analysis. The most notable feature of cryo-EM is the ability to determine virus structure in its native environment, which permits structural studies of metastable and transient functional intermediate states. The resolution of multiple states will be central to the discovery of molecular mechanisms. 3DR of icosahedral viruses at near-atomic resolution show that atomic models can be derived from cryo-EM images, yielding valuable new information about virus life cycles.

Cryo-ET allows the study of heterogeneous samples such as pleomorphic viruses, and analysis of virus structures in the cellular context, a frontier in structural virology. Analysis of virus-cell interactions will contribute to the development of strategies to combat a large variety of human, animal and plant viral diseases.

3.10 Perspectives

Although the molecular targets for X-ray crystallography and cryo-EM have begun to overlap as these techniques have been perfected, the synergy between them is clear (see also Chap. 7). Cryo-EM and cryo-ET will continue to provide insights into the structure of large, flexible, biologically relevant macromolecular viral assemblies in increasing, unprecedented detail. Whereas single-particle cryo-EM is a well-established and popular tool for structure analysis, with some anticipated improvements in instrumentation and software, cryo-ET is expected to evolve rapidly in coming years. Introduction of direct electron detectors instead of CCD will improve signal transfer efficiency in cryo-imaging. Likewise, the phase plate cryo-electron microscope has shown promising results in acquiring high-contrast, close-to-focus images with high-resolution information. The need remains for better computational methods for alignment, classification and CTF correction of low-contrast images.

Acknowledgements I thank Daniel Luque for stimulating discussions, José L. Carrascosa and José M. Valpuesta for continuous support, comments and careful reading of the manuscript, Alasdair C. Steven and Benes L. Trus for advice and encouragement and Catherine Mark for editorial help. I am indebted to current and former members of my group (Irene Saugar, Daniel Luque, Nerea Irigoyen, Elena Pascual, Josué Gómez-Blanco, Mariana Castrillo, Ana Correia and Carlos Pérez) and other colleagues for their hard work, skills and enthusiasm that made work possible and enjoyable. This work was supported by grant BFU2011-25902 from the Spanish Ministry of Science and Innovation.

References and Further Reading

1. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2007) Molecular biology of the cell. Garland Science, New York
2. Nogales E, Grigorieff N (2001) Molecular machines: putting the pieces together. J Cell Biol 152:F1–F10
3. Steven A, Belnap D (2005) Electron microscopy and image processing: an essential tool for structural analysis of macromolecules. Curr Protoc Protein Sci Chapter 17:Unit 17 12
4. Klug A (2010) From virus structure to chromatin: X-ray diffraction to three-dimensional electron microscopy. Annu Rev Biochem 79:1–35
5. Harrison SC (2004) Whither structural biology? Nat Struct Mol Biol 11:12–15
6. Baumeister W, Steven AC (2000) Macromolecular electron microscopy in the era of structural genomics. Trends Biochem Sci 25:624–631
7. Steven AC, Baumeister W (2008) The future is hybrid. J Struct Biol 163:186–195

8. Lasker K, Phillips JL, Russel D, Velazquez-Muriel J, Schneidman-Duhovny D, Tjioe E, Webb B, Schlessinger A, Sali A (2010) Integrative structure modeling of macromolecular assemblies from proteomics data. Mol Cell Proteomics 9:1689–1702
9. Goldsmith CS, Miller SE (2009) Modern uses of electron microscopy for detection of viruses. Clin Microbiol Rev 22:552–563
10. Bozzola JJ, Russell LD (1999) Electron microscopy. Principles and techniques for biologist. Jones and Bartlett Publishers, Boston
11. Dubochet J, Adrian M, Chang JJ, Homo JC, Lepault J, McDowall AW, Schultz P (1988) Cryo-electron microscopy of vitrified specimens. Q Rev Biophys 21:129–228
12. Amos LA, Henderson R, Unwin PN (1982) Three-dimensional structure determination by electron microscopy of two-dimensional crystals. Prog Biophys Mol Biol 39:183–231
13. Henderson R (2004) Realizing the potential of electron cryo-microscopy. Q Rev Biophys 37:3–13
14. Harris JR (1997) Negative staining and cryoelectron microscopy: the thin film techniques. BIOS Scientific Publishers Ltd., Oxford
15. Wild P (2008) Electron microscopy of viruses and virus-cell interactions. Methods Cell Biol 88:497–524
16. Leiman PG, Kanamaru S, Mesyanzhinov VV, Arisaka F, Rossmann MG (2003) Structure and morphogenesis of bacteriophage T4. Cell Mol Life Sci 60:2356–2370
17. Cavalier A, Spehner D, Humbel BM (eds) (2008) Handbook of cryo-preparation methods for electron microscopy. CRC Press, London
18. Baker TS, Olson NH, Fuller SD (1999) Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. Microbiol Mol Biol Rev 63:862–922
19. De Carlo S, Stark H (2010) Cryonegative staining of macromolecular assemblies. Methods Enzymol 481:127–145
20. Crowther RA (2008) The Leeuwenhoek lecture 2006. Microscopy goes cold: frozen viruses reveal their structural secrets. Philos Trans R Soc Lond B Biol Sci 363:2441–2451
21. De Rosier DJ, Klug A (1968) Reconstruction of three dimensional structures from electron micrographs. Nature 217:130–134
22. Moody MF (1990) Image analysis of electron micrographs. In: Hawkes PW, Valdré U (eds) Biophysical electron microscopy. Basis concepts and modern techniques. Academic Press, London, pp 145–288
23. Rochat RH, Chiu W (2012) Cryo-electron microscopy and tomography of virus particles. In: Egelman EH (ed) Comprehensive biophysics, biophysical techniques for structural characterization of macromolecules, vol 1. Academic Press, Oxford, pp 311–340
24. Chang J, Liu X, Rochat RH, Baker ML, Chiu W (2012) Reconstructing virus structures from nanometer to near-atomic resolutions with cryo-electron microscopy and tomography. Adv Exp Med Biol 726:49–90
25. Grigorieff N, Harrison SC (2011) Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. Curr Opin Struct Biol 21:265–273
26. Steven AC, Heymann JB, Cheng N, Trus BL, Conway JF (2005) Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. Curr Opin Struct Biol 15:227–236
27. Luque D, Gonzalez JM, Garriga D, Ghabrial SA, Havens WM, Trus B, Verdaguer N, Carrascosa JL, Castón JR (2010) The T=1 capsid protein of penicillium chrysogenum virus is formed by a repeated helix-rich core indicative of gene duplication. J Virol 84:7256–7266
28. Johnson JE (2008) Multi-disciplinary studies of viruses: the role of structure in shaping the questions and answers. J Struct Biol 163:246–253
29. Russel D, Lasker K, Phillips J, Schneidman-Duhovny D, Velazquez-Muriel JA, Sali A (2009) The structural dynamics of macromolecular processes. Curr Opin Cell Biol 21:97–108
30. Tang L, Johnson JE (2002) Structural biology of viruses by the combination of electron cryomicroscopy and X-ray crystallography. Biochemistry 41:11517–11524

31. Topf M, Lasker K, Webb B, Wolfson H, Chiu W, Sali A (2008) Protein structure fitting and refinement guided by cryo-EM density. Structure 16:295–307
32. Rossmann MG, Morais MC, Leiman PG, Zhang W (2005) Combining X-ray crystallography and electron microscopy. Structure 13:355–362
33. Topf M, Sali A (2005) Combining electron microscopy and comparative protein structure modeling. Curr Opin Struct Biol 15:578–585
34. Luque D, Saugar I, Rodriguez JF, Verdaguer N, Garriga D, Martin CS, Velazquez-Muriel JA, Trus BL, Carrascosa JL, Castón JR (2007) Infectious bursal disease virus capsid assembly and maturation by structural rearrangements of a transient molecular switch. J Virol 81:6869–6878
35. Zhou ZH (2011) Atomic resolution cryo electron microscopy of macromolecular complexes. Adv Protein Chem Struct Biol 82:1–35
36. Zhou ZH (2008) Towards atomic resolution structural determination by single-particle cryo-electron microscopy. Curr Opin Struct Biol 18:218–228
37. Hryc CF, Chen DH, Chiu W (2011) Near-atomic-resolution cryo-EM for molecular virology. Curr Opin Virol 1:110–117
38. Yu X, Jin L, Zhou ZH (2008) 3.88 A structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. Nature 453:415–419
39. Zhang X, Settembre E, Xu C, Dormitzer PR, Bellamy R, Harrison SC, Grigorieff N (2008) Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. Proc Natl Acad Sci U S A 105:1867–1872
40. Briggs JA, Huiskonen JT, Fernando KV, Gilbert RJ, Scotti P, Butcher SJ, Fuller SD (2005) Classification and three-dimensional reconstruction of unevenly distributed or symmetry mismatched features of icosahedral particles. J Struct Biol 150:332–339
41. Kostyuchenko VA, Chipman PR, Leiman PG, Arisaka F, Mesyanzhinov VV, Rossmann MG (2005) The tail structure of bacteriophage T4 and its mechanism of contraction. Nat Struct Mol Biol 12:810–813
42. Aksyuk AA, Rossmann MG (2011) Bacteriophage assembly. Viruses 3:172–203
43. Stubbs G, Kendall A (2012) Helical viruses. Adv Exp Med Biol 726:631–658
44. Ge P, Tsao J, Schein S, Green TJ, Luo M, Zhou ZH (2010) Cryo-EM model of the bullet-shaped vesicular stomatitis virus. Science 327:689–693
45. Ge P, Zhou ZH (2011) Hydrogen-bonding networks and RNA bases revealed by cryo electron microscopy suggest a triggering mechanism for calcium switches. Proc Natl Acad Sci U S A 108:9637–9642
46. Egelman EH (2007) Single-particle reconstruction from EM images of helical filaments. Curr Opin Struct Biol 17:556–561
47. Grunewald K, Medalia O, Gross A, Steven AC, Baumeister W (2003) Prospects of electron cryotomography to visualize macromolecular complexes inside cellular compartments: implications of crowding. Biophys Chem 100:577–591
48. Lucic V, Forster F, Baumeister W (2005) Structural studies by electron tomography: from cells to molecules. Annu Rev Biochem 74:833–865
49. Cope J, Heumann J, Hoenger A (2011) Cryo-electron tomography for structural characterization of macromolecular complexes. Curr Protoc Protein Sci Chapter 17:Unit17 13
50. Subramaniam S, Bartesaghi A, Liu J, Bennett AE, Sougrat R (2007) Electron tomography of viruses. Curr Opin Struct Biol 17:596–602
51. Grunewald K, Cyrklaff M (2006) Structure of complex viruses and virus-infected cells by electron cryo tomography. Curr Opin Microbiol 9:437–442
52. Grunewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, Baumeister W, Steven AC (2003) Three-dimensional structure of herpes simplex virus from cryo-electron tomography. Science 302:1396–1398
53. Butan C, Winkler DC, Heymann JB, Craven RC, Steven AC (2008) RSV capsid polymorphism correlates with polymerization efficiency and envelope glycoprotein content: implications that nucleation controls morphogenesis. J Mol Biol 376:1168–1181

54. Harris A, Cardone G, Winkler DC, Heymann JB, Brecher M, White JM, Steven AC (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. Proc Natl Acad Sci U S A 103:19123–19127
55. Cyrklaff M, Risco C, Fernandez JJ, Jimenez MV, Esteban M, Baumeister W, Carrascosa JL (2005) Cryo-electron tomography of vaccinia virus. Proc Natl Acad Sci U S A 102:2772–2777
56. Ibiricu I, Huiskonen JT, Dohner K, Bradke F, Sodeik B, Grunewald K (2011) Cryo electron tomography of herpes simplex virus during axonal transport and secondary envelopment in primary neurons. PLoS Pathog 7:e1002406
57. Fu CY, Johnson JE (2011) Viral life cycles captured in three-dimensions with electron microscopy tomography. Curr Opin Virol 1:125–133
58. Iwasaki K, Omura T (2010) Electron tomography of the supramolecular structure of virus-infected cells. Curr Opin Struct Biol 20:632–639
59. Crowther RA, Amos LA, Finch JT, De Rosier DJ, Klug A (1970) Three dimensional reconstructions of spherical viruses by fourier synthesis from electron micrographs. Nature 226:421–425

Further Reading

Agbandje-McKenna M, McKenna R (eds) (2011) Structural virology. RSC Publishing, Cambridge

Jensen GJ (ed) (2010) Cryo-EM, Part A. Sample preparation and data collection. Methods in Enzymology, vol 481; Cryo-EM, part B. 3-D Reconstruction. Methods in Enzymology, vol 482; Cryo-EM, part C. Analysis, interpretation and case studies. Methods in Enzymology, vol 483. Academic Press

Rossmann MG, Rao VB (eds) (2012) Viral molecular machines. Adv Exp Med Biol, vol 726. Springer, New York

Also especially recommended for further reading are references [3, 4, 11, 23, 25, 26, 27, 35] *listed above*.

Chapter 4
X-Ray Crystallography of Viruses

Nuria Verdaguer, Damià Garriga, and Ignacio Fita

Abstract For about 30 years X-ray crystallography has been by far the most powerful approach for determining virus structures at close to atomic resolutions. Information provided by these studies has deeply and extensively enriched and shaped our vision of the virus world. In turn, the ever increasing complexity and size of the virus structures being investigated have constituted a major driving force for methodological and conceptual developments in X-ray macromolecular crystallography. Landmarks of new virus structures determinations, such as the ones from the first animal viruses or from the first membrane-containing viruses, have often been associated to methodological breakthroughs in X-ray crystallography. In this chapter we present the common ground of proteins and virus crystallography with an emphasis in the peculiarities of virus studies. For example, the solution of the phase problem, a central issue in X-ray diffraction, has benefited enormously from the presence of non-crystallographic symmetry in virus crystals.

Keywords Molecular replacement • Non-crystallographic symmetry • Phase problem • Structural virology • Virus capsid • Virus X-ray crystallography

---

N. Verdaguer (✉) • I. Fita (✉)
Institut de Biología Molecular de Barcelona (CSIC),
Parc Científic de Barcelona, c/Baldiri i Reixac 10, 08028 Barcelona, Spain
e-mail: nvmcri@ibmb.csic.es; ifrcri@ibmb.csic.es

D. Garriga
Department of Molecular and Cell Biology, Centro Nacional
de Biotecnología (CSIC), c/Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain

M.G. Mateu (ed.), *Structure and Physics of Viruses: An Integrated Textbook*,
Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_4,

© Springer Science+Business Media Dordrecht 2013

Abbreviations

Ad	Adenovirus
Ad5	Adenovirus type 5
BTV	Bluetongue virus
EM	Electron microscopy
HRV2/HRV14	Human rhinovirus serotypes 2 or 14
MR	Molecular replacement
SBMV	Southern bean mosaic virus
STNV	Satellite tobacco necrosis virus
TBSV	Tomato bushy stunt virus
VLP	Virus-like particle

4.1 Introduction

"Seeing is believing": visualizing the three-dimensional structures of intact virus particles, as well as their constituent proteins and complexes, provides us a tool for a full understanding of their biological function. X-ray crystallography is one of the most powerful approaches for visualizing such macromolecular assemblies at atomic resolution. The first high-resolution pictures of intact virus particles were provided some 30 years ago with the crystal structures of small RNA plant viruses [1–3]. This seminal work was followed by two landmark studies in macromolecular crystallography, the structures of human rhinovirus (Fig. 4.1) and poliovirus [7, 8]. By the end of 1990s, virus crystallography achieved a new breakthrough, the structure of the large core particle (70 nm in diameter) of bluetongue virus (BTV) [9]. The cores of these viruses act as transcriptional machines in infected cells, holding their genomes and transcriptional enzymes within a protective shell so that the viral double-stranded RNA is never revealed to the infected cell, eliminating any chance of cellular response. The BTV structure (Fig. 4.2) showed the assembly of nearly 1,000 protein subunits: the core was made from two principal protein components: a thin skin covering the genome and enzymes, made from 120 copies of a large protein VP3 (triangulation number $T = 2$), which assembles into an icosahedral arrangement *via* conformational switching to yield a pattern of subunits not predicted in the theory of Caspar and Klug (see Chap. 2). This was clothed in 780 copies of the protein VP7 ($T = 13$) in an arrangement that follows the theory of quasi-equivalence.

Enveloped viruses (see Chap. 2) have rarely yielded crystals; however, solubilized protein components of their membranes have been crystallized and studied with great success. Only very regular structures can form single crystals, and in order to study the molecular details of larger and more complex virus

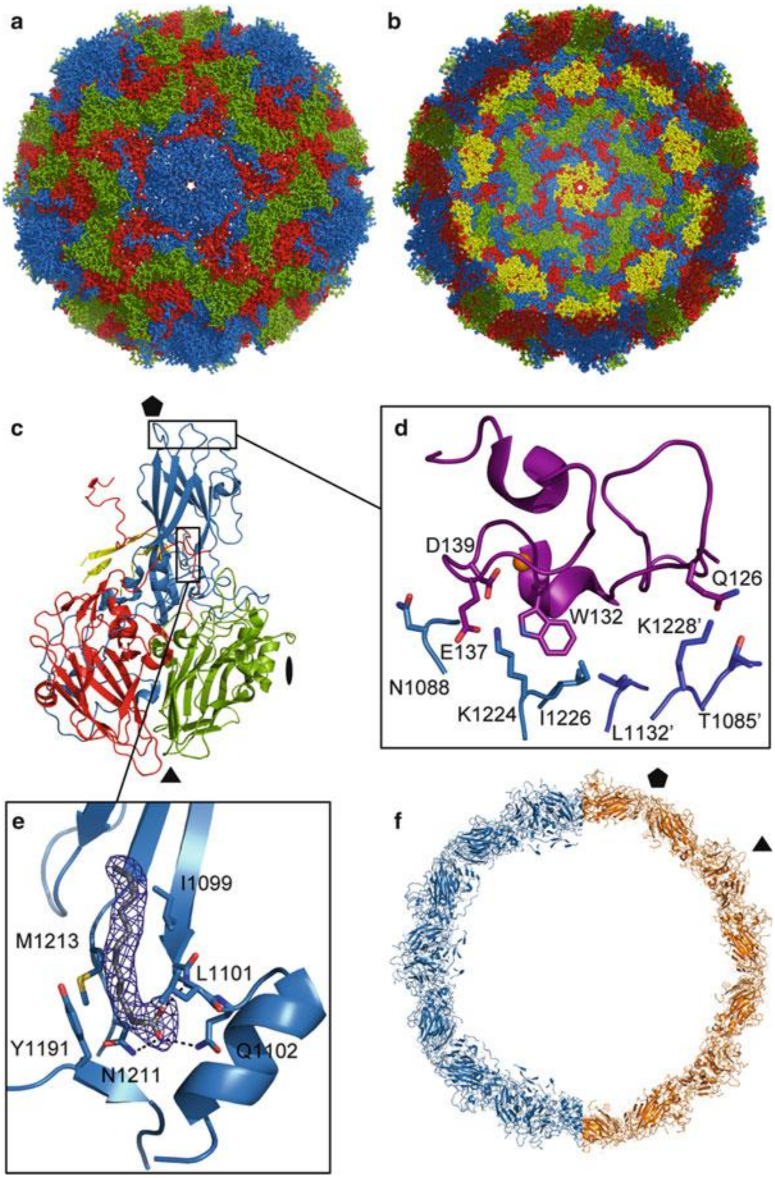


Fig. 4.1 The Human rhinovirus capsid. Representation of the external (**a**) and the internal (**b**) surfaces of the HRV2 capsid (PDB code 1FPN) [4], seen along a fivefold symmetry axis. The capsid, of about 30 nm in diameter, is built by 60 protomers, arranged in a T = 1 icosahedral lattice. Each protomer is composed by four different capsid proteins: VP1 (*colored in blue*), VP2 (*green*), VP3 (*red*) and VP4 (*yellow*). VP1, VP2, and VP3 proteins constitute the capsid shell, whereas VP4 is located at the interior of the capsid, lying on the inner surface, around the fivefold symmetry axes. A prominent feature of the HRV shell is a star-shaped dome on each of the fivefold axes surrounded by a large depression or 'canyon'. (**c**) Ribbon representation of the capsid protomer, colored as in **a** and **b** and with the icosahedral symmetry axes indicated with an *ellipsoid*

particles by X-ray crystallography, it is necessary to dissect them into well defined subunits or substructures. This dissection was originally done with proteases, by disassembly or by isolation of substructures from infected cells. For example, the influenza virus hemagglutinin, the first viral glycoprotein for which atomic details were visualized [10], was obtained from crystals of protein cleaved from the surface of purified virions. In addition, the structure of the adenovirus hexon was obtained from excess unassembled protein derived from adenovirus-infected cells [11]. Since the middle of 1990s, this dissection has more commonly been carried out using recombinant proteins (*i.e.* the gp120 protein from HIV-1; [12]). So far, only one membrane-containing virus X-ray structure has been published: PRD1, possessing an internal membrane, was successfully crystallized, and its structure was solved by X-ray crystallography [13, 14]. Four proteins (MPC, P16, P30 and P31) out of a total of 18 structural proteins identified in PRD1 were resolved in the icosahedrally-ordered capsid. A conformational switch of the MPC N- and C-terminus controls the quasi-equivalent interactions within the pseudo T = 25 lattice.

Today X-ray structural studies of these large assemblies (see also Chaps. 2, 10, 11, 12, 13, 15, 16 and 17) are facilitated by the accumulation of experience in their purification, crystallization and structure determination, as well as by the development of multidisciplinary technologies, including: (i) Crystallization robotics that automate the process; (ii) third generation synchrotron beamlines tailored for data collection from weakly diffracting crystals with large unit cells (see below); and (iii) faster computers and improved software for crystallographic data integration and for structure determination. Current frontiers include defining the architectures of very large assemblies, as exemplified by the 3.5 Å X-ray structure of a complete human adenovirus (Ad) particle [15]. This particle of 150 MDa and an average diameter of 91 nm represents the largest structure solved to date using X-ray

◄

Fig. 4.1 (continued) (twofold axis), a *triangle* (threefold) and a *pentamer* (fivefold). The high resolution crystal structure of a virus is used to gather information for more than the mere spatial disposition of the capsomers; they may also provide the molecular basis of key steps of the virus life cycle. Some examples of this are illustrated in panels (**d**), (**e**) and (**f**), (**d**) Ribbon representation of the structure the HRV2 capsid interacting with the V3 module of VLDL receptor. The cognate receptor for the minor group rhinoviruses (PDB code 1V9U) [5]. One V3 module (*colored in purple*) interacts with two VP1 molecules from neighboring protomers (*colored in different blue tones*). The residues involved in the virus-receptor interaction are represented as sticks and labeled. (**e**) Electron density map (*dark blue wire*) corresponding to the ‘pocket factor’, a cell-derived fatty acid that often occupies the hydrophobic pocket of VP1, at the bottom of the canyon. The VP1 residues interacting with this pocket factor are represented as sticks. The pocket factor can be displaced by a number of antiviral drugs that are thought to rigidify the capsid, preventing the required structural changes needed for viral genome uncoating during cell infection. (**f**) Comparison of the native HRV2 capsid (*blue*) with the 80S subviral B-particle (*orange*), the end-stage of RNA uncoating (PDB code 3TN9) [6]. Only half of a ~80 Å slab of each capsid shells are represented as to illustrate the expansion in the 80S particle with respect to the native virion. The positions of one threefold and one fivefold axis of the icosahedral particle are indicated with a *triangle* and a *pentamer*, respectively

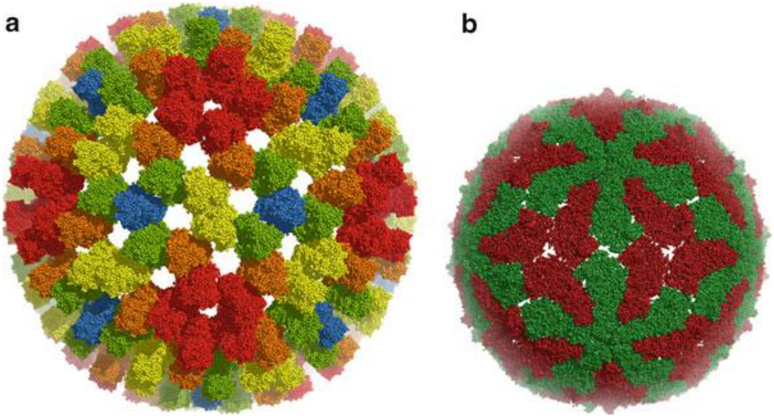


Fig. 4.2 Bluetongue virus core particle structure. Stick representation of the atomic structure of BTV1 virus (PDB code 2BTV) [9], seen along a twofold symmetry axis. The BTV core particle is composed by two differentiated protein shells. The outer shell (**a**) contains 780 copies of VP7 protein, arranged as trimers on a T = 13 quasi-equivalent lattice. In the figure, these VP7 trimers are colored according to their relative position to the symmetry axes (*blue* at the threefold axes, *yellow* at the twofold axes, *red* at the fivefold axes, *green* between the two- and threefold axes trimers, *orange* between the two- and fivefold axes trimers). The inner shell (**b**) consists in a thin subcore generated by 120 copies of VP3 protein organized in a T = 2 lattice. The two structural conformations of the VP3 protein are shown in *dark red* and *dark green*. Both layers of the core are represented at scale

crystallography (Fig. 4.3). Furthermore, comparative analyses of multiple structures of these assemblies, captured in different conformational states, greatly help to precisely define how they function and are regulated. The growing number of structures of virus-receptor complexes [16] (Fig. 4.1b) (see Chap. 15), mature and immature capsid proteins [17] (see Chap. 13), membrane fusion intermediates [18, 19] (see Chap. 16) or RNA uncoating intermediates [6, 20] (Fig. 4.1f) (see Chap. 15) illustrate some examples. It is now clear that solving the structure of a new virus at high resolution typically signifies the beginning, not the end, of a modern structural virology study.

4.2 Basic Concepts and General Experimental Design

4.2.1 Basic Concepts

We cannot form images of things that are much smaller than the wavelength of light we are using. X-rays are in the order of atom diameters and bond lengths, allowing atoms in a molecule or molecular complex to be individually resolved. However, X-ray scattering from a single molecule would be unimaginably weak. In contrast, crystals arrange a huge number of molecules (*e.g.*, virus particles) in the same

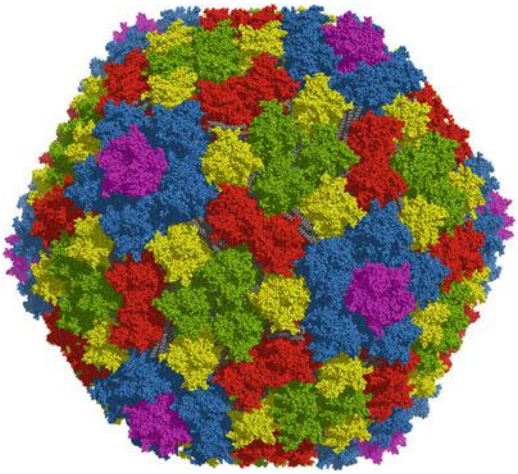


Fig. 4.3 Structure of Human adenovirus. Stick representation of the crystal structure of Ad5 (PDB code 1VSZ) [15]. The adenovirus capsid is mainly composed of hexon protein, with 720 subunits arranged as 240 trimers on a pseudo-T = 25 icosahedral lattice (12 hexon trimers per icosahedral facet). In the figure, the hexon trimers are colored according to their relative position to the symmetry axes: *cyan* around the fivefold axes, *green* around the threefold axes, *red* around the twofold axes and *yellow* between the two- and fivefold axes. Each of the 20 vertices of the capsid is occupied by five copies of the penton base protein (*colored in pink*), in which the Ad fiber shafts are inserted (not shown in the figure). Additional proteins (proteins IIIa, VI, VIII and IX, the so-called cement proteins, represented as *grey balls* in the figure) mediate and stabilize the interactions between neighboring hexon trimers

orientation and scattered waves add up in phase, increasing the signal to a level which can be measured.

Figure 4.4 shows two examples of packing of a virus in two different crystal types. The lattice is defined as a construct dividing space into regular translationally periodic units. The combination of the unit lattice with molecular motifs (*e.g.* a virus particle) generates the unit cell of the crystal that is built from translationally stacked unit cells. In addition, the molecules contained in each unit cell can be related by internal symmetry, generating a unit cell packed with multiple symmetry equivalent copies of the molecule (Fig. 4.4b). The number of different ways to pack molecules in space is limited. For asymmetric chiral molecules (such as proteins), there are only 65 discrete and distinct ways to assemble three-dimensional periodic crystals through combination of translational and rotational symmetry. These 65 types of arrangements form 65 space groups. Their symmetry properties and the rules for constructing each crystal structure are described in the International Tables for Crystallography, Volume A [21].

The asymmetric unit of a unit cell contains all the necessary information to generate the complete unit cell of a crystal structure by applying its symmetry operations to the asymmetric unit. Besides the strict crystallographic symmetry between molecules, additional non-crystallographic symmetry can locally exist

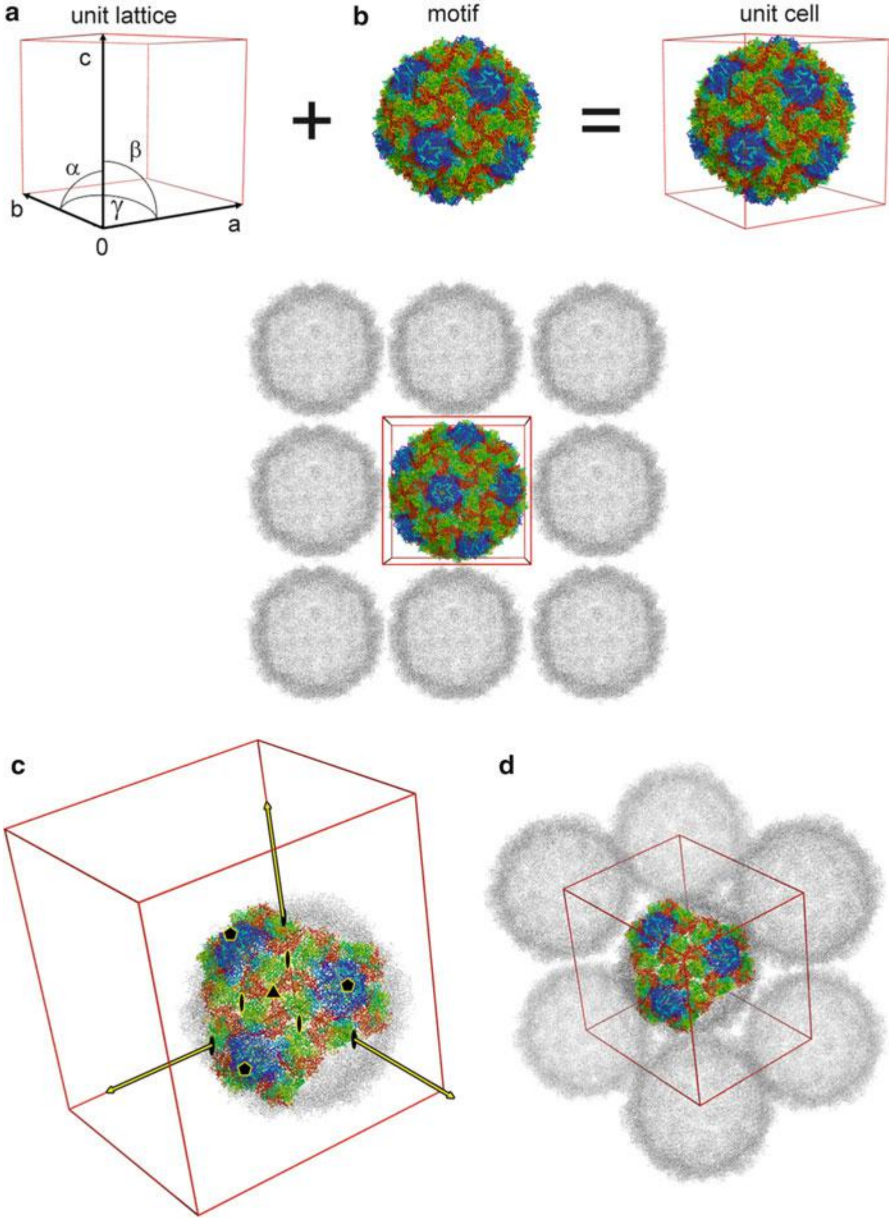


Fig. 4.4 Assembly of virus crystals. (**a**) The crystal unit cell is built by a motif (in this case, a whole virus) that fills the unit lattice. The unit lattice is defined by three basic vectors [**a**,**b**,**c**] that extend from a common origin [0] in a right handed system. The angle between **a** and **b** is γ, the angle between **b** and **c** is α and the angle between **a** and **c** is β. The stacking of multiple unit cells will produce a crystal. (**b**) Crystal packing of an icosahedral virus, in which the whole virus is the asymmetric unit (*colored*). The other viruses of the crystal (*grey*) can be obtained by translation of the unit cell. For clarity, only one two-dimensional layer of the crystal is shown. (**c**) An icosahedral capsid has 12 fivefold symmetry axes (indicated with pentamers, in the figure), 20 threefold

between molecules in the asymmetric unit. In virus crystals, the molecular symmetry results in most cases from the combination of local and crystallographic symmetry (Fig. 4.4c).

In a basic diffraction experiment, a single molecular crystal is placed into a finely focused X-ray beam and the diffraction images are recorded in a detector. The electron density representing the atomic structure of the molecules in the crystal is reconstructed by Fourier Methods from the diffraction data and an atomic model of the structure is built into the electron density (Fig. 4.5). It is important to note that the diffraction experiment is fundamentally different from imaging with an optical microscope. The crucial difference is that visible light scattered from objects can be focused through refractive lenses to create a magnified image of the object. In contrast, as no refractive lenses are available for X-rays, the X-ray diffraction patterns do not provide a direct image of the molecular structure. The electron density of the scattering molecular structure must be reconstructed by Fourier transform techniques. For this type of reconstruction, from reciprocal diffraction space back into the direct molecular space, two terms are required: (i) the structure factor amplitudes, directly accessible in the form of the square root of the intensities of each measured diffraction spot or reflection and (ii) the relative phase angle of each reflection. These phase angles are not directly accessible and must be supplied by additional phasing experiments. The absence of directly accessible phases constitutes the so-called phase problem in macromolecular crystallography (see Sect. 4.6).

At the end of the chapter (Section on *Further reading*) three reference books are provided for an extended description of the general principles and techniques in X-ray crystallography of biological molecules. The first two books are aimed at non-practitioners; the last one contains a more in-depth treatment.

4.2.2 General Experimental Design

The general approach to structure determination of viral particles (and other biomolecules and molecular complexes) by X-ray crystallography is schematically shown in Fig. 4.5. The procedure involves the following basic steps: (i) preparation and purification of virus particle samples; (ii) crystallization and mounting the virus

Fig. 4.4 (continued) symmetry axes (*triangles*) and 30 twofold symmetry axes (*ellipsoids*), relating the quasi-identical subunits that build up the virus capsid. When this virus is fitted in a crystal lattice, it may happen that some of the icosahedral symmetry axes coincide with the symmetry axes defined by the unit cell (*yellow arrows*); in this case, only part of the virus forms part of the crystal asymmetric unit (*colored*). The rest of the virus particle (*grey*) can be obtained applying the crystal symmetry operators to the asymmetric unit. (**d**) The crystal packing of an icosahedral virus, in which one forth of the virus forms the crystal asymmetric unit

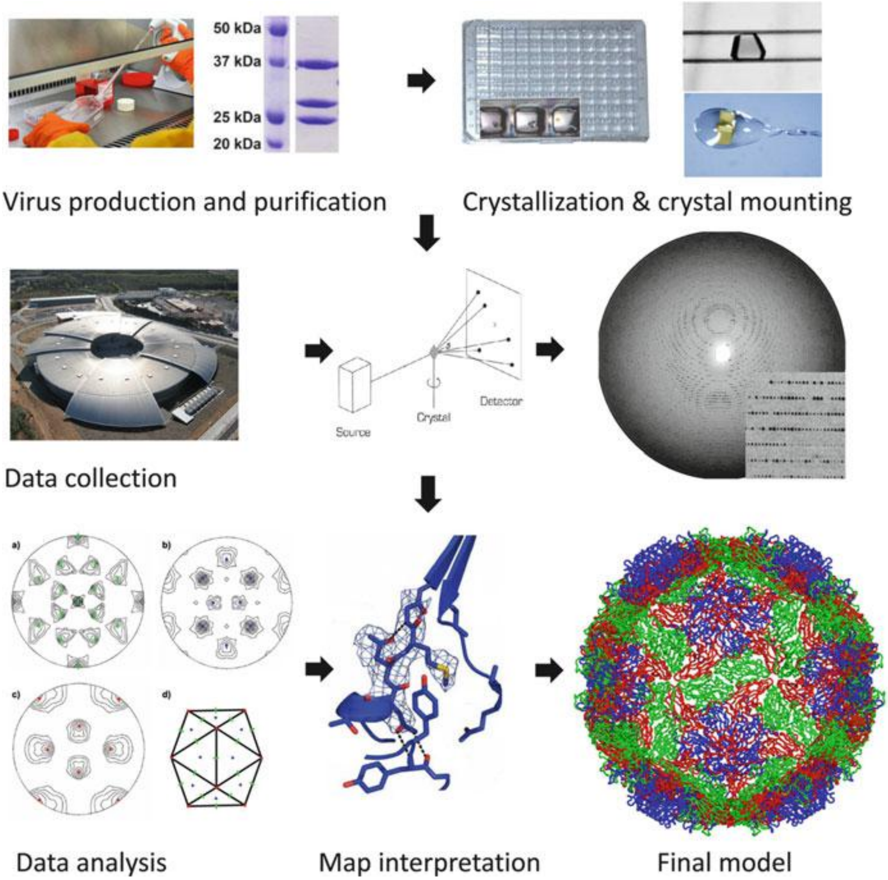


Fig. 4.5 Virus structure determination by X-ray crystallography. The solution of a virus structure involves the following steps: (i) preparation and purification of virus samples, (ii) crystallization and mounting the virus crystals, (iii) measurement of the diffraction data, (iv) Phase calculation usually by molecular replacement and, (v) map interpretation and model building

particle crystals; (iii) measurement of the diffraction data; (iv) phase calculation, usually by molecular replacement (MR); and (v) map interpretation and model building. These steps will be briefly described in the following sections in this chapter (Sects. 4.3, 4.4, 4.5, 4.6 and 4.7).

The first limitation for an X-ray experiment is the requirement of well-ordered three-dimensional crystals. This implies that we cannot determine the structure of a macromolecular complex that does not have a single, well-defined, structure or that is difficult to obtain in a pure and stable state. Furthermore, the structure determination of an intact virus particle entails a significant scaling up in most respects compared to a traditional crystallographic analysis of a single protein. At a fixed resolution, the number of reflection spots to be measured is proportional to the volume of the asymmetric unit of the crystal. For a virus particle, there are several hundred times as many data to be measured than for an isolated protein. Not only are there far more X-ray reflections to be measured, but these reflections will each be, on average, several hundred times weaker. Therefore, virus crystallography makes considerable demands on the qualities of X-ray sources, the performance and sensitivity of X-ray detectors and computer hardware and software [22].

Having a well-diffracting single crystal is a first step, though not necessarily sufficient, to solve the structure. As mentioned in Sect. 4.2.1, X-ray detectors are able to measure the position and intensity of reflections without recording the phase information. Therefore, in order to compute the Fourier transform, to go from the diffraction pattern to the electron density map of the structure, we need to obtain information about the phases. In virus crystallography, molecular replacement is now the method most frequently used to estimate the phases, needed to solve crystal structures (see Sect. 4.6). Spherical viruses are made from protein shells with icosahedral (532) symmetry (Chap. 2). According to the theory of quasi-equivalence, the icosahedral shell is composed of 60 × T chemically identical subunits which can then, in principle, make similar interactions to form the capsid (see Chap. 2). The presence of icosahedral symmetry leads to the presence of non-crystallographic redundancy (with a minimum of fivefold non-crystallographic symmetry, for those viruses that crystallize with the viral capsid lying on a point of 23 crystallographic symmetry). This redundancy is of enormous importance in a crystallographic analysis, providing extremely powerful constraints that facilitate many aspects of the analysis [22] (Sect. 4.6).

4.3 Production and Purification of Viral Particles and Proteins for Structural Studies

4.3.1 Growth and Purification of Intact Viruses and Viral Capsids

The first step in the route to solving the structure of a viral particle by X-ray crystallography (Fig. 4.5), and a prerequisite for crystallization of the particle, is the availability of sufficient amounts of the sample in purified form. For virus particles, the purification from native sources often relies on classical biochemical protein purification procedures. Virus yields vary and for animal viruses grown in a monolayer cell culture, a vast number of roller bottles or large culture plates may be required. A key step in purification is usually ultracentrifugation using sucrose and/or CsCl gradients. In addition, chromatographic FPLC techniques are now available to separate high-molecular-weight species. Many virus particles may be readily concentrated using 100-kDa microconcentrators or precipitated by ultracentrifugation. Standard spectrophotometric concentration measurements should be corrected for the nucleic acid content; an approximate formula is: Virus particle concentration $= (1.55\ A_{280}–0.76\ A_{260}) \times$ dilution [23].

With a macromolecular assembly, there is a possibility of crystallizing a substructure such as an assembly intermediate (see Chaps. 10 and 11). Homogeneity and integrity of the sample used for crystallization and of the crystallized molecule should be carefully analyzed. The crystallization of intact virus particles can be verified using SDS-PAGE to demonstrate the presence of the correct complement of proteins, using tests to show recovery of infectivity from dissolved crystals and by electron microscopic analysis of the crystals.

4.3.2 Expression and Purification of Recombinant Viral Proteins

Low yield virus production and sample heterogeneity sometimes impede purification of native virions with the quality and in the quantity required for structural studies. The developments in molecular biology techniques that allow cloning the viral genomes or capsid protein genes for overexpression and assembly of recombinant virus-like particles (VLPs) in heterologous systems are instrumental for the structural studies of these~~s~~ viral assemblies. Recombinant techniques can further provide the means to modify complexes at the gene level, resulting in alteration of the protein complex subunits to meet the high quality requirements for structural analyses, as well as for site-directed structure-function studies. A successful example in the application of these techniques was achieved in the X-ray structure determination of human adenovirus type 5 (Ad5). High quality crystals were obtained of a genetically modified virus, in which short penton fibers from another serotype have replaced those of the wild-type Ad5 [15].

Molecular cloning, protein expression and crystallization are the major preparatory steps of a structural study. All three processes are experimentally driven and its outcome cannot be predicted. A parallel approach pursuing multiple variants of the protein target significantly increases the chances of success. A number of high throughput initiatives have been developed using affordable methods and equipment to automate molecular cloning of expression constructs, recombinant expression screenings and purification. The advantages of automation are obvious: procedures can be carried out in parallel and scaled-up accordingly to process large amounts of samples at reasonable cost. The methodology is also available for overexpression of multicomponent protein complexes, either in prokatyotic or eucaryotic cells [24].

4.4 Crystallization

The biochemical and biophysical process of crystallization is still a largely empirical process. Crystals grow from an aqueous protein solution, when the solution is brought into supersaturation (Fig. 4.6). Supersaturation is achieved by varying concentrations of precipitant, protein, additives, and other parameters affecting solubility, such as pH and temperature. Crystallization proceeds in two phases,

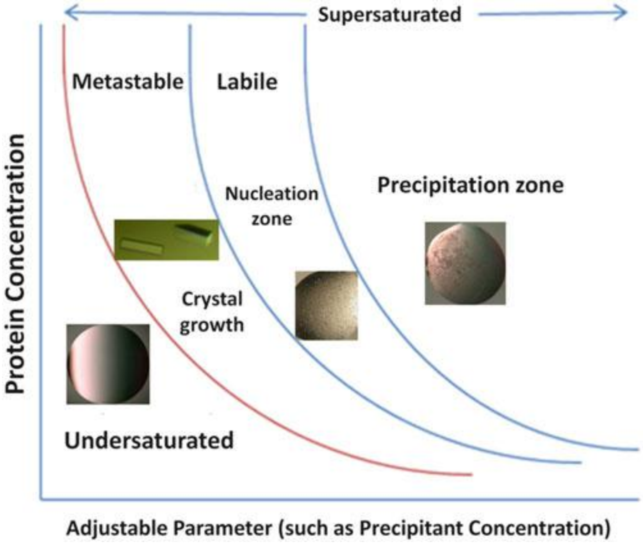


Fig. 4.6 Crystallization Phase Diagram. A schematic drawing of a protein crystallization phase diagram based on two of the most commonly varied parameters, protein and precipitant concentrations. The figure schematically illustrates four areas: (i) an area of very high supersaturation where the protein will precipitate; (ii) an area of moderate supersaturation where spontaneous nucleation takes place; (iii) an area of lower supersaturation, just below the nucleation zone where crystals are stable and may grow. This area, known as metastable zone, is thought to contain the best conditions for growth of large and well ordered crystal: (iv) an undersaturated area where the protein is fully dissolved and will never crystallize

nucleation and growth. Nucleation is the conception stage of crystallization during which a molecular aggregate of critical size forms, allowing for subsequent crystal growth. Once the crystal nucleus has formed, crystal growth follows spontaneously (see [25] for a complete view).

The crystallization process can be effectively illustrated using a phase diagram (Fig. 4.6). The phase diagram indicates which state is stable under different chemical and physical conditions. It provides means of quantifying the influence of parameters such as the concentrations of protein, precipitants, additives, etc. on the production of crystals, allowing the design of crystal optimization experiments. Crystallization conditions for viruses and other complex assemblies are similar to those of single proteins.

4.4.1 *Classical Methods*

There are four major techniques for growing crystals of biological macromolecules. These techniques have been in use for over 40 years and include: batch crystallization, dialysis, liquid-liquid diffusion (also known as free interface diffusion), and vapor diffusion [25]. The four methods work well, the later being the most extended technique for viral particle crystallization. The typical time of crystal growth, using any one of these techniques, is 1-3 weeks. However the time for obtaining crystals can vary from several hours to as long as a year. Standard virus concentrations for the first crystallization trials range around 10–20 mg/ml; however, if the solubility of the virus is limited to less than this, it is still worth trying at whatever concentration can be achieved.

The method of vapor diffusion begins with an aqueous crystallization drop containing the macromolecule and the crystallization agents at a concentration lower that that required for the formation of crystals. This drop (typically from 0.5 μl to 10 μl) is equilibrated against a reservoir (0.5–25 ml) containing a solution with crystallization agents at different concentrations than in the crystallization drop. Usually, the crystallization cocktail is used as reservoir solution but it is not a requirement, the reservoir solution and the crystallization cocktail can be different chemical species [26]. Equilibration proceeds by diffusion of the volatile species, typically water, until vapor-pressure equilibrium is reached between the experiment drop and the reservoir. As a consequence of the equilibration process, there is an increase in the relative concentration of solutes in the crystallization drop. By decreasing the amount of water in the drop, the reservoir acts to gradually concentrate the protein and other chemicals in the crystallization drop. This gradual change and increase in the concentration may lead to supersaturation of the macromolecule and subsequent crystallization. Three commonly used vapor diffusion methods are; hanging drop, where the drop is suspended over a reservoir from the underside of a hydrophobic surface; sitting drop, where the drops sits in a small concave surface; and sandwich drop, where the drop makes contact with two narrowly separated surfaces.

The first step of finding initial crystallization conditions for a new macromolecule is to set up screening trials, exposing the macromolecule to numerous different chemical agents in different drops in order to find ‘hits’ or ‘leads’ that point to conditions that may be conducive to crystallization. Crystals, crystalline precipitate and phase separation are all considered leads that are worth pursuing.

One of the main phenomena preventing crystallization is particle aggregation. Electron microscopy is a useful tool for checking this, in addition to methods such as dynamic light scattering. If aggregation proves to be a problem, it may be relieved by the addition of detergents. Other useful additives include divalent cations known to interact with the capsid. With viruses, highly ordered crystals are often obtained once the optimum crystallization conditions are found. This may be attributable to their isometric nature and reflects the perfection with which these systems assemble. There is a tendency toward high symmetry space groups, often with some of the capsid symmetry axes coinciding with those of the crystal [22].

4.4.2 New Approaches

Today structural studies of large molecular assemblies such as viruses are facilitated by the development of the robotic crystallization technology. The use of robotic nanoliter screenings allows exploring a vast number of different conditions with a small amount of sample [27]. High throughput technologies render the initial crystallization screening process, an often repetitive process utilizing a similar set of cocktail conditions, less labor-intensive and sample-consuming. Many stages of the procedure are automated, starting with the mixing of concentrated chemical stock solutions into a set of crystallization cocktails, to dispensing the macromolecule plus precipitant mixture droplets into plates, sealing the plates, monitoring, scoring the outcomes with the use of image recognition software and even suggesting new experiments based on those outcomes.

4.5 Data Collection and Processing

Once adequate crystals of the viral particle are available, they are mounted in an appropriate holder and irradiated with X-rays (Fig. 4.5). Macromolecular crystals commonly suffer rapid radiation damage during room temperature X-ray data collection. Therefore, data are now routinely collected with the sample held at around 100 K, usually resulting in higher resolution and better quality data. In addition to the cryo-protection techniques, other sophisticated post crystallization treatments, as crystal dehydration, resulted crucial for the structure determination of large macromolecular assemblies.

4.5.1 Cryo-Crystallography

Macromolecular crystallographic data are nowadays routinely collected at cryogenic temperatures (100 K). The key advantage of this technique is the enormous reduction in secondary radiation damage to the sample during X-ray data collection compared to data collection at room temperature. This is because, at cryo-temperatures, most of the free radicals produced by the incident radiation have much lower diffusion rates so they do not travel as far through the crystal, thus avoiding as much damage. The prolonged crystal life-times usually result in better quality and higher resolution data. Nevertheless, even in samples that are cryo-cooled, radiation damage is often observed at undulator synchrotron beamlines. It is a limiting problem in the optimum use of these high brilliant beams, which are routinely attenuated or defocused.

Currently, mounting a crystal for cryo-crystallography involves suspending it by surface tension in cryo-buffer within a thin fiber loop (Fig. 4.5) and then flash-cooling it in gaseous or liquid cryogen. The cryo-buffer is designed to vitrify and prevent crystalline ice formation [28]. Cryo-crystallography experiments are affected by two classes of variables: first, the physical and chemical environment of the crystal in the loop, including choice and concentration of cryo-protectant agent, solvent content of the crystal, size of the crystal and surface area/volume ratio of the crystal; second, the external variables, including the cooling regime (the cryogen used for flash-cooling, the cryogen flow rate and the temperature used during data collection) and the incident beam conditions of dose, dose rate, flux and wavelength.

The most popular cryoprotectant is glycerol at concentrations between 20 % and 25 %. This does not mean that it is the best cryoprotectant; most experiments involve little optimisation of cryoprotocols and glycerol is often just the easiest available solution that will 'do the job'. The optimum scenario, however, is to grow the crystal in a solution that is already cryoprotected.

Due to faster heat transfer, small crystals are generally easier to flash-cool than larger ones. Larger crystals, however, have a greater diffracting power and will yield more useable diffracted photons before reaching their dose limit. There is thus a minimum crystal size that can be expected to give a data set at a specified resolution for a certain unit cell.

4.5.2 Other Post-Crystallization Treatments

Data collection at cryogenic temperatures, which reduce radiation damage and increase crystal lifetime in the X-ray beam, is essential when dealing with crystals of large macromolecular assemblies. However, the development of protocols for cryo-crystallography in large viruses sometimes fails. Even in sufficiently cryoprotected crystals, rapid freezing can introduce lattice disorders that reduce the diffraction limit. This effect, most likely attributable to the unequal spatial contraction of the solvent and the crystal lattice, is more pronounced for larger crystals with high solvent content [29]. Furthermore, to reduce safety hazards, crystals of infectious pathogenic viruses may have to be mounted in sealed capillary tubes for data collection. In these cases, a large number of crystals will have to be examined in order to solve the structure. Alternative post-crystallization treatments may be helpful to reduce the amount of virus crystals needed for structural studies of human pathogens.

Crystal dehydration can lead to remarkable improvements of diffraction quality. The controlled reduction of the solvent content in a crystal decreases the volume of the unit cell, altering in most cases the crystal packing. These rearrangements may reinforce existing and create new crystal contacts and can thereby lead to better ordered and better diffracting crystals. A sophisticated example of this application has been found in the structural analysis of the membrane-containing bacteriophage PRD1. Crystals of PRD1 were very delicate and the mechanical handling involved in preparing them for data collection rendered them useless for this purpose. In this case a method for growing PRD1 crystals in thin-walled quartz capillary tubes by vapor diffusion were developed, which obviated the need for crystal handling and rendered X-ray data collection practical. These crystals grown in capillaries were further treated by the addition of small amounts of a 20 % (w/v) PEG20K solution injected into the capillary prior data collection [30].

The new developments in techniques for measuring diffraction data directly from crystals in their crystallization plates allows to by-pass a number of problems associated with crystal handling and cryo-protection. By eliminating crystal-handling, plate-based data collection in modern synchrotron beamlines provides a safe platform for the analysis of pathogenic particles, as recently reported in the structure solution of bovine enterovirus 2 [31].

4.5.3 Large Unit Cells and Fragile Crystals

Large unit cell dimensions are inherent to crystals of viruses. Therefore, various crystallographic challenges have to be undertaken when starting with data collection; crystals with large unit cells diffract more weakly because fewer unit cells per crystal volume contribute to scattering. Furthermore, as a consequence of long unit cell axes, the distances between reflections in reciprocal space are small, leading to closely spaced diffraction spots on the detector. Modern undulator beamlines in third generation synchrotrons offer the best instrumentation for successful data collection when working with large-unit cells. Their high beam intensity compensates for the weak diffraction while the small beam divergence and low energy bandwidth decreases spot size, reducing the problem of overlapping reflections. The increased surface area of modern detectors allows crystal to detector distances to be increased. This further reduces the problem of overlapping reflections while increasing the signal to noise ratio as the background radiation is inversely proportional to the square of the detector distance. In addition to the best possible instrumentation, an optimized data collection strategy is needed for a successful experiment. Among the general considerations that should be taken into account, the orientation of unit cell axes is of special importance. If one of the unit cell axes is substantially longer than the others it should be aligned with the spindle axis of crystal rotation to avoid overlapping reflections from neighboring lunes [29].

The strong beam intensity and long exposures necessary to collect high-resolution diffraction data from crystals of large macromolecular complexes causes severe radiation damage and frequently limits the amount of data that can be collected from one crystal, even at cryogenic temperatures. In favorable cases only a few crystals are needed to collect useful data (*i.e.* for frozen crystals of a virus belonging to a highly symmetric space groups). However, for more challenging analyses, a large number of crystals can be required (for example, in the studies on BTV and reovirus cores over 1,000 crystals, mounted in sealed capillaries, had to be examined [9, 32]). Potential problems with non-isomorphism of crystals and crystal quality can be associated with data collection from many crystals. Post crystallization treatments will reduce problems with non-isomorphism and variations in cell dimensions.

The measurement of low-resolution reflections is often overlooked in traditional protein crystallography. For virus crystallography these reflections can be very important, for instance in using a low-resolution cryo-electron microscopy reconstruction as a phasing model (see Sect. 4.6 and Chaps. 3 and 7), or for visualizing less well ordered structures [33].

4.5.4 Data Processing

The most common data collection method in macromolecular crystallography is the rotation method that involves, as the name indicates, a simple rotation of the crystal around a single axis in small increments (generally between 0.1° and 0.3° for viruses) while the crystal is exposed to X-rays and the reflections are recorded on an area detector (Fig. 4.5).

Once the crystal is centered, the first rotation image is recorded and the raw electronic signal from the detector is corrected by the detector software for detector-specific non-uniformity defects. The corrected raw image includes a header section with relevant experimental parameters, such as detector distance, beam center position, wavelength etc. and is displayed in a graphical display window. The image is then visually inspected for a number of qualitative key parameters of the crystal. Good diffraction implies single, resolved and strong spots, extending far out in the diffraction angle to high resolution.

Indexing is the assignment of a consistent set of unit cell vectors matching the diffraction pattern. In order for indexing to succeed, the position of the incident primary beam on the detector and the crystal-detector distance needs to be known with precision. For crystals with large unit cells, defining the correct starting values for the detector parameters is the most important for quality indexing. Once a unit cell has been selected, its lattice symmetry determines the minimal symmetry of the reciprocal space. Based on lattice symmetry a data collection strategy is determined. The goal of the strategy is to collect at least enough reflections to cover the entire asymmetric unit of the reciprocal space in order to obtain complete data. Once all image frames are collected, the data processing programs must convert the raw pixel intensity of the measured and indexed diffraction spots into a list of properly scaled intensities and their estimated standard errors. The quality of data is normally judged by the agreement of multiple recorded or symmetry related data.

4.6 Phase Determination

The large unit cell dimensions inherent to virus crystals posed also daunting challenges to the determination of phases required to compute the electron density maps from diffraction data using Fourier transformations (Fig. 4.5). However, the first animal virus crystal structures [7, 8] were landmarks confirming the power of density modification, in particular non-crystallographic symmetry averaging, to solve the phase problem, clearing the path towards more accurate and more complex virus structures that has continued till today [15]. In virus X-ray crystallography the solution to the phase problem usually proceeds in two differentiated steps with first the determination to a certain resolution of some initial phases that are later refined and extended to higher resolution.

This section contains a number of technicalities presenting the approaches used to solve the phase problem in virus crystallography. Readers interested mainly in an overall perspective should only retain the idea of the central importance of the phase problem in diffraction experiments and that some powerful procedures have been developed to solve it in virus crystallography. Interested readers can find more in-depth explanations in the references listed under *Further reading* at the end of this chapter, in particular the last reference.

4.6.1 Initial Phase Determination in Crystals from Icosahedral Virus

To obtain a set of initial phases for an icosahedral virus crystal it is, in general, necessary to determine first the orientations and positions of virus particles in the asymmetric unit of the crystal. Orientations can be determined using the self-rotation function (or even better the locked self-rotation: a self-rotation obeying the restraints imposed by the particle symmetry). Computation of the self-rotation, which requires information only about the structure factors amplitudes, is usually performed in polar coordinates and represented in sections at constants values of the rotation angle κ corresponding to the icosahedral symmetry axes (Fig. 4.5). Once orientations have been determined, information on the position of the center of the viral particles, the translational parameters, can generally be derived from packing considerations. In some cases electron microscopy analysis of thin-sectioned crystals has permitted direct visualization of packing with the determination of the number and position of particles in the unit cell [34]. Often the symmetry of the crystal can constraint the possible orientations and positions of the virus particles by requiring the coincidence of crystal and viral particle symmetries (Fig. 4.4c). For example, when three orthogonal crystal twofold axes superimpose with three twofold axes from the icosahedral particle only two alternative orientations are possible. In this particular case the position of the virus particle is also fixed by the packing that places the center of the icosahedra at the intersection of the three crystal axis. Coincidence of crystal and icosahedral axis reduces the multiplicity of the non-crystallographic symmetry that is 60 when a complete icosahedral particle is found in the crystal asymmetric unit. Fivefold axes are not among the possible crystal symmetries and consequently the fivefold axes from icosahedral particles will remain always as local, non-crystallographic, symmetries.

When orientations and positions of the viral particles in the unit cell are known, initial phases can be obtained using either isomorphous replacement (with heavy atom derivatives) or MR. Isomorphous replacement, historically the first method capable of solving the phase problem in protein crystallography, requires the preparation of heavy atoms derivatives, which are crystals isomorphous to the ones from the native sample (ideally with the same packing and the same molecular structures), but containing also a few extra heavy atoms (with high atomic numbers) [35]. Diffraction differences between native and isomorphous crystals can then provide information about the phases. In turn, MR requires the availability of independent structural information about the sample that could be used to obtain approximate phases or, at least, some constraints on the possible phases. This independent structural information can be very diverse ranging from atomic models or electron density maps of related samples to the presence of non-crystallographic symmetries. Anomalous diffraction, an alternative phasing methodology used frequently in protein crystallography, has essentially never been applied in virus crystallography mainly due to the experimental difficulties of collecting the weak anomalous signal with enough accuracy from virus crystals.

Preparation of isomorphous heavy atoms derivatives is similar to all protein crystals including virus crystals, but because there are 60 equivalent units in the icosahedral particle, at least 60 positions on the particle are expected to be modified, which results in a large number of vectors (at least 60×59) between heavy atoms within one particle. Information about the non-crystallographic symmetry is taken into account both during the heavy-atoms search, which can be restricted to an asymmetric unit of the icosahedron, and also during the refinement of the heavy-atom parameters [36]. Heavy atom derivatives was the method used to solve the first icosahedral viruses structures such as TBSV, SBMV, STNV or HRV14 [1–3]. However, for about the last 20 years the vast majority of new virus crystal structures have been solved without heavy atoms derivates. Instead a diversity of ingenious molecular replacement approaches in combination with the presence of non-crystallographic symmetry has revolutionized virus crystallography, allowing the explosive rise in the number of crystal structures from viruses that have been determined.

When the atomic coordinates of a structurally similar virus is available, classical MR techniques can be employed to place the searching model in the asymmetric unit of the cell of the unknown structure. In some cases it has been reported that summing two or more homologous structures can provide improved searching models [23]. The calculated structure factors provide then approximate phases and amplitudes, with correlation coefficients between observed and calculated amplitudes often used as an estimation of the quality of phases at a given resolution. Starting at this stage, very accurate phases for the unknown structure can be derived reaching near-atomic resolutions (as described below in Sect. 4.6.2) by density modification (Fig. 4.7), even in the presence of the minimal fivefold non-crystallographic symmetry and low similarity [37]. Improved searching models have also been obtained in conjunction with the information provided by three dimensional cryo-electron microscopy (cryo-EM), where the coordinates of a starting model built by similarity have been trimmed and adjusted to fit the cryo-EM map [38] (see also Chaps. 3 and 7). True quasi-atomic models, obtained by combining the X-ray coordinates of one of the capsid components with a cryo-EM reconstruction of the complete capsid, have provided successful starting models in challenging studies.

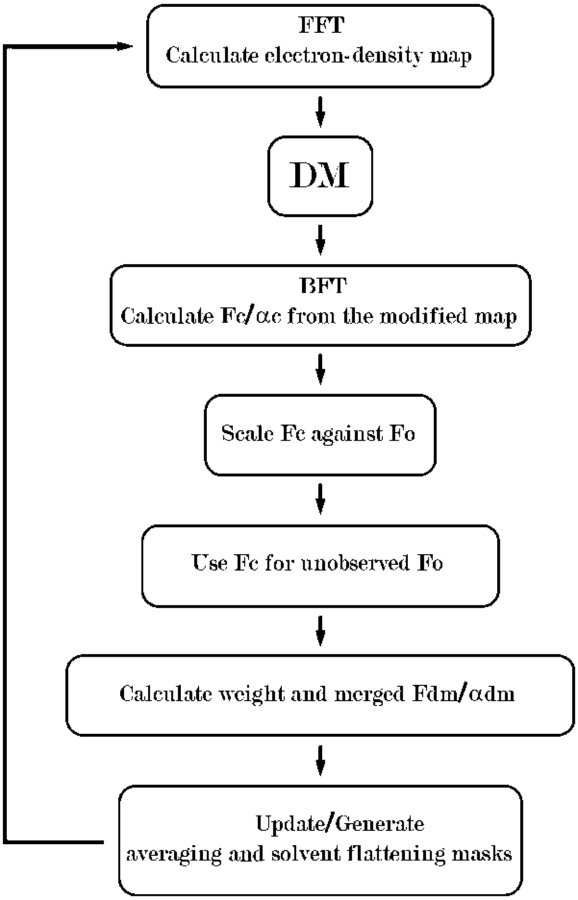


Fig. 4.7 Flow chart of the density modification iterative procedure. Electron density maps, obtained with a Fourier transformation (FFT step), are modified according to the density modification constraints available (DM step), particularly non-crystallographic symmetry averaging and solvent flattening in virus crystallography. The back Fourier transform is then applied (BFT step) to obtain the corresponding structure factors amplitudes and phases (F_c and α_c, respectively). After the experimental and calculated amplitudes are scaled together the calculated structure factors are used to fill the unmeasured reflections. Finally, new amplitudes and phases, ready to provide a new electron density map, are obtained combining and weighting reflections mainly according to differences between experimental and calculated amplitudes. At this point new solvent flattening and averaging masks, defining the volumes where these density constraints can be applied in the next DM cycle, calculated. The process can be repeated iteratively until reaching convergence

After the initial MR solution, the quasi-atomic model is generally improved by rigid body refinement of the capsid components against the experimental diffraction data from the crystal. Using cryo-EM maps of the BTV virion, a quasi-atomic model was built for the outer shell formed by protein VP7, which had been solved by X-ray crystallography [9]. The quasi-atomic model was then placed in the unit cell of BTV crystals providing initial phases, which were then refined and extended to higher resolution (see below) allowing the accurate structure determination not only of the outer VP7 shell, but also of the inner shell formed by 120 subunits of VP2. Similarly, initial phases were obtained by MR for the structure of the membrane-containing bacteriophage PRD1 using a quasi-atomic model of the viral capsid derived from fitting a cryo-EM reconstruction [13]. Position and orientation of the initial model were refined against the experimental diffraction data of PRD1 crystals, which was followed by rigid body refinement of the 240 P3 trimers included in the model. The PRD1 crystal solved at about 4 Å resolution, allowed to determine the conformation variants adopted by P3 and also the structures of proteins P16 and P31, forming the pentameric base of the vertex spike, and of P30. Initial phases for the crystal structure of human adenovirus solved at 3.5 Å resolution, were also obtained by MR using a quasi-atomic model built by fitting the known hexon and penton base structures into the cryo-EM density of the virus [15]. The 150-megadalton adenovirus capsid with nearly one million residues in the crystal asymmetric unit represents, as indicated before, the largest crystal structure ever solved, confirming the power of MR and non-crystallographic symmetry in virus crystallography.

MR has also been sufficient to solve a number of virus crystal structures when using as searching models the electron density of the virus obtained by cryo-EM. The crystal structure of human hepatitis B virus capsid, with a $T = 4$ triangulation icosahedral lattice, was solved using as searching model the 7 Å resolution cryo-EM map [39]. The initial phases to 8 Å resolution were then extended using density modification procedures to 3.3 Å resolution. At this point it became evident that in the electronic density α helices were left handed, showing that the wrong hand had been selected for the EM map used. Luckily in this case one of the icosahedral twofold axis was coincident with a crystal twofold and the phase extension protocol had not been affected except by producing the inverted hand map. Even the virus masks of cryo-EM maps, with constant density inside, have proved capable to provide searching models for good molecular replacement solutions with initial phases that have been then refined and extended to almost atomic resolution [32].

The feasibility of ab initio phase determination has been demonstrated for icosahedral viruses with initial phases calculated using spherical hollow shell models of the viruses [34, 40]. However, this fascinating possibility has seldom been fulfilled in practice because a number of difficulties mainly related with: (i) the sensibility to certain parameters, such as the outer and inner radius of the shell, (ii) the availability of accurate values for intensities of reflections at low resolution (at least two-thirds of the viral radius), or even to (iii) the packing of the virus particles, particularly when positions present a center of symmetry in the unit cell, as for example when an icosahedral twofold axis coincides with a crystallographic even-fold axis.

4.6.2 Phase Refinement and Extension: The Power of Non-crystallographic Symmetry

To improve and extend the resolution of the initial phases, obtained at resolutions lower or much lower than atomic resolution, iterative density modification procedures, in particular non-crystallographic averaging and solvent flattening, can be used (Fig. 4.7). The power of phase determination provided by an iteration of density modification is, according to the semi-quantitative expression derived by Arnold and Rossmann [41], directly related to the square root of the non-crystallographic redundancy (N). Therefore, non-crystallographic symmetry averaging with high non-crystallographic multiplicities, as often available in virus crystallography, can be an extremely powerful mechanism to refine and to extend the initial phases. The accuracy of the observed amplitudes and the proportion of data measured also affect the expected phasing power. In fact, high non-crystallographic redundancy can overcome efficiently the lack of completeness of the experimental diffraction data. Reciprocally, low redundancy can be compensated with the high accuracy and completeness of the measured intensities.

Many of the original problems in virus crystallography related with the averaging algorithms had to deal with limited computer memories and are now essentially solved with the standard software of protein crystallography. It was found that filling the missing experimental amplitudes by including calculated values (suitably weighted) from the back-transformed electron density map can be critical to allow the propagation of phase information and accordingly also critical for the success of the density modification procedure. The power of density averaging is greatly reinforced when used together with solvent flattening, setting the solvent and sometimes also the capsid interior to their respective mean electron densities. In fact, for crystals with a large solvent percentage volume, solvent flattening provides a considerable amount of phasing power, in particular for new reflections added at higher resolution. Thus the envelope or mask defining the solvent must allow the maximum volume without truncating the protein shells. Finding and optimizing the envelopes of the volumes within which the different density modification operations are applied is a critically important aspect of the density modification procedures as already mentioned for the ab initio phasing determination. A number of elaborated alternatives have been developed to refine automatically the envelopes that now are routinely used.

Density modification is mostly applied in real space despite many of the concepts are better controlled in the reciprocal space formulation [42]. In particular, the shell thickness for phase extension should be limited by consideration of the spherical viral particle radius and of the unit cell size. Weighting reflections according to the likely error in their phases can hasten the process. Phases for the newly phased reflections should be allowed a few cycles to converge before the unobserved data are added in. Poor agreement of the strong low-resolution terms may inhibit phase extension. It may therefore be necessary to impose low-resolution cut-offs on the observed data if, for instance, reflections close to the back-stop shadow have not been measured properly. Termination errors can be reduced by extending the reflections a little beyond the resolution limit of the data, thus allowing the calculated data to contribute to the proper phasing of the observed reflections at the highest resolution. The process of phase refinement and extension is normally monitored by reference to averaging agreement R factors and more reliably to the correlation coefficients between the observed and the calculated structure factors analyzed in terms of resolution.

4.7 Map Interpretation

The final step in the structure determination is the building of a chemically reasonable model that fits the electron density maps generated in the previous steps (Fig. 4.5). Refinement allows completing and optimizing the model maximizing the agreement with all the experimental information available.

4.7.1 Model Building

The accuracy of the phases obtained by the averaging procedure results in high quality electron density maps where the amino acids sequence is often recognized even despite the relatively low resolution (in some cases lower than 3.5 Å) of many of the virus crystal structures determined. In these maps a model is normally constructed for a single icosahedral subunit using standard protein crystallography tools (Fig. 4.5). For viruses with triangulation numbers higher than one the icosahedral subunit contains multiple structurally different copies of a given protein, but once a model for the first protein copy becomes available it can be adapted to provide an initial fit for the electron density corresponding to the remaining copies. Minimaps, skew maps and other map interpretation techniques, widely used in virus crystallography and referred in the literature till the early 1990s, have now been abandoned by the possibilities offered by powerful graphic systems and interactive software such as COOT [43]. For some virus structures solved at low resolution, like PRD1 at about 4 Å resolution [13], averaged difference Fourier maps corresponding to selenium methionine derivatives were calculated using the refined final phases of the native virus. Peaks in these maps, corresponding do the different number of electrons between the selenium and the sulfur atoms (from the selenium and the normal methionines, respectively), flagged the positions of methionines facilitating and providing controls to the sequence assignment during model building.

4.7.2 Model Refinement

The coordinates and atomic temperature factors can be refined to minimize the differences between the calculated and observed structure factors while maintaining reasonable stereochemistry. The number of refined parameters can be sharply reduced by applying strict non-crystallographic symmetry constraints resulting in a much higher data-to-parameters ratio, though constrained refinement is only available in a few refinement packages such as CNS and XPLOR [44] or, recently, REFMAC [45]. Since phases derived from the averaging process can be very accurate, these phases can also be used as experimental observations. In fact, due to the quality of the density maps obtained by averaging the aim of refinement is, in general, not to improve the maps, but simply to improve the atomic model. The quality of the averaged maps together with the large computational requirements of virus refinement in reciprocal space explains the successful application of real space refinement to a number of virus structures [3]. From the atomic models a solvent mask is now routinely constructed by most refinement programs, which has a major contribution to the diffraction data at low resolution data (below 8 Å resolution). In cases where high resolution and good data completeness is available the strict non-crystallographic symmetry constraints can be relaxed to analyze deviations from perfect symmetry among the subunits found in the crystal asymmetric unit. This can be particularly important for some external parts of the virus shells involved in crystal contacts [46].

4.8 Understanding Viruses: Some Major Contributions of X-ray Crystallography

The amount, diversity and complexity of viral capsid structures now available at near atomic resolution and stored in the PDB (Protein Data Bank; http://www.rcsb.org/pdb/) or in specialized structural data bases, such as VIPER (VIrus Particle ExploreR; http://viperdb.scripps.edu/), are the result of the outstanding achievements of structural virology, which now treasures the three dimensional structures of over 350 virus capsids, from more than 35 viral families. It is difficult to exaggerate the importance of X-ray crystallography in this process as until very recently X-ray crystallography was the only methodology, and it is still the dominant technique, allowing reaching high resolution with large systems. However, EM is now confirming its enormous potential by providing near-atomic resolution for a rapidly growing number of viral particles (see Chap. 3).

Many more viral protein structures, about 5,000, are also available and a very high percentage of them, more than 90 %, have been determined by protein X-ray crystallography (PDB sept/2012). These includes from structural proteins not forming icosahedral capsids, such as the proteins of cylindrical viruses or the different components of tailed phages, to many viral enzymes, such as viral proteinases or RNA and DNA polymerases.

The enormous amount of information contained in the structures of virus capsids and proteins, covering many of the relevant issues in virology and allowing many others types of research, is deeply changing our view of the virus world at an accelerated pace. Many of the structures available correspond to potential targets for antiviral drugs and structural virology has had a major role in the development of pharmaceuticals for clinically important viral diseases such as AIDS (see Chap. 20). Drugs directed against the HIV aspartic proteinase are a paradigmatic example of the relevance of structural biology in the development of viral medications, with the structure determination of thousands of crystals from the proteinase complexed with different compounds. The structures of essentially all the protein types encoded by some viruses, such as picornaviruses, have been determined, which allows evaluating the possibilities and limitations of the structural genomics strategies. In fact, it appears clear from these examples that the structural information available for the individual proteins provides only an initial glimpse of the viral functioning as the structures of a diversity of protein complexes and of the proteins interacting with nucleic acids together with alternative protein conformations are still required.

The near-atomic resolution structure and organization of the capsids of some viruses have enriched the information about the underlying building principles expanding the quasi-equivalence symmetry and icosahedral triangulation concepts towards, for example, the interpretation of departures from these principles (as in Polyomavirus or BTV) or to pseudo triangulation arrangements (as in phage PRD1, adenoviruses and many other viruses). This type of information has also contributed to the detailed modeling of the largest spherical viruses, despite they are not yet amenable to high resolution X-ray crystal studies [47, 48].

Specific contributions of X-ray crystallography for understanding particular aspects of stages in the virus life cycle are extremely numerous. We will not deal with particular examples in this chapter, because a substantial number of those contributions have been described in many other chapters of this book, specifically Chaps. 2, 7, 10, 11, 12, 13, 15, 16, 17, 20, 21 and 22. The reader is referred to those chapters to find relevant studies where high-resolution structures of viral particles or components determined by X-ray crystallography have proved instrumental for our understanding of the molecular biology of viruses.

4.9 Perspectives

At synchrotrons, new X-ray detectors and micro-beams together with “in situ” diffraction are leading to a step change that is central to virus X-ray crystallography. Even beyond this, ideas are being developed to grow viral crystals in infected cells, bypassing the requirements for purification and crystallization [49]. Such tiny crystals (typically sub-micrometer) might be amenable to diffraction with advanced synchrotrons or X-ray free-electron lasers. In fact, the major increases in X-ray brilliance, compared with the most brilliant third-generation SR sources, anticipated from the coming generation of X-ray lasers indicate that the sample size can be reduced to just a single molecule offering the possibility of single virus particles analysis. The oversampling of the continuous scattering originated by the diffraction of a non-periodic sample allows an alternative approach to solve the "phase problem" [50]. In these experiments a sufficiently brief (in the femtoseconds range) and intense X-ray pulse should provide sufficient photons for a useful pattern before radiation damage commences. There are now clear experimental evidences that this single molecule X-ray "diffract and destroy" analysis is feasible and can be applied in particular to virus, though at present only very low resolution has been achieved [51].

Since the first virus structures were determined at almost atomic resolution, about 35 years ago, virus X-ray crystallography has continued providing an always deeper and wider understanding of the virus world with a major impact in the innovation of the X-ray methodology. However, many of the questions remaining in virus structure and functioning will require an even bolder use of X-ray studies in combination with new experimental approaches, such as high-resolution electron microscopies and single-molecule techniques (see also Chaps. 3, 7, 8 and 9).

Acknowledgements We would like to dedicate this chapter to Prof. M.G. Rossmann for his pioneering and continuous ground-breaking results in structural virology. This work was supported by grants from the *Ministerio de Economia y Competitividad* to N.V. (BIO2011-24333) and to I.F. (BFU2009-09268).

References and Further Reading

1. Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G (1978) Tomato bushy stunt virus at 2.9 A resolution. Nature 276:368–373
2. Abad-Zapatero C, Abdel-Meguid SS, Johnson JE, Leslie AG, Rayment I, Rossmann MG, Suck D, Tsukihara T (1980) Structure of southern bean mosaic virus at 2.8 A resolution. Nature 286:33–39
3. Liljas L, Unge T, Jones TA, Fridborg K, Lövgren S, Skoglund U, Strandberg B (1982) Structure of satellite tobacco necrosis virus at 3.0 A resolution. J Mol Biol 159:93–108
4. Verdaguer N, Blaas D, Fita I (2000) Structure of human rhinovirus serotype 2 (HRV2). J Mol Biol 300:1179–1194
5. Verdaguer N, Fita I, Reithmayer M, Moser R, Blaas D (2004) X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. Nat Struct Mol Biol 11:429–434
6. Garriga D, Pickl-Herk A, Luque D, Wruss J, Castón JR, Blaas D, Verdaguer N (2012) Insights into minor group rhinovirus uncoating: the X-ray structure of the HRV2 empty capsid. PLoS Pathog 8:e1002473
7. Rossmann MG, Arnold E, Erickson JW et al (1985) Structure of a human common cold virus and functional relationship to other picornaviruses. Nature 317:145–153
8. Hogle JM, Chow M, Filman DJ (1985) Three-dimensional structure of poliovirus at 2.9 A resolution. Science 229:1358–1365

9. Grimes JM, Burroughs JN, Gouet P, Diprose JM, Malby R, Ziéntara S, Mertens PP, Stuart DI (1998) The atomic structure of the bluetongue virus core. Nature 395:470–478
10. Wilson IA, Skehel JJ, Wiley DC (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature 289:366–373
11. Roberts MM, White JL, Grütter MG, Burnett RM (1986) Three-dimensional structure of the adenovirus major coat protein hexon. Science 232:1148–1151
12. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature 393:648–659
13. Abrescia NG, Cockburn JJ, Grimes JM et al (2004) Insights into assembly from structural analysis of bacteriophage PRD1. Nature 432:68–74
14. Cockburn JJ, Abrescia NG, Grimes JM et al (2004) Membrane structure and interactions with protein and DNA in bacteriophage PRD1. Nature 2004(432):122–125
15. Reddy VS, Natchiar SK, Stewart PL, Nemerow GR (2010) Crystal structure of human adenovirus at 3.5 A resolution. Science 329:1071–1075
16. Stehle T, Casasnovas JM (2009) Specificity switching in virus-receptor complexes. Curr Opin Struct Biol 19:181–188
17. Ganser-Pornillos BK, Yeager M, Sundquist WI (2008) The structural biology of HIV assembly. Curr Opin Struct Biol 18:203–217
18. Harrison SC (2005) Mechanism of membrane fusion by viral envelope proteins. Adv Virus Res 64:231–261
19. Mukhopadhyay S, Kuhn RJ, Rossmann MG (2005) A structural perspective of the flavivirus life cycle. Nat Rev Microbiol 3:13–22
20. Wang X, Peng W, Ren J et al (2012) A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71. Nat Struct Mol Biol 19:424–429
21. Burzlaff H, Gruber B, Wolf PM et al (2002) Crystal lattices. In: Hahn T (ed) International tables for crystallography, vol A. Kluwer Academic Publishers, Dordrecht
22. Fry E, Grimes J, Stuart DI (1999) Virus crystallography. Mol Biotechnol 12:13–23
23. Fry E, Logan D, Stuart D (1996) Virus crystallography. In: Jones C, Mulloy B, Sanderson M (eds) Methods in molecular biology, vol 56: crystallographic methods and protocols. Humana Press, Totowa
24. Vijayachandran LS, Viola C, Garzoni F et al (2011) Robots, pipelines, polyproteins: enabling multiprotein expression in prokaryotic and eukaryotic cells. J Struct Biol 175:198–208
25. Chayen NE, Helliwell JR, Snell EH (2010) Macromolecular crystallization and crystal perfection. Oxford University Press, Oxford
26. Newman J (2005) Expanding screening space through the use of alternative reservoirs in vapor-diffusion experiments. Acta Crystallogr D Biol Crystallogr 61:490–493
27. Chayen NE (2005) Protein crystallization: automation robotization and miniaturization. In: Sundstrum M, Norin M, Edwards A (eds) Structural genomics and high throughput structural biology. CRC Press/Taylor & Francis, Abingdon
28. Garman E (2003) 'Cool' crystals: macromolecular cryocrystallography and radiation damage. Curr Opin Struct Biol 13:545–551
29. Mueller M, Jenni S, Ban N (2007) Strategies for crystallization and structure determination of very large macromolecular assemblies. Curr Opin Struct Biol 17:572–579
30. Cockburn JJ, Bamford JK, Grimes JM, Bamford DH, Stuart DI (2003) Crystallization of the membrane-containing bacteriophage PRD1 in quartz capillaries by vapour diffusion. Acta Crystallogr D Biol Crystallogr 59:538–540
31. Axford D, Owen RL, Aishima J et al (2012) In situ macromolecular crystallography using microbeams. Acta Crystallogr D Biol Crystallogr 68:592–600
32. Reinisch KM, Nibert ML, Harrison SC (2000) Structure of the reovirus core at 3.6 Å resolution. Nature 404:960–967
33. Gouet P, Diprose JM, Grimes JM et al (1999) The highly ordered double-stranded RNA Genome of bluetongue virus revealed by crystallography. Cell 97:481–490

34. Taka J, Naitow H, Yoshimura M, Miyazaki N, Nakagawa A, Tsukihara T (2005) *Ab initio* crystal structure determination of spherical viruses that exhibit a centrosymmetric location in the unit cell. Acta Crystallogr D Biol Crystallogr 61:1099–1106
35. Baker TS, Johnson JE (1997) Principles of virus structure determination. In: Chiu W, Burnett RM, Garcea RL (eds) Structural biology of viruses. Oxford University Press, New York
36. Arnold E, Vriend G, Luo M, Griffith JP, Kamer G, Erickson JW, Johnson JE, Rossmann MG (1987) The structure determination of a common cold virus, human rhinovirus 14. Acta Crystallogr A43:346–361
37. Fry E, Acharya R, Stuart D (1993) Methods used in the structure determination of foot-and-mouth disease virus. Acta Crystallogr A49:45–55
38. Speir JA, Munshi S, Wang G, Baker TS, Johnson JE (1995) Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. Structure 3:63–78
39. Wynne SA, Crowther RA, Leslie AGW (1999) The crystal structure of the human hepatitis B virus capsid. Mol Cell 3:771–780
40. Tsao J, Chapman MS, Wu H, Agbandje M, Keller W, Rossmann MG (1992) Structure determination of monoclinic canine parvovirus. Acta Crystallogr B48:75–88
41. Arnold E, Rossmann MG (1985) Effect of errors, redundancy and solvent content in the molecular replacement procedure for the structure determination of biological macromolecules. PNAS 83:5489–5493
42. Rossmann MG (1989) The molecular replacement method. Acta Crystallogr A46:73–82
43. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132
44. Brunger AT, Kuriyan J, Karplus M (1987) Crystallographic R factor refinement by molecular dynamics. Science 235:458–460
45. Murshudov GN, Skubak P, Lebedev AA et al (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D67:355–367
46. Garriga D, Querol-Audí J, Abaitua F et al (2006) The 2.6-Angstrom structure of infectious bursal disease virus-derived T=1 particles reveals new stabilizing elements of the virus capsid. J Virol 80:6895–6905
47. Nandhagopal N, Simpson AA, Gurnon JR et al (2002) The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. PNAS 99:14758–14763
48. Zhanga X, Xianga Y, Duniganb DD et al (2011) Three-dimensional structure and function of the Paramecium bursaria chlorella virus capsid. PNAS 108:14837–14842
49. Abrescia NGA, Bamford DH, Grimes JM, Stuart DI (2012) Structure unifies the viral universe. Annu Rev Biochem 81:795–822
50. Raines KS, Salha S, Sanberg RL, Jiang H, Rodríguez JA, Fahimian B, Kapteyn HC, Du J, Miao J (2010) Three-dimensional structure determination from a single view. Nature 463:214–217
51. Seibert MM, Ekeberg T, Maia FR et al (2011) Single mimivirus particles intercepted and imaged with an X-ray laser. Nature 470:78–81

Further Reading

Blow D (2002) Outline of crystallography for biologists. Oxford University Press, Oxford

Rhodes G (2006) Crystallography made crystal clear, 3rd edn. Academic Press, London

Rupp B (2010) Macromolecular crystallography. Garland Science, Taylor & Francis group, LLC, New York

Chapter 5
Nuclear Magnetic Resonance Spectroscopy to Study Virus Structure

José L. Neira

Abstract Nuclear magnetic resonance (NMR) is a spectroscopic technique based in the absorption of radiofrequency radiation by atomic nuclei in the presence of an external magnetic field. NMR has followed a "bottom-up" approach to solve the structures of isolated domains of viral proteins, including capsid protein subunits. NMR has been instrumental to describe conformational changes in viral proteins and nucleic acids, showing the presence of dynamic equilibria which are thought to be important at different stages of the virus life cycle; in this sense, NMR is also the only technique currently available to describe, in atomic detail, the conformational preferences of natively unfolded viral proteins. NMR has also complemented X-ray crystallography and has been combined with electron microscopy to obtain pseudo-atomic models of entire virus capsids. Finally, the joint use of liquid and solid-state NMR has allowed the identification of conformational changes in intact viral capsids on insertion in host membranes.

Keywords Assembly • Binding • Capsid • Chemical shifts • Conformational changes • Coupling constants • Dynamics • Equilibrium • Flexibility • Hydrogen-exchange • Membrane • Nuclear magnetic resonance (NMR) • Nucleic acid • Protein-protein interactions • Protein-nucleic-acid interactions • Solid-state • Solution structure • Unfolded protein • Virus

---

J.L. Neira (✉)
Instituto de Biología Molecular y Celular, Universidad Miguel Hernández,
03202 Elche, Alicante, Spain

Instituto de Biocomputación y Física de Sistemas Complejos, 50018 Zaragoza, Spain
e-mail: jlneira@umh.es

M.G. Mateu (ed.), *Structure and Physics of Viruses; An Integrated Textbook*,
Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_5,
© Springer Science+Business Media Dordrecht 2013

Abbreviations

BLV	Bovine leukemia virus
CA	Retrovirus capsid protein
CTD	C-terminal domain of CA
COSY	Correlation spectroscopy
EIAV	Equine infectious anemia virus
EM	Electron microscopy
FID	Free induction decay
HIV-1	Human immunodeficiency virus type 1
hLysRS	Human lysyl-tRNA synthetase
HTLV	Human T-cell leukemia virus
MA	Matrix protein
MMTV	Mouse mammary tumour virus
MoMuLV	Moloney murine leukemia virus
MV	Monkey virus
NC	Nucleocapsid
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NTD	N-terminal domain of CA
PIP2	Phosphatidylinositol-4,5-diphosphate
RSV	Rous sarcoma virus
SL	Stem loop
SIV	Simian immunodeficiency virus
SSNMR	Solid-state NMR
TOCSY	Total correlation spectroscopy
VPr	Accessory viral protein R

5.1 Introduction

Although X-ray crystallography (Chap. 4) has been extensively used for determination of the three-dimensional structures of virus particles and viral proteins at atomic resolution, not all biological macromolecules can lead to crystals, and many biomolecules can behave differently in the crystalline state as they do in solution; in those cases, nuclear magnetic resonance (NMR) spectroscopy has been used to determine their high-resolution structures. In addition, the structures of biomolecular complexes, the monitoring of enzymatic reactions, the description of protein folding reactions, or the exploration of the dynamic properties of biomolecules can be addressed by NMR. Therefore, NMR spectroscopy has emerged as an important complement to X-ray crystallography in structural biology in general, and in structural virology in particular.

In this chapter, I shall describe the foundations of nuclear magnetic resonance (NMR) spectroscopy and some of its applications to study viral proteins and virus particles. This technique has a serious limitation when faced to the study of large proteins or macromolecular assemblies: the signals from those biomolecules get too broad to be clearly observed in a spectrum. Due to this limitation, NMR has dealt with the study of whole virus particles, and large viral proteins, by solving the structure of small protein subunits or isolated protein domains of larger proteins (a “bottom-up” approach). NMR has been applied also to study the conformational dynamics of viral proteins in solution, and there have been attempts to tackle the dynamics of complete viral particles. Only recently, solid-state NMR (SSNMR) techniques have been applied to the structural study of whole, intact virus particles in “quasi-solution” conditions.

5.2 Physical Principles of NMR Spectroscopy

5.2.1 The Basis of NMR Spectroscopy

NMR is one of the few spectroscopies that can be explained from both a classic perspective and by using a quantum mechanics point of view. In-depth explanations of NMR theory are beyond the scope of this chapter, and I shall describe briefly only some basic aspects in the following paragraphs. There are excellent books where the interested reader can find a detailed treatment of both complementary views on the NMR foundations [1–5].

NMR is a physical phenomenon based on the magnetic properties of atomic nuclei. Those nuclei with odd numbers of protons and/or neutrons have a magnetic moment (μ) due to their spinning. From a classical point of view, in the presence of an external static magnetic field (B_0), the μ of the nuclei will orient parallel to (with) or anti-parallel to (against) B_0. This interaction originates a precession of the nuclear dipoles around the external field (Fig. 5.1a), in a manner analogous to a gyroscope spinning in a gravitational field. The angular frequency (ω) of this motion depends on the strength of B_0 and the properties of μ. The Boltzmann distribution dictates that the most stable state (nuclear magnetic dipoles oriented with B) will be the most populated at equilibrium, albeit by only a small fraction, because of the small difference in energy with the less stable state (dipoles oriented against B_0). If an appropriate amount of energy is applied, the nuclear magnetic dipoles will move from their equilibrium position and will become oriented against B_0. The corresponding absorption of energy by the nuclei is measured by NMR.

From a quantum perspective, if we consider the simplest nucleus (that of ^{1}H, with a single proton) in the absence of an external magnetic field B_0, the two possible orientations of the nuclear spin (± ½) have an equivalent energy (they are degenerate). In the presence of B_0, two different quantum levels will appear (Fig. 5.1b). The energy difference (E) between those levels is the dot product of

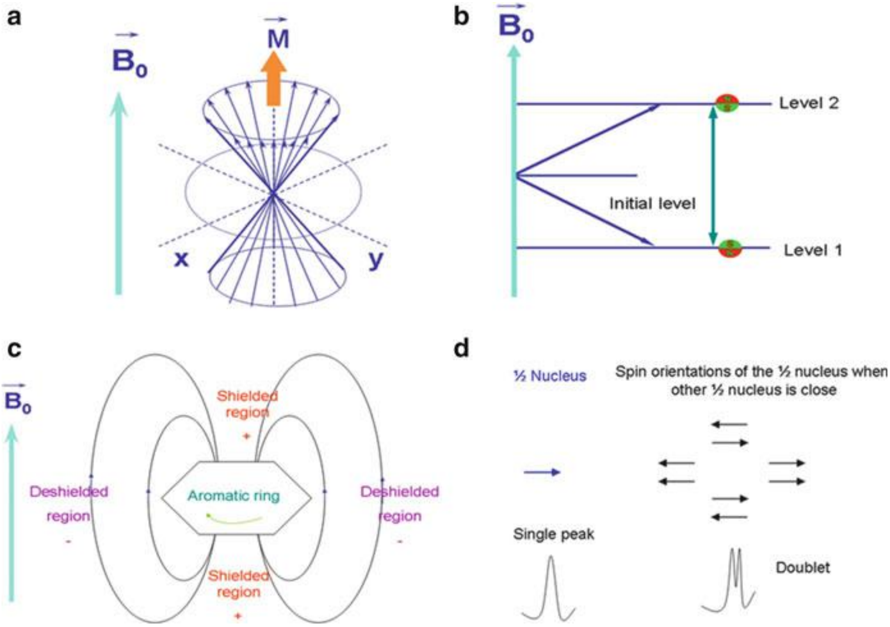


Fig. 5.1 Physical basis of NMR. (**a**) Classical view of NMR. Nuclear spins are precessing around the external magnetic field (B_0, *cyan arrow*); the net magnetization of the system (the sum of all the nuclear magnetic dipoles of the sample) is shown in *orange* (M). (**b**) Quantum view of NMR. The two quantum levels originated as a consequence of the existence of B_0 (*blue arrow*) are shown (for the sake of clarity, it has been assumed that the nuclear spin has two possible orientations only, as in ^{1}H); the orientation of the two spins is shown as *small circles* in each level indicating the polarity of the nuclear magnets. The *double arrow* indicates the jump from one quantum level to another. (**c**) The concept of shielding in aromatic side chains. The *small blue arrows* indicate the direction of the created ring-current magnetic field, for which two lines have also been drawn. The *green line* on the aromatic ring shows the movement of the electron rings. The *red region* is the shielded one, and the *purple zone* is the deshielded region. The direction of the external magnetic field is indicated by a *large cyan arrow*. (**d**) The coupling constant concept for a nucleus with spin ½ (two possible orientations), with a nearby nucleus also with spin ½. From a quantum perspective (see [2–7]), some of those possible orientations are indistinguishable, and they result in two lines (a doublet) in the NMR spectrum instead of the expected single line

B_0 and μ, and it matches the energy within the radiofrequency region; the small separation between both quantum levels is responsible for the low sensitivity of the NMR spectroscopy. To induce a transition between those levels it will be necessary to apply electromagnetic radiation, with an energy given by the Planck equation [1–7]. Thus, the technique is called "nuclear", because we are observing nucleus; it is called "magnetic", because we need a magnetic field to have the different non-degenerate quantum levels; and, it is called "resonance" because we need to apply the exact energy (in the radiofrequency region), provided by the Planck equation corresponding to the energetic separation between the split levels.

In general, nuclei with even numbers of protons and neutrons (such as the most abundant isotope of carbon, ${}_{6}{}^{12}C$) will have zero spin and they will not yield an NMR spectrum (because there are no levels to split in the presence of B_0). Those nuclei with an odd number of either protons or neutrons will have a half-integer spin; and those with odd numbers of both protons and neutrons will have an integer spin (larger than ½). If the spin of the nucleus is larger than ½ there are more than two quantum states in the presence of B_0 (or, in other words, two different ways of spinning in the classical view), and more spectral transitions, which makes those nuclei much less convenient for NMR analyses than ^{1}H and other nuclei with spin ½. Since the total number of protons and neutrons varies from one isotope to another, it is clear that for the same element there will be some NMR-active isotopes. The most interesting nuclei, from a biomolecular point of view, are: ^{1}H (100 % naturally abundant); ^{15}N, ^{13}C (natural abundance of 0.36 % and 1.11 %, respectively; biomacromolecules need to be artificially enriched in ^{15}N and/or ^{13}C if these isotopes are to be used in NMR experiments; see below); and ^{31}P (100 % naturally abundant; especially useful for studies on nucleic acids).

5.2.2 Chemical Shifts

The atomic nuclei in a molecule are not isolated: they interact with their own electrons, those of other atoms, and with other atomic nuclei. Since the electronic shells have their own spin, they will create an internal magnetic field which will be "felt", in different ways, by the nearby nuclei; therefore, each nucleus in a molecule will have a different "view" (the so-called "shielding") of the external magnetic field B_0, depending on its surroundings. Perhaps the clearest example of this shielding is provided by aromatic rings. In the presence of B_0, a planar aromatic ring with its delocalized π electrons "creates" its own magnetic field (Fig. 5.1c). A nucleus above or below the ring plane will be shielded from the external magnetic field, but it will be unshielded if it is located at other positions relative to the ring; thus, the magnitude of the total field "felt" by a nucleus informs on its location with respect to the ring. In the same way, hydrogen-bonding decreases the electronic shell around a nucleus and reduces the shielding effect. Therefore, it is possible to say that each nucleus in a molecule resonates at a slightly different energy (frequency) relative to others, because of its specific magnetic environment (that is, each nucleus needs a slightly different energy to be promoted to its quantum excited state).

It is impractical to measure the absolute frequency of resonance for every nucleus, and a relative frequency is provided instead. A standard small compound is introduced with the NMR sample, and resonance frequencies of every nucleus in the molecule under study are referred to that of the standard: we measure, in a first stage, the differences between the resonance frequency of each nucleus and that of the standard. Since these measurements of frequencies are provided in hertz (Hz) they are B_0-dependent, and then, the spectrometer magnetic field must be provided.

To avoid an instrument-dependent parameter, a unit-less chemical shift (δ) is provided, given in parts per million (ppm). Thus, δ has the same operational meaning than the wavelength (λ) or frequency (υ) in other spectroscopies (see Chap. 6).

5.2.3 Spin-Spin Coupling: Seeing Other Covalently-Bound Nuclei

The chemical shift, δ, at which a nucleus resonates is a measure of nucleus-electron-clouds interactions. Another important parameter in NMR is the *coupling constant, J*, which is a measure of the magnetic interaction between the spins of two or more covalently bound nuclei (and thus, the interaction occurs *through* the electrons of the bond). The magnetic field "felt" by a nucleus depends on whether the covalently bound nucleus, with a μ, is aligned or not with B_0 (Fig. 5.1d). Therefore, the local field "felt" will be either slightly reduced or increased (when compared to the original magnetic field "felt" when only electronic shielding is taken into account), due to the presence of the neighbouring, covalently-bond nuclei. Then, the original absorption line in the NMR spectrum is split into several lines (two for the case of a single nearby nucleus of hydrogen, in the simplest case, Fig. 5.1d); the strength of the splitting is given by J. An NMR-active nucleus covalently bound to nuclei with a null μ (such as protons bound to ${}_6{}^{12}C$) will not show a splitting of its NMR spectral lines; conversely, a nuclei with a nuclear spin larger than ½ (such as ^{14}N) will tend to broaden neighbouring nuclei lines in the NMR spectrum, rather than splitting them.

The J is measured in Hz; it is independent of B_0, and depends on the nuclei involved and the number of covalent bonds between the nuclei. For instance, the values of J for protons range from 0 to about 20 Hz, but the one-bond J for the ^{15}N-^{1}H is close to 90 Hz. As J depends on the covalently bound nuclei, it provides information on the dihedral angle connecting them (for protons, for example, this is given by a mathematical relationship termed the Karplus equation [1]). This finding has profound implications, since it allows us to obtain structural information: for instance, the J between the amide proton of an amino acid residue in a protein and its H_α proton varies, depending on the type of secondary structure element (α-helix or β-sheet) where the residue is located: 3–5 Hz for residues in α-helices and 8–9 Hz for residues in β-sheets. Therefore, J allows to obtain direct structural information on the dihedral angles between covalently-bound atoms.

5.2.4 Relaxation Mechanisms

For simplicity, let us consider a nucleus such as ^{1}H with a nuclear spin of ½, and then, two quantum levels in the presence of a magnetic field. Electromagnetic radiation of the appropriate υ (in the radiofrequency region) causes spins to jump from the lower to the higher quantum level. As the lower level (μ aligned with the external magnetic field, according to the classical view) has a larger population of spins than the higher one, there will be a net absorption of radiofrequency radiation, leading to a signal in the NMR spectrum. If the electromagnetic radiation is intense enough, the number of spins in the lower and higher levels becomes equal, and therefore, there is no net absorption of energy, because there is no room in the higher level. At the quantum view, we say that the signal is saturated: as many photons are emitted as they are absorbed, and no signal is observed in the NMR spectrum.

However, there are always spontaneous transitions allowing the spins in the saturated higher level to return to the equilibrium population (to the lower level). In other spectroscopic techniques, the recovery of the equilibrium population occurs by "crashing" with other molecules, and losing the excess energy by heat release (see Chap. 6). In NMR, a nucleus relaxes by interaction (through the μ) with a nearby nucleus (from either the same or different molecules); the *spin-lattice* relaxation time, T_1, is the time the equilibrium spin population takes to recover by exchanging energy with the surroundings (*i.e.*, the *lattice*). For a nucleus to contribute effectively to the relaxation of another nucleus, it should be close, and it should belong to a molecule(s) which rotate(s) with an angular frequency similar to the magnetic resonance transition frequency of the studied nucleus (the quantum level separation).

There is another characteristic time, the *spin-spin* relaxation time, T_2, which also measures the recovery of the equilibrium population. The time T_2 refers to interactions *between* spins and measures the rate with which spins exchange energy with each other, which do not involve interactions with the environment (*i.e.*, with the *lattice*). The time T_2 depends, like T_1, on the molecular size and the speed of rotation of the molecule (which is related to the molecular tumbling). Furthermore, T_2 is inversely proportional to the NMR resonance linewidths: the more efficient the T_2-relaxation process, the larger the linewidths (as observed in large molecules, or small ones in rigid environments, for example in SSNMR, see Sect. 5.4.4). Both T-relaxation times are important, since they provide information about the dynamics of the molecule.

5.2.5 The Nuclear Overhauser Effect: Seeing Other Nuclei Through Space

A very important parameter in NMR is the *nuclear Overhauser effect* (NOE), which is due to a through-space interaction between nuclear spins. An NOE is a change of intensity of an NMR resonance corresponding to a particular nucleus, when a nearby nucleus is irradiated. Those nuclei do not have to be covalently bound; it is necessary and sufficient that they are spatially close enough. Thus, determination of NOEs for different nuclei in a macromolecule allows us to determine which nuclei are near each other in space, a crucial step in determining its three-dimensional structure by NMR (see below).

When intense radiation corresponding to the resonance transition frequency of one nucleus is applied to the sample, spin saturation occurs and, therefore, equal populations of spins are found in the lower and higher levels (as described in Sect. 5.2.4). However, that nucleus is not isolated, it can interact with other nuclei nearby (belonging to the same or different molecule). This results in a change in the equilibrium population of spins. Therefore, when we irradiate a nucleus with the proper energy to cause a transition between its levels (and then, to observe its corresponding peak in the NMR spectrum), we shall observe a variation in the intensity of the peaks of nearby nuclei, compared to the intensity observed if the first nucleus had not been excited. Since the interaction between the corresponding magnetic dipoles depends on r^{-6} (r being the distance between them), an NOE is only observed for close neighbours (usually less than 5Å); the NOE short-range limits its applications, but it provides great specificity in the determination of molecular structures by NMR (see below).

5.2.6 Pulse Techniques and Multi-dimensional NMR Spectroscopy

Most spectroscopic techniques (such as fluorescence and CD, see Chap. 6) collect the spectra by varying the wavelength of the excitation source. We could do the same in NMR: varying the wavelength of the radiofrequency source would sequentially induce transitions of different atomic nuclei between their quantum levels (Fig. 5.1b); however, modern NMR spectrometers use radiofrequency pulses, due to improved signal-to-noise ratio and versatility. Since most of the tiny magnetic dipoles – but not all, due to the equilibrium Boltzmann distribution – are aligned with the external field (Fig. 5.1a), there is a net magnetization, M, along the direction defined by B_0 (in the classical view), and most of the spins will start precessing along that direction (see Sect. 5.2.1) (Fig. 5.2a, left side). If we apply a short (microsecond) radiofrequency perturbation (a *pulse*) along the x-axis (Fig. 5.2a), B_1, with a frequency matching the separation energy between the quantum levels (or, in other words, in the classical view, the frequency of precession of the magnetization), the spins aligned initially with the z-axis will start moving away from that axis, and will tend to align with the y-axis (and therefore, in the classical view they will start precessing along that axis). When we remove the perturbation (the pulse) and as the time passes, the different spins forming the net magnetization M, will start to separate at their different precession frequencies (due to their diverse δs), returning to the z-axis (Fig. 5.2a, right side). Thus, the magnetization vector will decrease with a T_1-relaxation time constant (see Sect. 5.2.4.) leading to an exponentially decreasing signal termed a *free induction decay* (FID). The FID is usually observed during a time called t_2 (note that this is an *observation* time, different to the relaxation time T_2) (Fig. 5.2a, right side). Each spin (nucleus) makes a different contribution to the FID due to its slightly

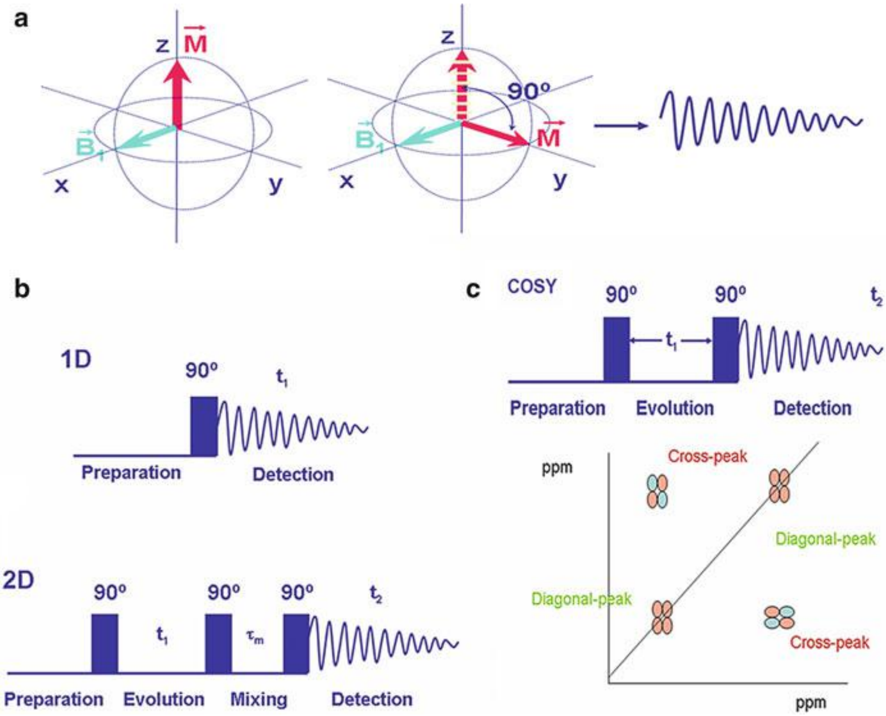


Fig. 5.2 Pulses and n-dimensional NMR. (**a**) How does a pulse work? Initially (*left image*) the system is in equilibrium in the presence of an external magnetic field (which is not indicated); that is, this image shows the state of the system as in Fig. 5.1a. A pulse (B_1, cyan) is applied along the x-axis, and its effect (*right image*) is to tilt the magnetization away from the z-axis. I have represented a pulse which tilts completely the magnetization on the y-axis, but it does not necessarily do so (see, for instance [2–7], for the different lengths of the pulses). From the evolution with time of this tilted magnetization to recover its equilibrium position a FID is obtained (*shown at far right*). (**b**) A series of pulse sequences (pulses are indicated by *rectangles*): *top*, scheme to obtain a 1D-NMR spectrum; *bottom*, scheme to obtain a 2D-NOESY spectrum. The time periods during the experiments are labelled; by convention, the acquisition time during a 1D-NMR spectrum is called t_1 (but it is called t_2 in the 2D-NMR spectra). The numbers on top of the rectangles (the pulses) indicate the tilt caused by the pulse (that is, a 90° pulse tilts the magnetization completely on the y-axis). (**c**) A scheme to obtain a 2D-COSY spectrum. The pulse sequence is depicted on top, where the time t_1 is linearly increased in each experiment. The resulting spectrum showing the diagonal peaks and the cross-peaks is shown at the bottom. The clusters of peaks in a COSY are four-lobed, with alternated signs (positive and negative), due to the effect of the coupling constant among the covalently-bound nuclei

different δ; the individual contributions of the different spins can be separated by means of a Fourier transformation of the time-dependent signal (which relies on the mathematical conclusion that any time-dependent function is equivalent to a frequency-dependent function) [2–7].

With the application of one pulse to the sample, we obtain a one-dimensional NMR (1D-NMR) (Fig. 5.2b, top panel) spectrum, which in graphic terms is, as in any other spectroscopy, a plot of absorption intensity *versus* δ (or ν). As the molecular size increases, the number of nuclei is larger and, therefore, a large overlapping of δs among the different nuclei will occur when collecting a 1D-NMR spectrum, just because the magnetic environment around some of them will be similar, by chance, even if they are located in different parts of the molecule. Therefore, if we want to identify (and isolate) the δ (or ν) of each particular nucleus, we must try to separate those overlapping frequencies for the nuclei which are found in a similar environment. If a series of pulses are applied to the sample, we can split those overlapping frequencies in two (or more) axes, increasing the spectral resolution; in that way, we obtain a multidimensional spectrum, in which absorption intensity is plotted as a function of two (in 2D-NMR) or more (n-dimensional NMR) δ (or ν) values (Fig. 5.2b, bottom panel). A multidimensional spectrum can be used to monitor the interactions of nuclei through bonds (obtaining J values) or through space (obtaining NOEs). The through-bond (or through-space) interactions can be homonuclear (for instance, ^{1}H-^{1}H) or heteronuclear (such as ^{1}H-^{15}N). I shall describe briefly in the next paragraphs two types of homonuclear 2D-NMR experiments; a clear description of some of the most important 2D-NMR experiments used in structure determination can be found in [2–8].

The simplest 2D-NMR spectrum (a homonuclear one called 2D-COSY) is obtained by applying two pulses before observation of the FID during the time t_2 (Fig. 5.2c). The time t_1 that separates the two pulses (Fig. 5.2c, top panel) is varied in successive experiments, and as a result the FID observed during t_2 will change, since the spins will have evolved differently during that small increment in t_1. That is, in the end, we have k (the number of times we increase the time t_1) 1D-NMR experiments, where in each of them, a parameter, t_1, is modified. Therefore, the FIDs obtained are a function of two times: t_1 and t_2. A Fourier transformation of this matrix leads to a 2D-NMR spectrum in which resonances are obtained as a function of two frequencies ν (or chemical shifts δ), or "dimensions" (Fig. 5.2c, bottom panel). In the end, we have a contour plot-level (analogous to those used in the weather forecast) where we have the different resonances of protons split in two dimensions (in two axis, as in a Cartesian coordinate map). In the resulting spectrum, each resonance from the 1D produces a correlation with itself, leading the so-called diagonal-peaks (because there is a diagonal through the two opposite vertices of the contour map); these peaks do not yield useful information, since they provide the same information as in the 1D-NMR spectrum. On the other hand, we have off-diagonal cross-peaks, which provide the useful information, since they correlate nuclei (protons, in the 2D-COSY) which are related by (in this case) the covalent-bonds between them (Fig. 5.2c, bottom panel).

On the other hand, a homonuclear 2D-NOESY spectrum consists of the addition of a third pulse, after the last pulse of a COSY, and from which it is separated by a fixed time (mixing time) maintained throughout the experiment (Fig. 5.2b bottom panel); during this time, the spatially close spins are allowed to interact through the NOE (see Sect. 5.2.5). We can also design, for instance, heteronuclear spectra, where instead of observing in both dimensions the same nucleus (the ^{1}H), we observe in the t_2 dimension (the horizontal axis) the ^{1}H and in the t_1 (the vertical axis) another type of nucleus (such as ^{15}N or ^{13}C). With biomolecules, these heteronuclear experiments (especially, one called HSQC, which shows the correlations between ^{1}H (horizontal axis) and ^{15}N (or ^{13}C) (vertical axis)) are useful to map biomolecule-biomolecule interactions (see below), but it must be kept in mind that due to the low sensitivity of the NMR technique, and the low natural abundance of the latter NMR-active isotopes, the biomolecule must be isotopically enriched in those nuclei.

The key to 2D-NMR spectroscopy is having a time (t_1) which is varied in each of the 1D-NMR spectra acquired (Figs. 5.2b, c), during an observation time, t_2. Therefore, if instead of varying a single time (t_1), we modify two times t among the pulses, we have 3D-NMR. In 3D-NMR we have "cubes" with diagonals between two opposite vertices, and there are off-diagonal, cross-peaks, which contain useful information. To obtain a good resolution, the 3D-NMR experiments require the use of labelled samples. The 3D-NMR spectroscopy is widely used in the assignment (see Sect. 5.3) and structural elucidation of biomolecules. If we modify three times t among the pulses, we have 4D-NMR spectroscopy, and so on for n-NMR spectroscopy. In most of the examples described in the literature, the use of 2D and 3D-NMR spectra with double (^{15}N and ^{13}C) or even triple (^{15}N, ^{13}C and ^{2}H) labelled samples is enough to obtain the structure of a biomolecule.

5.3 Determination of Biomolecular Structures by NMR Spectroscopy

5.3.1 Sample Preparation

Most NMR experiments are carried out in H_2O (to which a 5–10 % of $^{2}H_2O$ is added to monitor the spectrum of ^{2}H during the long acquisition times; see [3–5] for a detailed explanation of such addition, which is beyond the scope of this chapter). One major exception are those experiments aimed at describing the hydrogen-bond scaffolding in a molecule (see Sect. 5.3.2). In some cases, especially when working with peptides or protein fragments, organic solvents are used (such as trifluoroethanol or di-methyl sulfoxide). On the other hand, in proteins associated with membranes or other lipids, the use of detergents, such as sodium dodecyl sulphate (SDS), is required, and in some cases conventional high-resolution NMR methods cannot be used: special techniques as SSNMR can, then, be used to gain structural insight.

As NMR is an insensitive spectroscopic technique, experiments aiming at solving molecular structures tend to be performed at concentrations of about 1 mM (for current spectroscopic configurations). Therefore the biomolecule must be soluble (*i.e.*, monodisperse) at those concentrations, and stable enough during the long acquisition times. The sample is introduced in a tube (made of special, non-paramagnetic substances [3–5]), usually 5-mm-wide, which must be free of contaminants or other paramagnetic compounds. If the biomolecule is stable, NMR spectra tend to be acquired at low pH, since the hydrogen-exchange of amide protons (on which is based the assignment process, see Sect. 5.3.2) with water is diminished; the buffer must not contain any signal which could obscure any of the resonances of the sample (and then, deuterated buffers are usually employed in homonuclear (^{1}H-detected) experiments). The presence of a high ionic strength can decrease the signal-to-noise ratio, due to the presence of local magnetic fields caused by the ions present in solution. Finally, as the resonance linewidth decreases (and therefore, resolution increases) with temperature, the NMR experiments are acquired at (moderately) elevated temperatures (provided that the biomolecule does not unfold or aggregates).

5.3.2 Structure Determination

The steps in determining the structure of a protein or nucleic acid by NMR follow a straightforward logic. We assume that we know the primary structure of the biomolecule, and we want to obtain the secondary and tertiary structures. There are two ways of assigning a biomolecule: (i) by using the natural abundance of ^{1}H (the homonuclear procedure), which relies on the use, mainly, of 2D-NMR spectroscopy; and (ii) by using labelled samples (the heteronuclear approach), which relies in the use of 3D (or even larger dimensions)-NMR spectroscopy. The difference relies on the type of experiments used to assign the spectra, since the restrictions required to calculate the structure are basically the same.

Although there are many variations to the assignment process (which are beyond the scope of this chapter, and which can be consulted by the interested reader in [8, 9]) either in the homo- or heteronuclear approaches, this process can, in essence, be resumed in a series of steps. First, the multidimensional NMR spectra obtained must be *assigned*. This means that every intensity peak at a defined δ in the spectra must be identified as belonging to one particular nucleus in the molecule.

In the homonuclear approach assignment consists, in a first stage, on the identification of the different types of residues (Fig. 5.3). Using an analogy, during this first step we shall assign a *name* (*e.g*., Asn, or any other amino acid type) to each different set of signals related with the same amide proton observed in a TOCSY experiment (a more sophisticated type of COSY experiment); but, probably, we shall not be able to put a *surname* to each set of signals assigned to a *name* (amino acid type). In other words, we shall not be able to say whether that assigned residue in the TOCSY spectrum, with the features of an Asn is Asn25, Asn72, Asn80 or Asn100 in the 200-residues long protein we are studying. Then, in a second step, we will assign a *surname* to each set of signals with an already assigned *name*, with the acquisition of a NOESY spectrum, which allows the identification of the connectivity chain through space (see Sect. 5.2.5). Therefore in the homonuclear approach, at the same time we are “naming” and “surnaming” each of the residues of the polypeptide chain, we can conclude, from the pattern of

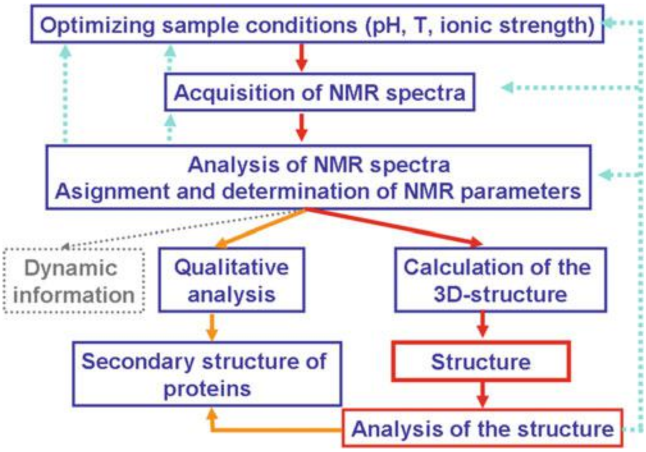


Fig. 5.3 Solving structures by NMR. Flow chart indicating the different steps in the determination of the three-dimensional structure of a protein by NMR

observed NOEs, the type of secondary structure in which the assigned residues are located [2, 5, 7–10] (Fig. 5.3). It is important to note that either the "naming" or "surnaming" steps are mainly based in the presence of amide protons.

In a heteronuclear approach, the assignment step does not involve the use of NOESY spectra [5, 8–10], as a stage required in assignments, since assignment relies on the determination of J values for ^{1}H with ^{13}C, ^{1}H with ^{15}N, ^{13}C with ^{15}N, and ^{13}C with ^{13}C nuclei; thus, no ambiguities are introduced, and the assignment process can be relatively easily computerized [9]. In this second approach, NOESYs are used in a late stage, to confirm the through-bond assignment and to provide the data set of restraints (distances between nuclei) for the next step in the structure solving protocol [9] (see below). As it happens with the homonuclear approach, the assignment of the resonances is based (although not exclusively) in the presence of the amide protons.

Once the spectra have been assigned, a number of parameters for the different identified nuclei can be obtained (see Sect. 5.2). The most important are the chemical shifts (δ), coupling constants (J), peak intensities, NOEs and, for dynamic studies, the relaxation times (in some cases, even for the heteronuclei). In particular, distances between atomic nuclei (the so-called distance restrictions) can be obtained from the NOEs, and dihedral angles between bound nuclei can be obtained from J values. Probably, the most important restriction set to obtain the structure is that provided by the distances, which make the NOESY spectra a key component of the pulse-program tool-kit of every NMR structural determination.

There is another final parameter which is also used in structural determination, because it provides information on the hydrogen-bond scaffolding of the biomolecule (secondary and tertiary structure in proteins, and hydrogen-bonding pattern in nucleic acids). For clarity, I shall focus on proteins in the next lines. In proteins, the amide hydrogens, and other hydrogens bound to nitrogens (as in the side chains of

Lys, Arg, Asn or Gln) are weakly bound due to the hybridization of nitrogen and its electronegativity, as are the hydrogens in the water molecule. Therefore, the N-H bonds (in proteins) and the O-H bonds (in water) can break apart due to the thermal fluctuations, and can immediately be captured either by an oxygen of other water molecule or by a nitrogen of the protein, leading to *hydrogen-exchange* between protein and water. If the protein is dissolved in ^2H_2O instead of H_2O, we should observe an exchange between the labile ^1H of the protein and the ^2H of the solvent. Furthermore, since the resonance (ν) of the ^2H is not the same as that of the ^1H, if we are observing the latter nucleus, we shall not be able to observe the former by NMR; therefore, acquiring spectra at different times after dissolving the biomolecule in ^2H_2O, we shall observe the disappearance of the ^1H, involved in N-H bonds, in those 2D-NMR spectra. However, not all the ^1H will exchange with the same easiness, since those hydrogen-bonded will be more "reluctant" to exchange with the solvent. For structure determination, the information on which protons are more "reluctant" to disappear in the presence of ^2H_2O is extremely worthy, but the speed of exchange also provides information on the stability of the exchanged N-H hydrogen-bonds. Thus, the pattern of exchange of ^1H with ^2H will be hydrogen-bond- and protein-stability dependent, and the number of incorporated ^2H nuclei will allow us to have an estimation of the overall protein stability. A similar approach is used to measure the stability of viruses capsids by using mass spectrometry (see Chap. 6), since the mass of ^2H is roughly twice that of a ^1H, the number of incorporated ^2H nuclei can be easily determined within the precision of current mass spectrometers.

It is important to note that I have described here the structural restrains most frequently used in NMR, but there are other parameters, which are used in partially folded proteins or intrinsically disordered proteins, that are beyond the scope of this chapter. The interested reader can have a wider look in [5, 7, 9].

To produce a structural model of the macromolecule by NMR, we have to find the conformation(s) that is (are) consistent with the experimentally obtained restraints (distances, dihedral angles and hydrogen-bonds) (Fig. 5.3). Usually, these restraints are incorporated into potential energy functions to calculate the molecular dynamics and minimum energy structures [9]. Structures with the fewest violations of the NMR restraints are found by this computing-intensive process. If there are enough restrictions only one structure (or set of very similar structures) can be obtained; with fewer restraints a number of somewhat different structures is found, all of which are consistent with the experimental data. During this computing process, the obtained structures are used to refine the restrictions, and to reassign many of the NOEs observed in the NOESY spectra, which could not be unambiguously determined before (Fig. 5.3, dotted line). The final result is an ensemble of refined, similar structural models that resemble the true minimum energy three-dimensional structure of the macromolecule. Sometimes these models are averaged to provide a representative model of the entire ensemble. Those structures will appear with well-defined regions (where usually there is a large number of structural restrictions *per* amino acid or base) and other zones where the backbone of the protein or the nucleic acid will appear less restricted. These ill-defined regions may correspond to zones with a high mobility, and as such they are frequently involved in molecular recognition processes; but also, these ill-defined regions might be so due to the absence of local and/or long-range restrictions

5.4 NMR Structures of Viral Macromolecules

In this section, I shall describe several examples of the use of NMR spectroscopy to solve the structure and investigate the dynamics of viral proteins and nucleic acids involved in defining the structure of virus particles (the so-called "structural biomolecules"). A number of proteins from different viruses have been studied by NMR. However, preferred targets have included structural proteins and nucleic acid domains from the human immunodeficiency virus type 1 (HIV-1) or other retroviruses, mainly due to their interest as targets for the development of new antiretroviral therapies (see Chap. 20). In this section, I shall summarize a number of structural studies on retroviral proteins and nucleic acids as case studies to show what NMR can reveal about virus structure. In addition, a few examples on proteins from other viruses will be used to illustrate some other applications of NMR in structural virology.

5.4.1 NMR Structure of Viral Proteins

The main protein components of retrovirus particles are synthesized as three polyproteins that produce the internal structure of the virion (Gag), the viral enzymes (Pol) or the glycoproteins of the virion envelope (Env). The Gag polyprotein is required and sufficient for virus particle assembly and budding. The Gag proteins capture the viral RNA and assemble either in the cytosol (B- and D-type retroviruses) or at the cell membrane (C-type, human T-cell leukaemia virus (HTLV), bovine leukemia virus (BLV), and lentiviruses) to form a RNA-containing spherical capsid in the immature virion. During virion maturation, processing of the HIV-1 Gag polyprotein by the viral protease generates the matrix (MA), the (mature) capsid (CA), the nucleocapsid (NC) and the p6 proteins.

The Capsid Protein of Retroviruses

In HIV-1, the CA domain mediates Gag-Gag interactions that are critical for immature particle formation [11]. HIV-1 CA is also known to interact with a variety of host cell factors, such as cyclophilin A, a host restriction factor and human lysyl-tRNA synthetase (hLysRS) [12]. CA is a two-domain protein, composed of largely helical N-and C-terminal domains (NTD and CTD, respectively) separated by a flexible linker [13, 14]. Upon proteolytic processing during virion maturation, CA is released from the Gag precursor, and rearranges to form a conical capsid that surrounds the viral genome and associated proteins [13, 15]. The capsid is later disassembled after virus entry into the host cell, to allow uncoating of the viral genome. Mutations that impair CA assembly or alter its stability lead to dramatic reductions in viral infectivity. Assembly and stability of CA is dependent on inter- and intra-molecular interactions that involve the NTD and/or CTD of CA [14, 16]. The NTD (CA residues 1-146), which is a monomer in solution, forms hexamers in the virion capsid; the CTD (CA residues 146-231), which is dimeric in solution, links adjacent hexamers [14]. The atomic structure of CA of HIV-1 has been studied by both X-ray crystallography and NMR using a "bottom-up" approach, based on solving first the structures of the isolated domains. Recently, the NMR structure of the full-length CA has been solved by NMR, confirming the previous results with the isolated domains.

The structure of the HIV-1 NTD. The structures of the HIV-1 NTD determined by NMR and X-ray crystallography ([17] and references therein) are in good agreement. NTD consists of seven α-helices, two β-hairpins, and an exposed, partially ordered loop (Fig. 5.4a). The domain is shaped like an arrowhead; the β-hairpins and loop project from the trailing edge of the arrowhead and the carboxyl-terminal helix projects from the tip. This structure is different to those of other RNA virus coat protein structures [17]. At its N terminus there is a terminal proline, which forms a salt-bridge with a conserved, buried aspartate residue (Asp51). This salt-bridge cannot exist in the Gag precursor polyprotein, and it has been proposed that the β-hairpin itself is important for triggering condensation of the core particle. The binding site for cyclophilin A is located on the exposed loop and encompasses the Pro90, which in the free monomeric domain in solution adopts both *cis* (absent in the X-ray structures) and *trans* conformations due to the loop flexibility. In the X-ray structure of the complex with cyclophilin A, Pro90 adopts a *trans* conformation [17].

The structure of the HIV-1 CTD. The structure of the HIV-1 CTD has been also determined by NMR and X-ray crystallography. CTD is composed of a short 3_{10}-helix, a highly conserved sheet-turn-helix element called the major homology region and four helices (8–11, in the numbering of the intact CA protein) (Fig. 5.4b) which are connected by short loops or turn-like structures. The dimerization interface of CTD is formed by the mutual docking of α-helix 9 from each monomer (with the side chains of each tryptophan (Trp184) deeply buried in the dimer interface), and hydrophobic interactions between the 3_{10}-helix of one subunit and α-helix 9 of the other. Although there are small variations in the crystal and solution structures of dimeric CTD (see next paragraph), the basic lattice-stabilizing dimer interactions in the mature capsid of the virus are the same as in the solution dimer interface.

Although the X-ray and NMR structures of active dimeric CTD reveals basically the same structure within each monomer (the NMR structure was solved in 2009, and the X-ray one in 1996), there are small variations in the structure of α-helix 9 ([14] and references therein). These variations occur at tertiary level (packing of α-helix 9 against other helices within the same monomer), and at quaternary level (the angle formed by the α-helix 9 across the dimer dyad). Furthermore, in solution

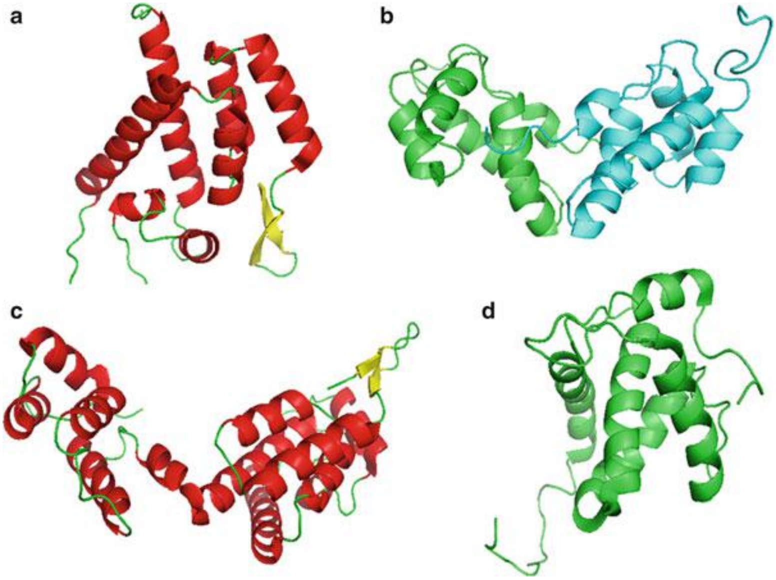


Fig. 5.4 Structures of some HIV-1 proteins solved by NMR. (**a**) Isolated NTD monomer (PDB number 1GWP) [17, 18]. (**b**) Isolated CTD dimer (*monomers are shown in green and cyan*) (PDB 2KOD) [19]. (**c**) Full-length monomeric CA mutant (PDN number 2LF4) [20]; the NTD and CTD are respectively on the *right* or *left side* of the figure. (**d**) The un-myrystilated MA (PDB number 2H3F) [21]. Images in this and following figures were produced with PyMol [22]. In panels (**a**) and (**c**) helices are shown in *red*, β-sheets in *yellow* and loops in *green*

CTD appears to have two conformers, whose population is determined by the protonation of Glu175, which is involved in the dimerization interface [23]. In addition, structurally different dimers have been characterized: namely the one described above, and a domain-swapped arrangement involving mainly the MHR [24]. Somewhat different structures of mutant, monomeric species of CTD have been also determined by NMR. The above and other results, taken together, reveal a high structural plasticity of the CTD that may be biologically important for capsid assembly, maturation and function [14].

Structure of the intact CA proteins of some retroviruses. The structure of the full-length HTLV type I CA and Rous sarcoma virus (RSV) CA have been determined by NMR [25, 26]. As for HIV-1 CA, the proteins contain two independent domains, connected by a flexible linker. Some minor structural differences between the CTDs may account for the fact that HIV-1 CA is dimeric, whereas HTLV-I CA and RSV CA are monomeric. NMR studies of these proteins show also that the two domains are structurally independently, with no inter-domain NOEs; furthermore, the dynamics of the whole domains (as measured by T_1-and T_2- relaxation times) suggest that both regions are also dynamically independent.

The structure of full-length monomeric HIV-1 CA (with an inactive dimerization interface) has also been solved by NMR (Fig. 5.4c). Although the overall structures of both domains are similar to those reported for the isolated regions, the CTD in whole CA shows differences in the dimerization helix [20]. In the crystal and NMR structures of dimeric CA, the dimerization helix (helix 9), is longer and shows a bend, whereas in the monomeric mutant the α-helix 9 is shorter and straight; this observation supports previous results that suggest substantial conformational rearrangements during CA dimerization ([14] and references therein). The NTD and CTD are connected by a five-residue linker, formed by residues 145–149 (Fig. 5.4c). As observed for other CA structures, the linker is highly flexible, as evidenced by the random coil δs for the central three residues, and the T_1-and T_2-relaxation data. This high interdomain flexibility may be important for facilitating the docking of other Gag domains during immature capsid assembly [27] and allowing the intersubunit structural variations observed in mature retroviral capsids [14].

The Matrix Protein of Retroviruses

The MA protein has similar roles in all retroviruses, directing membrane localization of the assembling viral particle, and after virus maturation, forming a stable structural shell associated with the inner surface of the mature viral lipid envelope ([14] and references therein).

The three-dimensional structures of the HIV-1, BLV and HTLV type II MA proteins have been determined by NMR ([17] and references therein); there are also X-ray structures of the simian immunodeficiency virus (SIV) and HIV-1 MA [17]. The HIV-1 MA consists of five α-helices, two short 3_{10}-helical stretches and a three stranded mixed β-sheet (Fig. 5.4d shows the NMR structure, where only the helices are observed). Close to the C terminus of the second helix there is a putative membrane binding surface, which could anchor the protein to the membrane surface, in concert with an N-terminal myristoyl group. HIV-1 MA is a monomer in solution but forms trimers in the crystal; the HIV-1 MA trimers may be also observed in lower-resolution EM images and could serve as a fundamental building block for formation of the MA shell within the mature virion. The individual proteins of the trimer are arranged to create a basic surface, which interacts directly with the acidic inner layer of the lipid ([17] and references therein). Comparison of a monomer in the crystallographic trimer with the NMR structure of the monomer in solution shows a displacement of the 3_{10}-helix, suggesting that this structural element could act as a “hinge” in virion assembly and disassembly; this resembles the rearrangements proposed for CA dimerization and, again, underscores the importance of protein flexibility in retrovirus morphogenesis. It should be interesting to explore the dynamics of that helix by measurements of T_1 and T_2-relaxation times by NMR, which may serve to unambiguously characterize MA flexibility.

The Nucleocapsid Protein of Retroviruses

NC is the domain of the Gag polyprotein that directs genome packaging. Except for the spumaviruses, all retroviral NC proteins contain one or two copies of a conserved CCHC motif (Cys-*X*2-Cys-X4-His-X4-Cys; where X = non-Cys/His residue), that are Zn-binding sites (the so-called "zinc knuckle"). In HIV-1, NC does not form direct lattice stabilizing interactions, although it contributes to particle formation by tethering Gag to the RNA genome (see [14] and references therein).

NMR structures are available for synthetic peptides with sequences encompassing the HIV-1 NC (the zinc knuckle domains), and for intact NCs from HIV-1, the Moloney murine leukemia virus (MoMuLV) – which contains a single CCHC "zinc knuckle" -, the mouse mammary tumour virus (MMTV), and the monkey virus (MV) [17, 28]. In the HIV-1 and MoMuLV NC proteins, the CCHC zinc knuckle domains adopt similar folds. The N-terminal residues form a metal-coordinating reverse turn called a "rubredoxin knuckle" due to its similarity to metal-coordinating substructures observed in the iron domain of rubredoxin. Subsequent residues form a loop, which leads to a carboxyl-terminal 3_{10}-helix (see Fig. 5.5c for an example, also showing the interaction of NC with a nucleic acid). The two domains in HIV-1 NC show weak NOE inter-knuckle contacts between two bulky aromatic side chains, but dynamic studies (by measuring the T_1- and T_2-relaxation times) suggest that the two "zinc knuckles" behave independently (as it happens with the NTD and CTD in CA, see above). Thus, NC in solution may explore different conformations with interacting and non-interacting knuckle domains [28]. On the other hand, one of the "zinc knuckles" of MMTV and of MV shows a conformational transition involving an additional C-terminal β-hairpin (albeit with different orientation in both proteins). Therefore, the NMR studies suggest that different structural features of the zinc knuckles in diverse retroviruses may facilitate the packaging of the different genomes.

Intrinsically Unfolded Proteins

Here, I shall briefly describe one of the major contributions of NMR to the field of virus structure: the identification and characterization of proteins which are nearly or completely unfolded. Such proteins cannot be crystallized and are not amenable to X-ray structure determination. Currently, NMR is the only technique providing atomic-level insights into this special class of proteins.

The p6 protein of HIV-1 is a docking site for several viral and cellular binding factors, facilitating virus budding [32], and regulating CA processing and virus core assembly [33]. The peptide is disordered in aqueous solution, but under hydrophobic solution conditions (*i.e.*, in the presence of trifluoroethanol) that may mimic a lipidic environment, p6 adopts a helix-hinge-helix structure: a short α-helix (residues 14–18) is connected by a flexible hinge region to a longer helix (residues

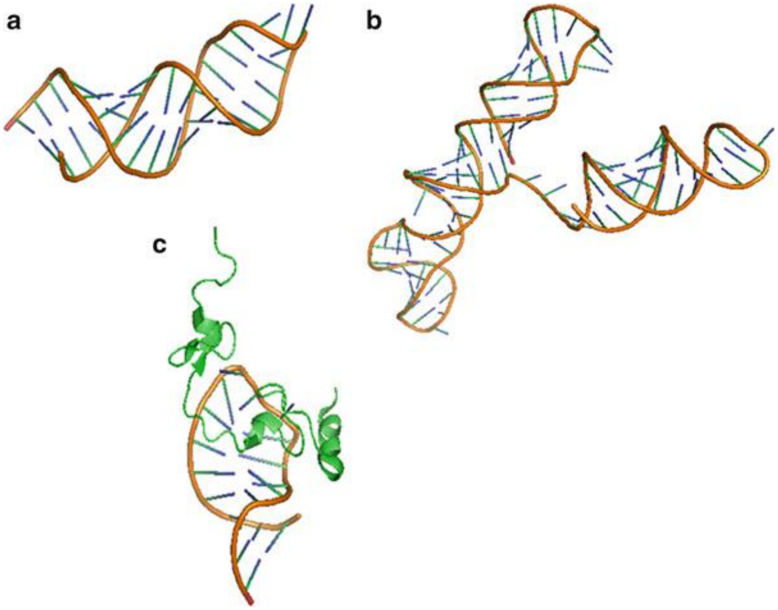


Fig. 5.5 NMR structures of isolated viral RNA or viral RNA-protein complexes. (**a**) Structure of the SL1 of the HIV-1 (PDB number 1N8X) [29]. (**b**) Structure of the MLV Ψ-RNA site (PDB number 1S9S) [30]. (**c**) Structure of the HIV-1 NC bound to SL2 of the RNA (PDB number 1FSU) [31]

33–44). The latter region binds in this conformation to a vacuolar protein sorting factor, and to the HIV-1 accessory viral protein (VPr) [34].

The Gag polyproteins of G-retroviruses contain a conserved p12 domain between MA and CA, which has key roles in virus assembly and nuclear transcription. A MoMuLV fragment, containing the p12 region and the NTD, is unfolded; even the β-hairpin at the N-terminus of the NTD region is unfolded [35]. The p12 region and the hairpin may fold during interaction with other molecules during capsid assembly.

Some non-structural viral proteins contain also substantial intrinsically unfolded regions. For example, the protein W from bacteriophage λ is required for DNA stabilization within the capsid (head) and, attachment of the tail onto the head during virion morphogenesis. The NMR structure of protein W consists of two α-helices and a two-stranded β-sheet arranged around a well packed hydrophobic core, whereas the C-terminal segment, which is essential for DNA stabilization, is unfolded [36]. The 117-residues-long bacteriophage λ FII protein, which is required for the joining of phage head and tail at a late step of phage morphogenesis, has a 3D-structure in solution formed by seven β-strands and a short α-helix. However, two large regions are unstructured: the N terminus (residues 1–24) and a large loop near the middle of the protein (residues 46–62) [37]. The authors suggest that both regions may fold during assembly. As a final example, NMR studies of the

Gn-membrane glycoprotein of hantaviruses show that the zinc finger domain is the sole folded polypeptide patch in the cytoplasmic tail of the protein; this region interacts with the ribonucleoprotein during virus assembly [38].

To sum up, the examples above described, and others found in the literature, suggest that large regions in some structural or non-structural viral proteins are intrinsically unfolded and only become folded upon interaction with other biomolecules. As for other intrinsically unfolded proteins, this lack of order entails an entropic penalty to be paid on binding the ligand; however, this strategy may allow a more promiscuous interaction with different ligands and, in some examples, it could increase the value of the kinetic association rate for complex formation.

5.4.2 NMR Structures of Viral Nucleic Acids

All retroviruses specifically package two copies of their genomes during virus assembly. Packaging is mediated by the Gag proteins, which in the absence of viral RNA, can incorporate cellular RNAs. However, viral RNA contains specific packaging signals (the Ψ-sites), which are preferentially packaged; it seems that genome selection proceeds, albeit not exclusively, *via* the direct binding of NC to those sites. The Ψ-sites are sequestered by intramolecular base pairing in the monomeric RNA, and then, become exposed upon dimerization, allowing high-affinity binding of the RNA dimer to the NC [39, 40]. Probably, a ribonucleoprotein complex formed by a small number of Gag molecules and two copies of the genome is directed towards the plasma membrane assembly sites, where other thousand additional Gag molecules localize and assemble to form the immature virus particle.

Many NMR structural studies on retroviral RNA have focused on the 5′-untranslated region where packaging signals reside. Several studies have shown that the 120 nucleotides upstream of the Gag start codon are required for efficient packaging. This nucleotide segment was proposed to form four stem-loops (the SL1-SL4), with different structures and different affinities for NC [29, 41, 42] (Figs. 5.5a, b). For instance, the NMR structure of the SL2 stem-loop reveals a “platform motif”, where the rings of the sequential U14 and A15 bases are approximately co-planar; this motif is thought to mediate long-range interactions that could stabilize the Ψ-RNA association or facilitate splicing and/or packaging [41]. A similar “platform motif” has been observed in the NMR structure of the 101-nucleotide core encapsidation signal of the MoMuLV, together with an “A-minor K-turn”, where unpaired bases of a bulge pack against the groove of a proximal stem [30].

Recently, the structure of the whole HIV-1 5′-leader RNA (the 712-nucleotide dimer) has been studied by NMR [43]. Residues spanning the Gag start codon (the AUG) form a hairpin in the monomeric leader and base-pair with the bases of the unique-5′ region in the dimer. This association causes dimerization by displacing and exposing a dimer-promoting hairpin, and enhances binding to NC. These and other results taken together suggest a mechanism in which translation, RNA dimerization, RNA-NC binding and RNA packaging within the assembled capsid are triggered by a common RNA structural switch. This switch would involve the shift of a conformational equilibrium between different conformations in the RNA molecule.

5.4.3 NMR Structures of Viral Protein-Ligand Complexes

Protein-Nucleic Acid Complexes

In RSV, the dynamics of an RNA segment in the free and bound states with a fragment of Gag (comprising CTD, NC and the in-between spacer region (SP)) has been described. Both NC and CTD regions show an independent dynamic behaviour. The contacts with the RNA are distributed along the three regions: one of the zinc fingers in NC seems to make key contacts with the RNA; the SP region experiences a slow conformational equilibrium in the absence of the RNA, which is restrained upon binding; and finally, the first and fourth helices in CTD are also affected by binding to RNA. These results suggest that NC-RNA binding disrupts intramolecular interactions between SP and CTD regions; this disruption enhances the accessibility of the Gag dimer contact surface, promoting intermolecular organization [44].

The structures of several stem-loop sites in the retroviral RNA in complex with the NC of HIV-1 have been determined by NMR [31, 42, 45] (Fig. 5.5c), both zinc knuckles bind to exposed guanine bases of the tetraloop, whereas basic residues at the N terminus of NC pack against the closer "zinc knuckle" being near to the RNA stem; in addition, the N-terminal "zinc knuckle" interacts with the "platform motif" (see above). Some of these interactions are similar to those observed in SL3, but others are substantially different; for instance, the NC interacts with the major groove of SL3, and the orientations of the two zinc domains are different in both complexes [31].

However, not all retrovirus bind with a relative high affinity to the individual SL of RNA; for instance, in MoMuLV, binding of NC only occurs to the three SLs (with a nM affinity) suggesting a high specificity towards a region, which is only present with the intact three loops [42]. The NMR structure of the MoMuLV-NC in complex with the 101-nucleotide core encapsidation segment of the MoMuLV Ψ-site suggests a network of interactions which promote specific binding by the sole "zinc knuckle" of the NC.

Taken together, the structural studies of retroviral RNA-NC complexes reveal that genome recognition and NC-RNA binding differ considerably among the different viruses [39, 43]. Furthermore, these data suggest that the mechanism of genome recognition/packaging involves one or more NC domains of several Gag molecules with multiple Ψ-sites stem-loops [42].

Finally, it is interesting to note that NMR have been used also to detect and map new interactions between viral proteins and nucleic acids without a need to solve the structure of the complex. For instance, MA-DNA interactions have been

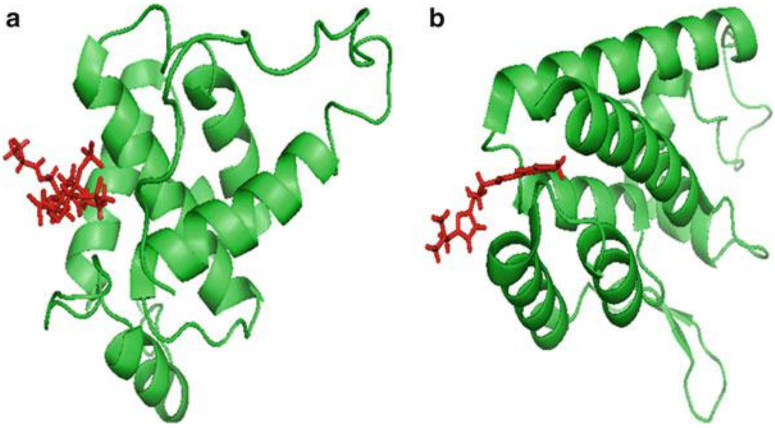


Fig. 5.6 NMR structures of the complexes of small ligands with viral proteins. (**a**) NMR structure of the myristylated HIV-2 MA in complex with the PIP2 lipid (PDB number 2K4I) [48]. The PIP2 lipid moiety is shown in *red sticks*. (**b**) NMR structure, obtained by joint refinement with X-ray data, of the CAP-1 inhibitor in complex with CA of HIV-1 (PDB number 2JPR) [49]. The CAP-1 compound is shown in *red sticks*

detected proved recently to interact with DNA by NMR, but the structure of the complex has not been solved. Instead the DNA-interacting region has been determined by identifying changes in the chemical shifts of the assigned resonances of the protein upon addition of nucleic acid by using heteronuclear HSQCs [46]. This is an example of a more general technique, which allows an easy and very fast detection of protein-ligand interactions in high-throughput screening methods (since it only requires the protein to be labelled with ^{15}N).

Protein-Lipid Complexes

NMR has been used also to study interactions between retroviral MA and lipids. Interactions of MA from equine infectious anaemia virus (EIAV) or HIV-1 with phosphatidylinositol-4,5-diphosphate (PIP2) have been described [47]. The interaction between PIP2 and HIV-1 MA promotes membrane binding and induces conformational changes at the N-terminus of MA (exposing the myristate moiety). The binding region involves the loop between the first and second helices, and the cleft between the second and the fourth helices. In un-myristylated EIAV MA, the binding site of the lipid includes the loop between the second and third helices, although there are conformational changes (as measured by the T_1- and T_2-relaxation times) which may trigger MA trimerization. The structure of myristylated MA of HIV-2 in complex with PIP2 has also been solved [48] (Fig. 5.6a); as it happens with the HIV-1 MA, the myristyl group of HIV-2 MA is partially sequestered within a narrow hydrophobic tunnel formed by side chains of helices 1, 2, 3, and 5, although the myristate of HIV-2 MA is more tightly sequestered than that of the HIV-1 protein; the addition of PIP2 induces minor changes in the structure of HIV-2 MA. These NMR studies help to understand MA-envelope interactions in retroviruses ([14] and references therein).

Protein-Peptide and Protein-Organic Molecule Complexes

The CA of HIV-1 is being considered as a new therapeutic target (see Chap. 20). To help in understanding how binding to CA of certain peptides and organic molecules may inhibit the assembly of the HIV-1 capsid, the structures of complexes between some of these compounds and CTD of HIV-1 have been solved by NMR ([50] and references therein). Next, I describe two examples.

In the first example, NMR was used to investigate the interactions of peptide CAI, and a stapled derivative (conformationally restricted by a covalent bond between residues that are non-contiguous in the sequence) with a monomeric mutant of CTD. Binding of these peptides to CA results in the inhibition of HIV-1 capsid assembly, and in virion infectivity [50]. The binding interface of CAI with CA was mapped by the changes of chemical shifts of ^{15}N-labelled CA in a HSQC (as it happened with the binding between DNA and MA, see above) ([51, 50] and references therein). The peptides bind in a hydrophobic pocket of CTD delineated by residues of the four helices in the domain, and it adopts a helical conformation where its hydrophobic side chains make extensive contacts with specific hydrophobic patches in CTD. The structural results suggest that binding of these peptides sterically interferes with CTD-NTD interactions in the hexameric rings that form the capsid, and distort the CTD dimerization interface. They also support, once more, the structural plasticity of CTD ([50] and references therein).

The molecule N-(3-chloro-4-methylphenyl)-N´-{2-[({5-[(dimethylamino)-methyl]-2-furyl}-methyl)-sulfanyl]ethyl}-urea) (CAP-1) binds HIV-1 CA and inhibits capsid assembly and virus infectivity. In this example, both X-ray and NMR studies show that CAP-1 binds in a pocket delineated by parts of helices 1, 2, 4 and 7 in NTD [49] (Fig. 5.6b); this region is spatially close to the groove created by α-helices 8 and 9 (the dimerization helix) of CTD. The residue Phe32 of NTD pulls away from its position, creating a large cavity where the aromatic ring of CAP-1 inserts making hydrophobic contacts with other residues in the protein [49]. One general conclusion of these studies is that assembly of the HIV-1 capsid can be inhibited by small molecules (peptides or organic compounds) hydrophobic enough to bind small pockets in CA, to induce local structural distorsions and to hamper CA-CA interactions ([14, 50] and references therein; see also Chap. 20).

This study of CAP-1 in complex with CA also underscores an important aspect of NMR. The joint use of this spectroscopy with other biophysical techniques, such as X-ray crystallography (in the above example) or EM, has led in the last years to promising achievements. For instance, the NMR structure of a viral capsid protein can be combined with the lower-resolution cryo-EM structure of the complete capsid to obtain a pseudo-atomic model. This approach was used in a detailed structural study of the mature HIV-1 capsid that revealed a small protein-protein interface that had been previously not detected [19]. In another example, the structure of a retroviral RNA packaging element was jointly solved by NMR and cryo-electron tomography [52].

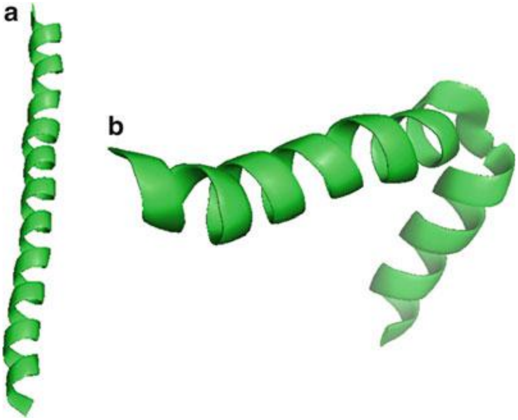
5.4.4 Structure and Dynamics of Viral Particles by Solid-State NMR Spectroscopy

The study of intact viral particles and complexes between viral particles and other viral or cellular macromolecules offers a challenge for solution-state NMR and, in some cases, even for X-ray crystallography (see Chap. 4), since they are very large in size, may be not soluble enough, and may prove difficult to crystallize. There have been approaches to tackle the structure of whole virus capsids by NMR, but in all cases, only information about highly dynamic regions of the protein, forming the corresponding capsid, was obtained [53, 54].

In addition to the above approaches using NMR in solution, SSNMR spectroscopy is emerging as a promising technique to study the structure and dynamics of viral particles because it is not limited by solubility, molecular size or lack of long-range order. Major difficulties encountered for obtaining NMR structures of large macromolecules and supramolecular assemblies come from the short T_2-relaxation time, and large linewidth (see Sect. 5.2.4) or, in other words, from the correlation time of the molecule (which is the time used by the molecule in rotating ~57°). The ability of SSNMR to replace molecular motions with radiofrequency irradiations are key to obtain spectra of large molecular assemblies, or aligned molecules within membrane-like environments. The physical details of these methods are out of the scope of this chapter, and the interested reader is referred to recent literature [55–57].

Filamentous bacteriophages are rod-shaped particles encasing a single-stranded circular DNA, within a cylindrical capsid. This capsid consists almost entirely of several thousand copies of the major capsid protein. The virions are not enveloped, but their life cycle involves interactions with the host cell membrane, and the capsid protein also acts as a membrane protein. The best-studied filamentous bacteriophage is fd (class I), which infects *Escherichia coli*. Its capsid is expressed as a procapsid protein, with a 23-residue sequence. *In vivo* the procapsid is inserted into the cell membrane becoming the most abundant membrane protein in infected cells; it is formed by a transmembrane helix connected by a short loop to an amphipatic helix. During virus assembly, the capsid is extruded from the membrane and the DNA is packaged within the capsid. The structure of the capsid protein in filamentous bacteriophage particles was determined by SSNMR [55, 58] (Fig. 5.7a). The first five residues are unfolded and highly mobile. After the Pro6-Ala7 patch the capsid protein is a continuous ideal straight helix between residues 8–38, although it is not a single curved helix: near residue 39 (in the middle of the hydrophobic

Fig. 5.7 Protein structures solved by SSNMR. (**a**) Structure of a fd capsid protein subunit in intact bacteriophage particles (PDB number 1NH4) [58]. (**b**) Structure of the membrane-bound species of the fd capsid protein (PDB number 1MZT) [59]



helix), there is a kink, causing a change in the helix rotation. Interestingly, the structure of the capsid protein in the virion is not identical to that determined by X-ray fiber diffraction. In the membrane-bound protein, the helix experiences a substantial rearrangement around residue 21 (which was involved in the straight helix in the bacteriophage), that accompanies assembly. The rotation of the trans-membrane helix (residues Gln16-Ala46) around its long axis changes by 16° to obtain the proper alignment for packing in viral particles (Fig. 5.7b) [60].

To allow for a comparison among different bacteriophages, the capsid protein of Pf1 (Class II), which infects *Pseudomonas aeruginosa* has also been studied by SSNMR [55]. The capsid protein of Pf1 is similar to that of fd, with a similar distribution of charged and hydrophobic residues. In contrast to the structure of fd, the six N-terminal residues of Pf1 form a hook; this structure disagrees with the results obtained by X-ray fiber diffraction studies: in fact, multiple chemical shifts were observed for the side-chain and alkyl backbone atoms of Thr5, whereas the rest of the residues had a single chemical shift (suggesting a unique conformation for the 7,300 subunits in the 36 MDa virion [61]. Furthermore, residues around Gln16 (where the hydrophobic helix starts) show a complex dynamic behaviour, with evidence of a bulge (also in disagreement to the X-ray fiber diffraction studies). This behaviour is clearly different to that observed in the capsid protein of fd; instead of a hinge around residue 21, a straight helix was detected.

The dynamic behaviour of the membrane-bound coat protein is also clearly different in Pf1 and fd [55, 60]. For instance, the Pf1 capsid shows a temperature-dependent transition [55, 62]; this transition leads to structural changes in the interfaces between neighbouring capsid subunits, adjusting the hydrophobic interfaces between fairly rigid subunits (as shown by the changes in chemical shifts). Moreover, SSNMR studies of ^{31}P of the DNA backbone suggest that the DNA is packaged differently: in fd the DNA is packaged randomly, while in Pf1 it is uniformly oriented [54]. Therefore, the SSNMR suggest that both bacteriophages are structurally similar but not identical.

SSNMR has also been used, in conjunction with solution-state NMR, for structural studies of the mature HIV-1 capsid. The capsid structure can be described as a fullerene cone made up from a lattice of CA hexamers containing a number of pentamers at defined positions (see [14] for details). SSNMR studies have shown that both open tubes made only of hexamers and the authentic conical capsid are conformationally homogeneous (even allowing partial assignment of the chemical shifts of several residues in the conical capsid). The structure of CA is not different in both morphologies [16, 23]; furthermore, the structure appears to be similar to that in solution. The linker region connecting the NTD and CTD was, again, found to be very flexible (as shown by the T_1- and T_2-relaxation times) (see Sect. 5.4.1) As already mentioned, this flexibility allows the protein to access multiple conformations essential for assembly of the pleiomorphic retroviral capsid.

5.5 Understanding Viruses: Some Major Contributions of NMR Spectroscopy

In this chapter, I have described how NMR spectroscopy can face the challenging problem of solving the structures of large viral proteins, viral particles and other viral complexes and assemblies by following a “bottom-up” approach. This technique has been used to solve the structure of a number of isolated domains of different viral proteins. The results have generally provided new clues on the relationships between protein structure and viral function, even in cases where the proteins had also been solved by X-ray crystallography. NMR has been essential to reveal the conformational dynamics of several viral proteins, and even of complete viral particles. NMR spectroscopy has shown the presence of alternate conformations, delineating the high intrinsic mobility of viral protein domains, and how such dynamic behaviour contributes to virion morphogenesis. Furthermore, NMR has been revealed as an unmatched technique to show, at atomic level, the natively-unfolded character of some viral proteins. The intrinsic mobility and/or the unfolded character of many (if not all) regions of some viral proteins seem to be landmark features of these viral biomolecules. Probably these features are necessary to allow interactions with other biomolecules (not only during virus assembly, but also at other stages of viral cycle), since a large mobility allow residues to survey a larger number of conformations, and among those, the adequate active structure will be more easily explored. The joint efforts of X-ray crystallography and NMR have allowed to define at a more specific detail the presence of those dynamic equilibria, and even have allowed to identify, and explain some structural ambiguities between both techniques. The combined used of NMR, X-ray crystallography and, in some cases, cryo-EM have allowed a more complete and detailed view of the interactions during assembly. Finally, new advances in SSNMR have contributed to understand the structure of filamentous phage particles and the mechanisms of membrane insertion of these particles.

5.6 Perspectives

A promising future can be envisioned for NMR spectroscopy in structural virology, as exemplified by the techniques and studies described in this chapter, and suggested also by other developments in NMR, which have not been described here because of space limitations. I believe that in the near future, NMR will contribute and/or help in studies dealing with:

i. The structure and, especially, the conformational dynamics of viral proteins.
ii. The structure of viral particles or details of viral proteins in solution through the combined use of cryo-EM (Chap. 3) or X-ray crystallography (Chap. 4) with either solution or solid-state NMR approaches.
iii. The structure and dynamics of viral particles using SSNMR, in conjunction with X-ray fiber or neutron diffraction, for example, to allow a deeper description of the structure of the membrane-bound conformations of viral proteins. Probably, SSNMR could be extended for the description of the first steps of infection (*e.g.*, the formation of fusion complexes).
iv. The structure and dynamics of viral capsids through further advances in solution NMR.
v. The study of structural aspects of a protein within a live cell. The use of in-cell NMR could be applied to viral proteins and assemblies to understand in further structural detail different stages of the life cycle *in vivo*.
vi. The design of new and more efficient drugs to interfere with the activity of any component of the viral machinery (although the first attempts will be probably targeted towards the viral capsid). This will imply the development of new techniques to characterize binding from "the-point-of-view" of the ligand (the so-called *saturation transfer difference* NMR spectroscopy).

Acknowledgements I thank Luis Pérez and Francisco N. Barrera for careful reading of the manuscript, suggestions and new ideas. I apologize to those colleagues whose work was not cited due to lack of space or inadvertent omission. Work in the author's laboratory was supported partially by the Spanish Ministerio de Ciencia e Innovación (MCINN) (CTQ2011-24393 and CSD2008-00005), the FIPSE foundation (2006/0008), and intramural BIFI 2011 projects.

References and Further Reading

1. Atkins PW, de Paula J (2008) Physical chemistry for the life sciences. Oxford University Press, Oxford
2. Wüthrich K (1986) NMR of proteins and nucleic acids. Wiley, New York
3. Derome AE (1987) Modern NMR techniques for chemistry research. Pergamon Press, Oxford
4. Claridge TDW (1999) High-resolution NMR techniques in organic chemistry. Pergamon Press, Oxford
5. Cavanagh J, Fairbrother WJ, Palmer AG III, Skelton N (1996) Protein NMR spectroscopy. Principles and practice, 1st edn. Academic Press, New York
6. Keeler J (2006) Understanding NMR spectroscopy, 2nd edn. Wiley, New York

7. Evans JE (1996) Biomolecular NMR spectroscopy. Oxford University Press, Oxford
8. Bax A, Lerner L (1986) Two-dimensional NMR spectroscopy. Science 232:960–967
9. Reid DG (ed) (1997) Protein NMR techniques. Methods in molecular biology, vol 60. Humana Press, Totowa
10. Dingley AJ, Lorenzen I, Grötzinger J (2008) NMR analysis of viral protein structures. In: Foster GD, Johansen IE, Hong Y, Nagy PD (eds) Plant virology protocols: from viral sequence to protein function. Humana Press, Totowa
11. Adamson CF, Freed EO (2007) Human immunodeficiency virus type 1 assembly, release and maturation. Adv Pharmacol 55:347–387
12. Mascarenhas AP, Musier-Forsyth K (2009) The capsid protein of human immunodeficiency virus: interactions of HIV-1 capsid with host protein factors. FEBS J 276:6118–6127
13. Ganser-Pornillos BK, Yeager M, Sundquist WI (2008) The structural biology of HIV assembly. Curr Opin Struct Biol 18:203–217
14. Ganser-Pornillos BK, Yeager M, Pornillos O (2012) Assembly and architecture of the HIV. Adv Exp Med Biol 726:441–465
15. Adamson CS, Salzwedel K, Freed EO (2009) Virus maturation as a new HIV-1 therapeutic target. Expert Opin Ther Targets 13:895–908
16. Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, Sundquist WY, Hill CP, Yeager M (2009) X ray structures of the hexameric building block of the HIV capsid. Cell 137:182–192
17. Turner BG, Summers MF (1999) Structural biology of HIV-1. J Mol Biol 285:1–32
18. Tang C, Ndassa Y, Summers MF (2002) Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. Nat Struct Biol 9:537–543
19. Byeon IJ, Meng X, Jung J, Jung J, Zhao G, Yang R, Ahn J, Shi J, Concel J, Aiken C, Zhang P, Gronenborn AM (2009) Structural convergence between cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. Cell 139:178–190
20. Shin R, Zhou YM, Krishna NR (2011) Structure of a monomeric mutant of the HIV-1 capsid. Biochemistry 50:9457–9467
21. Saad J, Miller J, Tai J, Kim A, Granam RH, Summers MF (2006) Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. Proc Natl Acad Sci USA 103:11364–11369
22. DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos
23. Byeon IJ, Hou G, Han Y, Suiter CL, Ahn J, Jung J, Byeon CH, Gronenborn AM, Polenova T (2012) Motions on the millisecond time scale and multiple conformations on HIV-1 capsid protein: implications for structural polymorphism of CA assemblies. J Am Chem Soc 134:6455–6466
24. Ivanov D, Tsodikov OV, Kasanov J, Elnberger T, Wagner G, Collins T (2007) Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. Proc Natl Acad Sci USA 104:4353–4358
25. Khorasanizadeh S, Campos-Olivas R, Summers MF (1999) Solution structure of the capsid protein from the human T-cell leukemia virus type-I. J Mol Biol 291:491–505
26. Campos-Olivas R, Newman L, Summers MF (2000) Solution structure and dynamics of the Rous sarcoma virus capsid protein and comparison with capsid proteins of other retroviruses. J Mol Biol 296:633–649
27. Jiang J, Ablan SD, Derebail S, Hercick K, Soheilian F, Thomas JA, Tang S, Hewlett I, Nagashima K, Gorelick RJ et al (2011) The interdomain linker region of HIV-1 capsid protein is a critical determinant of proper core assembly and stability. Virology 421:253–265
28. Klein DJ, Johnson PE, Zollars ES, De Guzman RN, Summers MF (2000) The NMR structure of the nucleocapsid protein from the mouse mammary tumor virus reveals unusual folding of the C-terminal zinc knuckle. Biochemistry 39:1604–1612
29. Lawrence DC, Stover CC, Noznitsky J, Wu Z, Summers MF (2003) Structure of the intact stem and bulge of HIV-1 Psi-RNA stem-loop SL1. J Mol Biol 326:529–542

30. D’Souza V, Dey A, Habib D, Summers MF (2004) NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus. J Mol Biol 337:427–442
31. Amarasinghe GK, De Guzmán RN, Turner RB, Chancellor KJ, Wu ZR, Summers MF (2000) NMR structure of HIV-1 nucleocapsid protein bound to stem-loop SL2 of the Psi-RNA packaging signal. Implications for genome recognition. J Mol Biol 301:491–511
32. Yu XF, Matsuda Z, Yu QC, Lee TH, Essex M (1995) Role of the C terminus Gag protein in human immunodeficiency virus type I virion assembly and maturation. J Gen Virol 76:3171–3179
33. Votteler J, Neumman L, Hahn S, Hahn F, Rauch P, Schmidt K, Studtrucker N, Solbak SM, Fossen T, Henklein P et al (2011) Highly conserved serine residue 40 in HIV-1 p6 regulates capsid processing and virus capsid assembly. Retrovirology 8:11
34. Fossen T, Wray V, Bruns K, Rachmat J, Henklein P, Tessmer U, Maczurek A, Klinger P, Schubert U (2005) Solution structure of the human immunodeficiency virus type 1 p6 protein. J Biol Chem 280:42515–42527
35. Kyere SK, Joseph PR, Summers MF (2008) The p12 domain is unstructured in a murine leukemia virus p12-CA(N) Gag construct. PLoS One 3:e1902
36. Maxwell KL, Yee AA, Booth V, Arrowsmith CH, Gold M, Davidson AR (2001) The solution structure of bacteriophage lambda protein W, a small morphogenetic protein possessing a novel fold. J Mol Biol 308:9–14
37. Maxwell KL, Yee AA, Arrowsmith CH, Gold M, Davidson AR (2002) The solution structure of the bacteriophage head-tail joining protein. J Mol Biol 318:1395–1404
38. Estrada DF, Conner M, Jeor SC, Guzmán RN (2011) The structure of the hantavirus zinc finger domain is conserved and represents the only natively folded region of the Gn cytoplasmic tail. Front Microbiol 2:251–262
39. D’Souza V, Summers MF (2005) How retroviruses select their genomes. Nat Rev Microbiol 3:643–655
40. D’Souza V, Summers MF (2004) Structural basis for packaging the dimeric genome of Moloney murine leukemia virus. Nature 431:586–590
41. Amarasinghe GK, De Guzmán RN, Turner RB, Summers MF (2000) NMR structure of stem-loop SL2 of the HIV-1 Psi-RNA packaging signal reveals a novel A-U-A base-triple platform. J Mol Biol 299:145–156
42. Lu K, Heng X, Summers MF (2011) Structural determinants and mechanism of HIV-1 genome packaging. J Mol Biol 410:609–633
43. Lu K, Heng X, Garyu L, Monti S, Garcia EL, Kharytonchyk S, Dorjsuren B, Kulandaivel G, Jones S, Hiremath A, Divakaruni SS et al (2011) NMR detection of the structures in the HIV-1 5′ leader RNA that regulate genome packaging. Science 334:242–245
44. Taylor GM, Ma L, Vogt VM, Post CB (2010) NMR relaxation studies of an RNA-binding segment of the rous sarcoma virus gag polyprotein in free and bound states: a model for autoinhibition of assembly. Biochemistry 49:4006–4017
45. De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ-RNA recognition element. Science 279:384–388
46. Cai M, Huang Y, Craigie R, Clore GM (2010) Structural basis for the association of HIV-1 matrix protein with DNA. PLoS One 5:e15675
47. Chen K, Bachtiar I, Psizcek G, Bouamr F, Carter C, Tjandra N (2008) Solution NMR characterizations of oligomerization and dynamics of equine infectious anemia virus matrix protein and its interaction with PIP2. Biochemistry 47:1928–1937
48. Saad JS, Ablan SD, Ghanam RH, Kim A, Andrews K, Nagashima K, Soheilian F, Freed EO, Summers MF (2008) Structure of the myristylated human immunodeficiency virus type 2 matrix protein and the role of the phosphatidylinositol-(4,5)-biphospahte in membrane targeting. J Mol Biol 382:434–437

49. Kelly BN, Kyere S, Kinde I, Tang C, Howard BR, Robinson H, Sundquist WI, Summers MF, Hill CP (2007) Structure of the antiviral assembly inhibitor CAP-1 complex with the HIV-1 CA protein. J Mol Biol 373:355–356
50. Neira JL (2009) The capsid protein of human immunodeficiency virus: designing inhibitors of capsid assembly. FEBS J 276:6110–6117
51. Bocanegra R, Nevot M, Doménech R, López I, Abián O, Rodríguez-Huete A, Cavasotto CN, Velázquez-Campoy A, Gómez J, Martínez MÁ et al (2011) Rationally designed interfacial peptides are efficient *in vitro* inhibitors of HIV-1 capsid assembly with antiviral activity. PLoS One 6:e23877
52. Summers MF, Irovalieba RN, Tolbert B, Smalls-Mantey A, Iyalla K, Loeliger K, D'Souza V, Khant H, Schmid MF, Garcia EL et al (2010) Structure of a conserved retroviral RNA packaging element by NMR spectroscopy and cryo-electron tomography. J Mol Biol 404:751–772
53. Vriend G, Hemminga MA, Verduin BJM, De Wit JL, Schaafsma TJ (1981) Segmental mobility involved in protein-RNA interaction in cowpea chlorotic mottle virus. FEBS Lett 134:167–171
54. Szymczyna BR, Gan L, Johnson JE, Williamson JR (2007) Solution NMR studies of the maturation intermediates of a 13 MDa viral capsid. J Am Chem Soc 129:7867–7876
55. Opella SJ, Zeri AC, Park SH (2008) Structure, dynamics, and assembly of filamentous bacteriophages by nuclear magnetic resonance spectroscopy. Annu Rev Phys Chem 59:635–637
56. Sun S, Han Y, Paramasivam M, Yan S, Siglin AE, Williams JC, Byeon IJ, Ahn J, Gronenborn AM, Polenova T (2012) Solid-state NMR spectroscopy of protein complexes. Methods Mol Biol 831:303–331
57. Duer MJ (2004) Introduction to solid state NMR spectroscopy. Blackwell, London
58. Zeri AC, Mesleh MF, Nevzorov AA, Opella SJ (2003) Structure of the coat protein in fd filamentous bacteriophage particles determined by solid-state NMR spectroscopy. Proc Natl Acad Sci USA 100:6458–6463
59. Marassi FM, Opella SJ (2003) Simultaneous assignment and structure determination of a membrane protein from NMR orientational restrains. Protein Sci 12:403–411
60. Park SH, Marassi FM, Black D, Opella SJ (2010) Structure and dynamics of the membrane-bound form of Pf1 coat protein: implications of structural rearrangements for virus assembly. Biophys J 99:1465–1474
61. Goldbourt A, Gross B, Day LA, McDermott AE (2007) Filamentous phage studied by magic-angle spinning NMR: resonance assignment and secondary structure of the coat protein in Pf1. J Am Chem Soc 129:2388–2344
62. Goldbourt A, Day LA, McDermott AE (2010) Intersubunit hydrophobic interactions in Pf1 filamentous phage. J Biol Chem 285:37051–37059

Further Reading

For the basis of NMR:

Brey WS (ed) (1996) Magnetic resonance in perspective: highlights of a quarter of century. Academic Press, New York

Homans SW (1992) A dictionary of concepts in NMR. Oxford University Press, Oxford

Neuhaus D, Williamson MP (2000) The nuclear overhauser effect in structural and conformational analysis, 2nd edn. Wiley-VCH, New York

For intrinsically disordered proteins in viruses:

Dyson HJ (2011) Expanding the proteome: disordered and alternatively folded proteins. Q Rev Biophys 44:467–518

Xue B, Williams RW, Oldfield CJ, Goh GK, Dunker AK, Uversky VN (2010) Viral disorder or disordered viruses: do viral proteins posses unique features? Protein Peptide Lett 17:932–951

Also especially recommended for further reading are references [2–9, 57] *listed above*.

Chapter 6
Fluorescence, Circular Dichroism and Mass Spectrometry as Tools to Study Virus Structure

José L. Neira

Abstract Fluorescence and circular dichroism, as analytical spectroscopic techniques, and mass spectrometry as an analytical tool to determine the molecular mass, provide important biophysical approaches in structural virology. Although they do not provide atomic, or near-atomic, details as electron microscopy, X-ray crystallography or nuclear magnetic resonance spectroscopy can do, they do provide important insights into virus particle composition, structure, conformational stability and dynamics, assembly and maturation, and interactions with other viral and cellular biomolecules. They can be used also to investigate the molecular determinants of virus particle structure and properties, and the changes induced in them by external factors. In this chapter, I describe the physical bases of these three techniques, and some examples on how they have helped us to understand virus particle structure and physicochemical properties.

Keywords Assembly • Calorimetry • Capsid • Chemical denaturants • Chirality • Circular dichroism • Conformational changes • Dissociation • Dynamics • Flexibility • Fluorescence • Interactions • Intermediate • Ligand • Mass spectrometry • Protein stability • Refractive index • Structure • Virus

Abbreviations

ANS	1-anilino-8-naphtalene sulfonate
bisANS	bis(4-anilinonaphtalene-5-sulfonic acid)
CD	Circular dichroism

J.L. Neira (✉)
Instituto de Biología Molecular y Celular, Universidad Miguel Hernández,
03202 Elche, Alicante, Spain

Instituto de Biocomputación y Física de Sistemas Complejos, 50018 Zaragoza, Spain
e-mail: jlneira@umh.es

M.G. Mateu (ed.), *Structure and Physics of Viruses: An Integrated Textbook*,
Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_6,
© Springer Science+Business Media Dordrecht 2013

DSC	Differential scanning calorimetry
EM	Electron microscopy
ESI	Electrospray ionization
FHV	Flock house virus
GdmCl	Guanidine hydrochloride
HBV	Hepatitis B virus
HIV	Human immunodeficiency virus
HK97	Hong-Kong 97 virus
HPV	Human papilloma virus
HRV	Human rhinovirus
HSV	Herpes simplex virus
HX	Hydrogen-exchange
m/z	Mass-to-charge
MALDI	Matrix-assisted laser desorption ionization
MS	Mass spectrometry
MVM	Minute virus of mice
NMR	Nuclear magnetic resonance
PKC	Protein kinase C
RYMV	Rice yellow mottle virus
SL	Stem loop
T_m	Thermal denaturation midpoint
TMV	Tobacco mosaic virus
TOF	Time of flight
UV	Ultraviolet

6.1 Introduction

In Chap. 5 I have described nuclear magnetic resonance (NMR) spectroscopy, a powerful structural and analytical technique to study viruses and viral proteins in solution. In this chapter, I shall describe some other analytical techniques in solution, including fluorescence and circular dichroism (CD) spectroscopies and mass spectrometry (MS). These techniques do not provide information at or close to atomic resolution, as electron microscopy (EM) (Chap. 3), X-ray crystallography (Chap. 4) and NMR (Chap. 5) can do; but they allow us to investigate several aspects of virus structure and properties such as stability and conformational dynamics.

In the first part of this chapter, I shall describe fluorescence spectroscopy and CD techniques. Both are faster (experiments can be completed in minutes) and require lower amounts of biomolecules than EM, X-ray crystallography or NMR. Thus, they can be used as a first approach to obtain some information on the structure and properties of the biomolecule, which can, then, be studied in structural detail by using the latter techniques. Both fluorescence and CD are based in the absorption of light, which induces transits between the electrons of the external shells of the molecules involved. Their uses to study virus structure, stability, assembly, and interactions are described here.

In the second part of this chapter, I shall describe the foundations and some uses of MS in structural virology. This technique is based on changes in the trajectory experienced by ions derived from the target molecule in the presence of electric and magnetic fields. Contrary to fluorescence, CD or NMR, MS does not use the electromagnetic radiation to induce leaps among quantum levels, and then it is not a spectroscopic technique. From the data obtained, MS provides information about the mass of the target biomolecules, even of intact virus particles. Data acquisition is very fast, and the amounts of sample required are even lower than for CD or fluorescence spectroscopies. The use of MS in virus structure, stability, dynamics and assembly, and virus-ligand interactions are also described here.

6.2 Physical Principles of Fluorescence and Circular Dichroism (CD) Spectroscopies

6.2.1 Fluorescence

A detailed account of the general principles of fluorescence can be found in the book by Lakowicz [1], and its applications to biochemistry in the textbook by Albany [2]. I shall briefly revise some basic principles to explain the basis of fluorescence.

The energy of a couple of bound atoms (E(total)) depends on: E(total) = E(translation) + E(rotational) + E(vibrational) + E(electronic) + E(electronic spin orientation) + E (nuclear spin orientation).

The translational energy involves the changes in the position of the molecule; the rotational energy involves the molecular rotation around the gravity center of the molecule; the vibrational energy involves the deformation (compression or stretching) of bonds and angles; the electronic energy involves the movement of electrons within the molecular shell. The last two terms involve the change of the spin of the electron and the change of the nuclear spin (see Chap. 5), respectively. The determination of changes in each of those types of energy leads to different kinds of spectroscopies; the fluorescence and CD techniques are based in electronic transitions.

Each of those types of energies does not distribute continuously between the two atoms, but are "quantized". For instance, the energetic separation between the rungs in the electronic energy ladder is 10^{-17} J; those in the vibrational one are 10^{-20} J; those in the rotational one are 10^{-24} J, and those in the translational one are 10^{-36} J. This energetic distribution leads to two important conclusions. First, inducing an

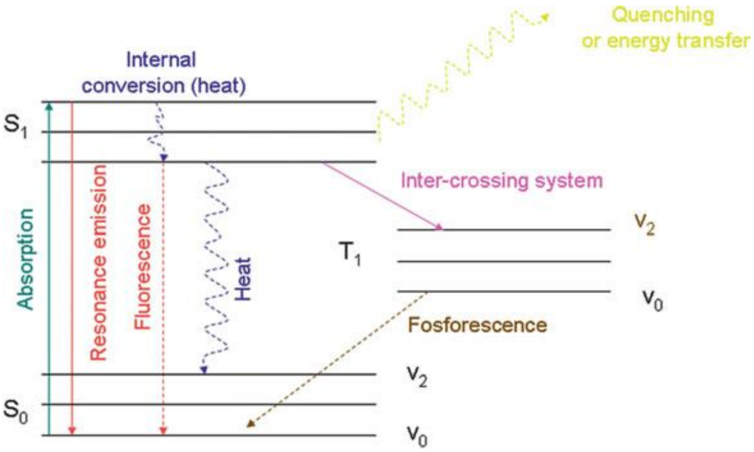


Fig. 6.1 The electronic energy levels and transitions in a molecule. Two electronic levels are shown (S_0 and S_1) with the corresponding vibrational levels (the number of vibrational levels drawn is arbitrary, and they are called $V_0 \ldots V_i$). The absorption of energy is shown by a continuous *cyan line*, and the emission by a *red* one; the internal conversion processes (such as the emission of heat) are shown by *curved dotted lines*. The emission fluorescence is shown by a *dotted red line*. The inter-crossing system, which involves the jumping from an electronic state to another unstable one (T1) with different spin properties, is shown in *pink*, and the fosforescence emission in a *dotted brown line*. The quenching or energy transfer is shown by a *curved dotted yellow line*. The meaning and nomenclature of both electronic terms, S and T, are beyond the scope of this book; the interested reader can have a wider look on the meaning of the electronic states in [1]

electronic transition involves also the induction of rotational, vibrational and translational transitions, because they have smaller energy. And second, for an electronic level there are vibrational ones; for each vibrational level there are rotational ones, and so on.

The Basis of Fluorescence

We deal in fluorescence with photon absorption and electrons located in specific orbitals in a molecule. The absorption of photons induces leaps from a ground electronic level to an excited state. In this excited state, the molecule can return to the ground state following any of these processes (Fig. 6.1):

(a) The molecules dissipate the excess of energy in the surrounding environment, usually, but not exclusively, as heat.
(b) The molecules can release the same amount of absorbed energy by the emission of a photon. This is called resonance.
(c) Part of the absorbed energy is dissipated in the medium as heat. Then, the molecules are in the lower vibrational level of the electronic excited state, and from this rung in the energetic ladder, they release a photon to return to the ground electronic level. The released photon will have a lower energy than the absorbed one, and then, its wavelength will be larger. This process is fluorescence.

(d) A transient passage occurs to an unstable electronic state, from where the molecule can pass to the ground state, by photon emission. This process is phosphorescence.
(e) Some molecules in the excited state can release energy to nearby molecules (by collisional quenching or energy transfer).

It is clear that the fluorescence emission spectrum is the mirror image, shifted towards longer wavelengths, of the absorption spectrum, provided that the vibrational distribution in both ground and excited states is basically the same. However, we cannot conclude that all the absorbing molecules will have fluorescence, because energy can be dissipated within the environment as collisional or thermal energy, or as energy transfer to other molecules. For instance, the heme group does absorb, but it does not show fluorescence, due to transfer of energy from the porphyrin ring to the iron.

Fluorescence of Biomolecules

In proteins, tryptophan, tyrosine and phenylalanine residues are responsible for the UV properties of the molecule, but whereas Trp absorbs more intensely than the two others, Tyr has the largest fluorescence. When both residues are present in a protein, Tyr fluorescence is weaker, due mainly to energy-transfer from tyrosine to tryptophan. The tryptophan fluorescence can be quenched not only by external agents (such as oxygen), but also by nearby residues. Furthermore, the environment of a tryptophan affects the position of its fluorescence spectrum maximum (which does not occur for tyrosine residues). Thus, variations in fluorescence intensity and/or wavelength of maximum intensity can be used to monitor stability, conformational rearrangements, dissociation, unfolding, and ligand binding and affinity of proteins and protein complexes, including viral particles.

Isolated sugars show no fluorescence, and the fluorescence work with lipids is rather limited due to the possibility of dispersion due to vesicle size. Puric and pyrimidinic bases in nucleic acids show a weak fluorescence (although they absorb strongly at 260 nm) at physiological pH, but they fluoresce strongly at acidic pHs and low temperature. Puric bases (especially guanine) show fluorescence at least three times more intense than pyrimidinic bases. In the native state, DNA or RNA show very weak fluorescence (except transfer RNAs which contains a highly fluorescent modified base); however, extrinsic fluorescent agents, that is binding to nucleic acids (such as ethidium bromide), can be used to detect nucleic acids. In this way, for example, the structural stability of some virions can be followed by quantifying the amount of nucleic acid released on capsid dissociation (and which is accessible to the external fluorophore).

Extrinsic fluorophores, such as ANS that binds to proteins, can change the protein fluorescence emission spectrum depending on the conformation of the macromolecule, and can be used to monitor unfolding, detect intermediate states or conformational rearrangements in proteins and protein complexes, including virus capsids. Furthermore, extrinsic fluorophores can be covalently attached to proteins, and used to map protein-ligand interactions within living systems [3, 4], after excitation at the appropriate wavelength.

6.2.2 Circular Dichroism

The Basis of CD

A property of many molecules is their chirality or molecular asymmetry. A description of the properties and characterization of organic chiral molecules can be found in organic chemistry texts [5]. The simplest cases result from the presence of asymmetric carbons; however, molecular chirality does not rely exclusively on the asymmetrically substituted carbon atoms, but rather in any molecular feature providing "handedness". For instance, proteins in the α-helical conformation wind in a right-handed manner; however, it is important to note that this helical "handedness" comes, at the last instance, from the stereochemistry of the amino acids. In nucleic acids, the helical polynucleotide winds in a left- or right-handed manner [6].

The chiral structures can be characterized by polarized light. In non-polarized light (Fig. 6.2a), the electric vector of the electromagnetic radiation can oscillate in any direction perpendicular to the way of propagation (for simplicity I shall reduce the following brief description to the electric field since, by using Maxwell's equations, we can characterize the electromagnetic radiation just by knowing the electrical component). Conversely, in a plane-polarized beam, the electric field oscillates only in a plane (Fig. 6.2b); if we moved with the field, we would see the vector representing the electric field oscillating back and forth along a line (Fig. 6.2b, right panel). In addition, in a circularly polarized beam, the single plane of oscillation rotates (Fig. 6.2c, left side): if we moved with the field, we would see the electric vector moving clockwise (right-handed circularly polarized light) or counter-clockwise (left-handed circularly polarized light). This single rotating plane can be decomposed in two components: the left and right circularly polarized vectors.

The two enantiomers of a chiral molecule have different refractive index, *n*; this phenomenon is called *birefringence*. That is, the variation of *n* is not the same through any direction in space in that material, and so, it is said that the variation of *n* is anisotropic. Biological cells or crystals have different *n*, depending of: (i) the direction of the beam, which illuminates them; and, (ii) the different planes of material they are made of. Chiral molecules show *circular* birefringence, and we say that the molecule shows circular dichroism. A good and rigorous description of the physical basis of CD can be found in several textbooks [7–9]. From a quantum perspective, CD is related to a transition between both electronic states in a molecule; this transition is active if, for an instant, the electric and magnetic components of both states are not perpendicular to each other (and this only occurs if the molecule is chiral). The CD spectrum results from the different absorption (or changes in the phase of the wave) of right-handed or left-handed circularly polarized light by a chiral molecule. In addition to de-phasing (relative to the incident beam), the amplitude of the emerging electric field is also modified differently for the left-handed and right-handed circularly polarized light~~s~~. Thus, the resulting electric field amplitude (the sum of the left-handed and right-handed circularly polarized electric field amplitudes) does not follow a circle, but it does an

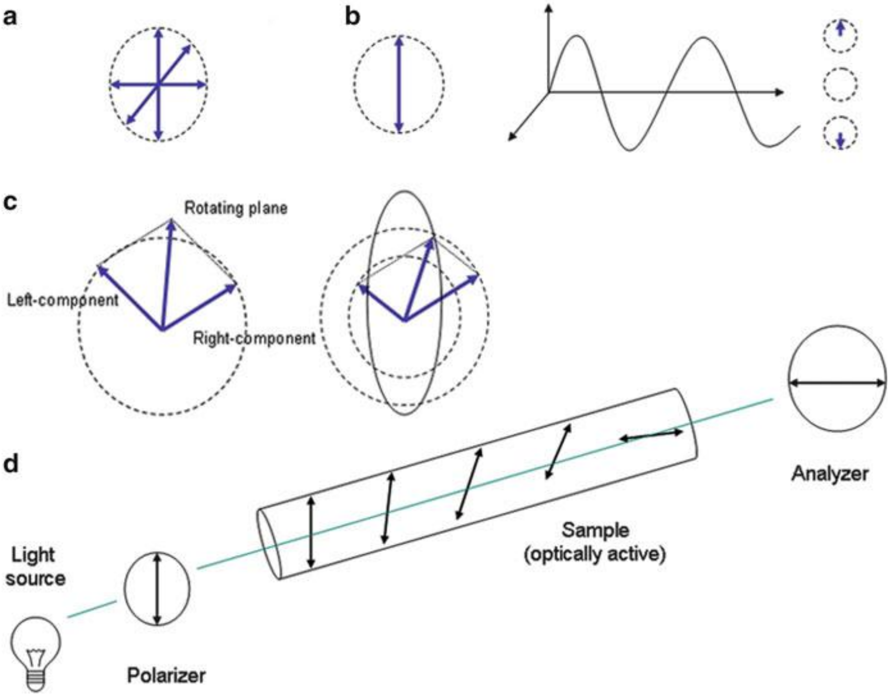


Fig. 6.2 Polarized light and CD. For simplicity, only the electric field (and not the magnetic field) of the electromagnetic radiation is shown (*double arrow head*): (**a**) Un-polarized light, as seen by a stationary observer. (**b**) Linearly (or plane) polarized light (*left panel*) as seen by a stationary observer; (*right panel*) the *small circles* show how the beam would be seen by an observer travelling with the beam in a linearly polarized light. (**c**) A circularly polarized light, as seen by a stationary observer, with the left and right circularly polarized components (*left panel*), which when added generate plane (linear) polarized radiation; (*right panel*) when the circularly polarized beam passes through a sample optically active, the left and right components have different amplitude and the resultant addition component is elliptically polarized (*continuous line*). (**d**) Set-up of a CD spectropolarimeter. The *green line* indicates the beam

ellipse (Fig. 6.2c, right side). It is said that the emerging beam is elliptically polarized, and then, the units of the CD are provided as a measured of the degree of the ellipticity of such ellipse. The changes in the ellipticity for a molecule depend on the wavelengths of the incident light and result in a CD spectrum (see examples in Fig. 6.3). There are specific rules able to predict which component of the polarized light is preferentially absorbed [9–11]. An experimental set-up for measuring CD in an optically active sample is shown in Fig. 6.2d.

CD of Biomolecules

Traditionally, the regions of the electromagnetic spectrum useful for CD analysis of biomolecules have been divided in: (i) the far-UV region (from 190 to 250 nm); and, (ii) the near-UV region (from 250 to 320 nm). In proteins, the peptide bonds

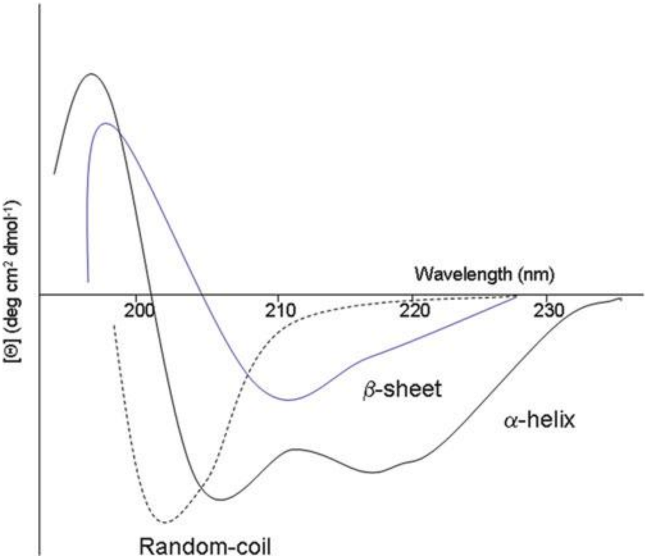


Fig. 6.3 Far-UV CD spectra of different model proteins: pure α-helix (*black continuous line*), pure β-sheet (*blue continuous line*), or random coil (*black dashed line*)

absorb UV light below 240 nm; the aromatic side chains absorb mainly in the 260–320 range (they can also absorb around 222 nm [11]); and the disulphide bridges show a broad absorption band at 260 nm. In addition, the spectra of the different types of secondary structure have characteristic shapes (see below). Some CD-based studies use wavelengths longer than those indicated above but, in most cases, they involve molecules with a strong coloured chromophore.

Secondary structure elements in proteins (α-helices, β-sheets or stretches with irregular structure such as loops or random-coil, denatured polypeptides) show a characteristic far-UV CD spectra (Fig. 6.3); perhaps the most characteristic is that of the α-helix, with two minima at 208 and 222 nm. Obviously, those proteins containing a mixture of the three types of structure will show a CD spectrum that is the resultant of the three spectra mentioned above, according to the percentages of the corresponding types of secondary structure elements. Deconvolution of the spectrum of a protein allows an estimation of the percentage of each type of secondary structure element, and there are programs implemented online to do so [12, 13]. Caution, of course, must be exercised (as with any mathematical procedure applied to complex experimental data), since the presence of different overlapping bands may lead to wrong conclusions on the secondary structure prediction of the protein.

The near-UV spectrum of proteins is usually rather weak, except for those with a large number of aromatic residues (thus, large protein concentrations in a wide path-length cell are generally required) [11]; contrary to what happens in the far-UV region, the near-UV spectrum of proteins is mainly associated with the asymmetric environment of the aromatic residues. This region is only used in specific cases, since the information provided is similar to that obtained by fluorescence spectroscopy. On the other hand, the near-UV spectrum of double-stranded DNA is very intense due to the packing of the bases. There are also “model-spectra” for the different nucleic acid conformations [9, 11]. As it happens with proteins, the transitions among different conformers can be monitored by following changes in the spectrum. The CD analysis of lipids is rather limited, due to the possibility of dispersion because of the size of the vesicles formed.

6.2.3 Fluorescence and CD Spectroscopies to Study the Stability and Conformational Dynamics of Biomolecules

In addition to investigating some structural aspects of biomolecules and biomolecular complexes including viruses, fluorescence and CD spectroscopies can be used to detect conformational changes in these molecules and complexes by following changes in their fluorescence and/or CD spectra. One particular example is the determination of their conformational stability. As the far-UV CD spectrum of a random-coil is so different from those of the other types of secondary structure in proteins (Fig. 6.3), it is possible to follow the variations in the protein secondary structure by measuring the changes in the ellipticity at a selected wavelength (usually at 222 nm). Likewise, as the Trp or Tyr fluorescence in a protein is sensitive to changes in the exposure to solvent, it is possible to follow changes in the protein tertiary (and quaternary) structure by measuring the changes in fluorescence intensity at a selected wavelength (or, in the case of Trp, the changes in the wavelength of the maximum intensity). For the determination of protein stability by using these approaches, melting of the protein structure can be induced by a chemical denaturant (such as pH, urea or GdmCl) or temperature. In the simplest cases, if the melting is quite cooperative (an all-or-none transition) a sigmoidal curve will be obtained that can be fitted to a two-state model, and the thermodynamic parameters of interest can be obtained [14, 15].

6.3 Fluorescence and CD Spectroscopies to Study Virus Structure

CD and fluorescence spectroscopy have been used to study structural features, stability and conformational changes of isolated viral capsid proteins, the stable oligomers they form, and occasionally, even those of entire virus particles. In addition, both techniques have been used to study the interactions of viral proteins with other biomolecules. I shall divide the studies in the literature in each of those three different aspects.

6.3.1 Conformational Changes in Viral Proteins and Capsids

A general use of far-UV CD and fluorescence in structural virology is to investigate the overall secondary and tertiary structures of virus capsid proteins and their possible changes on capsid assembly. For example, the hepatitis B virus (HBV) capsid protein was found to be mainly helical, and showed little conformational restructuring on assembly [16]. There are many examples where mutations in the capsid proteins, which result in impaired self-association, do not lead to a significantly different fluorescence or far-UV spectra of the mutants, compared to those of the wild-type protein. In these cases, CD and/or fluorescence analyses showed that capsid assembly may proceed without large structural changes in the tertiary (or secondary) structures of the capsid building blocks; they also showed that subtle structural changes are enough to impair capsid assembly, which favours the possibility of developing antiviral strategies targeting this process (see Chap. 20). Conversely, dimerization of the capsid protein of human immunodeficiency virus type 1 (HIV-1) involves a transient monomeric intermediate and conformational rearrangements, although the overall fold of the monomer is maintained [17]. Changes in the structures of the capsid proteins occurred, for example, upon the pressure-induced dissociation of human rhinovirus (HRV14) [18], or the multimerization of the N-terminal region of Mason-Pfizer monkey virus p12 protein [19]. These and other examples using CD or fluorescence revealed that dissociation or association of some viral capsids may entail substantial conformational rearrangements, a fact that has been repeatedly confirmed and studied in structural detail using EM and X-ray crystallography (see Chaps. 3, 4 and 13).

Another general use of far-UV CD and fluorescence in structural virology is the study of conformational rearrangements of virus proteins in response to factors in the *in vivo* environment, by using agents which mimic *in vitro* the action of those factors within the host cell or virion. For instance, the HIV-1 polyprotein Gag contains a spacer polypeptide (SP1), whose first five residues are critical to allow the proper assembly of Gag into an immature capsid. The SP1 is unfolded in aqueous solution, but at high concentrations it shows a helix-like far-UV CD spectrum; thus, Gag aggregation possibly promotes the acquisition of secondary structure by SP1 which may mediate its function [20]. The effect of other intracellular or intravirion factors on virus protein conformation have been also mimicked *in vitro* by using salts (ions) at high concentrations, or organic solvents, changes in pH, or even by the addition of lipids; the conformational changes have been monitored by either CD and fluorescence. The results have contributed to reveal a substantial conformational dynamics of virus structural proteins, although no general trends are apparent, and the conclusions depend largely on the particular protein explored.

Likewise, far-UV CD and, more frequently, fluorescence spectroscopy have been used also to detect conformational rearrangements in assembled virus capsids, in response to factors such as heat or changes in pH that may mimic physiological conditions. As an example, fluorescence spectroscopy was used to detect a subtle conformational rearrangement in the capsid of the minute virus of mice (MVM) that was associated with a biologically relevant peptide translocation through capsid pores [21].

6.3.2 Stability and Association of Viral Capsid Proteins

Both CD and fluorescence spectroscopy have been used to determine the stability of viral proteins and particles in thermal or chemical denaturation assays (see Sect. 6.2.3). An excellent review on assembly as a binding process can be found in the work by Katen and Zlotnick [22] (see also Chap. 19; ref. [22] also contains an up-to-date review on the stability of several viruses followed by different spectroscopic techniques). For example, it was found that the wild-type HBV capsid protein dimer is highly stable, showing a T_m above 95 °C [16]. The ability of this dimer to self-assemble is improved by the naturally occurring mutation F97L, but this mutant has the same structure (as monitored by both techniques), and the same stability, as the wild-type [23]. The authors suggested that this mutation could affect the conformational dynamics of the dimer.

Among the comparatively few examples where the stability of whole virus particles was studied, urea-denaturation analyses of HRV14 followed by far-UV CD and fluorescence, showed a small stability [18]. However, the dissociated capsid proteins of HRV14, as analyzed by far-UV CD, had residual structure even at 8 M urea [18]. Interestingly, it was also possible to observe the signal in the near-UV of the viral RNA, and there was a decrease in the nucleic acid signal at 260 nm as the urea concentration was increased. This result suggests that the RNA becomes more ordered in the assembled capsid (see also Chap. 5).

Far-UV CD and fluorescence have been used also to identify the molecular determinants of folding, stability and oligomerization of capsid proteins by mutational analysis. For example, the thermodynamic roles of every residue at the dimerization interface in the HIV-1 capsid protein were probed by far-UV CD (secondary structure), fluorescence (tertiary structure) and gel filtration (quaternary structure) [24]. Mutational studies to analyze the folding, stability and/or oligomerization of viral capsid proteins have been reported also for bacteriophage P22 [25], *Physalis* mottle tymovirus [26] and MVM [27]. In all cases, thermal- and chemical-denaturations of the mutated residues, involved in inter-subunit interactions, disrupt both subunit protein folding and particle assembly. From these studies, it can be concluded that the distribution of critical residues and binding energy in intersubunit interfaces in virus capsids may be complex and quite different among viruses, and involve, in some cases, a few key residues, interactions and allosteric effects.

Studies on the folding, stability and association of viral proteins monitored by fluorescence and CD can be complemented by thermal denaturation analysis followed by differential scanning calorimetry (DSC); in this biophysical technique, instead of following a spectroscopic property, the heat exchanged in the dissociation and/or unfolding reaction of a viral capsid or protein is monitored. For instance, the thermal stability of the empty capsid of MVM have been studied by using fluorescence and DSC [21]. The DSC results indicated that the dissociation/denaturation transition involves a high enthalpy change and proceeds through one or more intermediates, which disappear in the presence of GdmCl. Furthermore, thermal dissociation of the capsid involves a partially folded species of the capsid protein [27], with a high flexibility and practically devoid of tertiary structure. A similar partially folded conformation has been observed, by using fluorescence and CD, for the isolated capsid protein of herpes simplex virus (HSV) [28], and during maturation of the bacteriophage HK97 capsid [29]. These results, together with the existence of natively unfolded proteins (see Chap. 5), suggest that partially folded conformations (or completely disordered structures) of viral proteins may be very important in the virus life cycle [30]. Probably, flexibility may permit virus proteins to explore a wider conformational space, which facilitate the selection of the required conformation for self-association or interaction with other biomolecules (see Chap. 5).

The thermal stability of some viral capsid proteins have been analyzed at different pHs to elucidate the importance of several hydrogen-bonds, which disappear (or are formed) as the concentration of protons is modified. For instance, it has been shown that the thermal stability of the capsid protein of adenovirus type 2 is increased in mildly acidic conditions [31]. The effect of covalent modifications in viral capsid proteins on their thermal stability has also been analyzed. For example, PKC-mediated phosphorylation of the HBV core protein was shown to stabilize the protein, and also improved its self-association [32].

6.3.3 Binding of Ligands to Viral Proteins

If binding of a ligand induces a change in the secondary and/or the tertiary structure of the protein, the binding reaction can be monitored by far-UV CD and/or fluorescence, respectively. In a general sense, the addition of ions, or changes in pH can be considered particular examples of binding. Near-UV CD, together with fluorescence, can be used to detect the changes occurring in the protein (or the nucleic acid) upon binding of DNA or RNA. Alternatively, an increase in the T_m of the thermal denaturation of a protein monitored by far-UV CD, upon addition of a ligand [18] may indicate binding [33]: usually, upon binding, there is an increase in the T_m of the formed complex, when compared to that of the isolated protein. However, if the dissociation constant of the complex is large, or the heat released upon binding is very small, there might not be any changes in the measured T_m.

Both CD and fluorescence can be used to monitor the binding of ions to capsid proteins, which in some cases can trigger capsid assembly. A particularly well-studied example is that of HBV, where the affinity for Zn(II) has been determined by following changes in the fluorescence spectrum [22]. These findings suggested that: (i) the assembly reaction is kinetically trapped; and, (ii) the overall process of assembly triggered by the ion is allosterically regulated. Binding also can be also explored by using minimalist models of the capsid protein, as it has been described with fragments of human papilloma virus capsid [34]. Changes in fluorescence can be monitored not only on the capsid protein, but also on other molecules that can be used as reporters, as it has been shown, for example, to demonstrate the inhibition of capsid assembly of *Salmonella typhimurium* phage P22 by 1,1-bis(4-anilinonaphtalene-5-sulfonic acid) [35]: the bisANS-bound coat or scaffolding proteins are incapable of assembly into procapsids.

Binding of antivirals or other molecules that affect virus capsid assembly (see Chap. 20) have been also monitored by fluorescence and/or CD [36]. In the cases reported, binding induces conformational changes in the protein, which can hamper proper assembly [37]; an excellent application of the design and use of these molecules can be found in the work by Zlotnick's group on HBV [38, 39].

6.3.4 Understanding Viruses: Some Major Contributions of Fluorescence and CD

The experiments with different viral systems, such as those described above, have proved the importance of spectroscopic analyses based on CD and fluorescence to analyze the folding, stability, conformational dynamics, and association of viral capsid proteins, and also to study the stability and conformational dynamics of viral particles, the effects of different factors, and the molecular determinants of these processes. The results using these and other biophysical techniques have revealed some rather general basic themes and a wide diversity of specific strategies followed by different viruses to respond to the selective pressures acting on the conformational stability and dynamics of capsids and virions.

6.4 Mass Spectrometry (MS) as an Analytical Tool

6.4.1 The Basis of MS

MS identifies molecules on the basis of their mass and charge and has become one of the most sensitive methods to study some aspects of virus structure. The first mass spectrometer was designed at the turn of the twentieth century, by Sir Joseph J. Thomson. Thomson observed the deflection of a beam of positively-charged ions, when subjected to a combination of magnetic and electrostatic fields. Ion deflection occurred through a series of parabolic curves, corresponding to particular mass/charge (m/z) ratios of the ions. The particular position of the curve was related to the velocity of the ion, and then, to its mass and acceleration. Although MS has evolved since those experiments, the basic set-up of the technique remains: all MS have three key elements: (i) an ion source; (ii) an analyser (that is, a sieve) of ion beams according to the m/z ratio; and, (iii) a detector monitoring the deflected beams (Fig. 6.4, top).

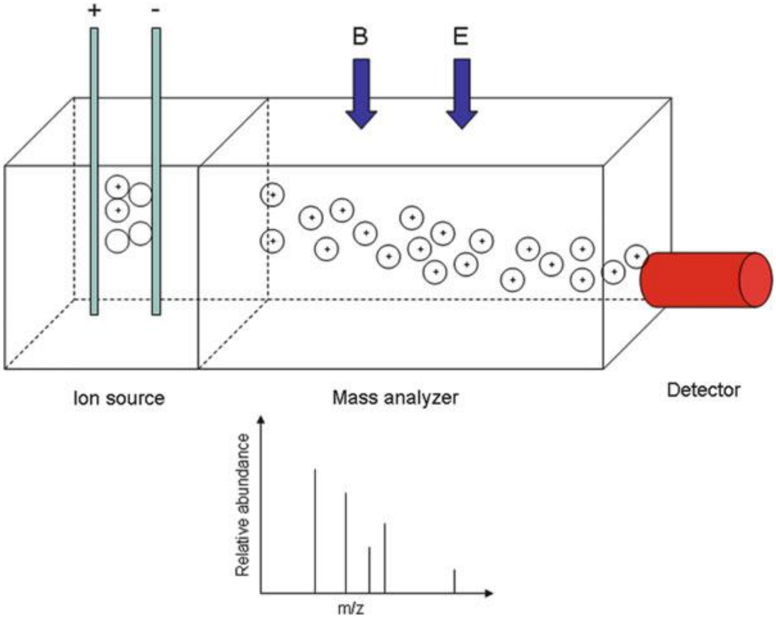


Fig. 6.4 Mass spectrometry. (*Top*) a mass spectrometer set-up, showing the ion source, the ion analyzer and the detector. In the ion analyzer, the deflection of the ion beams is shown. Initially, the sample is formed by neutral and charged molecules. (*Bottom*) example of a mass "spectrum" of a model protein

In the first stage in MS, molecules are vaporized and, in some experimental devices, also broken down (usually, but not exclusively, by the use of highly energetic particles); this molecular breaking produces ions. In the second stage, the ions are accelerated in a chamber where a high vacuum exists, under the influence of a combination of electric and magnetic fields; according to Newton's second law, lighter ions will deflect further than heavier ones, and they will follow different trajectories, depending on the ion masses and their charges. The third stage involves the monitoring of the ion deflections, which are used to calculate the m/z ratio, generating the mass finger-print, or the mass "spectrum". Graphically, a typical mass "spectrum" is a plot of intensity (signal) data *versus* the m/z ratio (Fig. 6.4, bottom) (and it is called "spectrum" in a wide sense, since no light is involved); thus, the nature of the mass "spectrum" will depend on how the spectrometer handles the sample: whereas some MS instruments break the molecules into several fragments, others record intact ones (that is, a single ion is produced).

6.4.2 Current MS Techniques

An excellent overview of the current methods in MS can be found in [40, 41], and its application to analyze protein complexes (focused in ESI-MS, see below) in

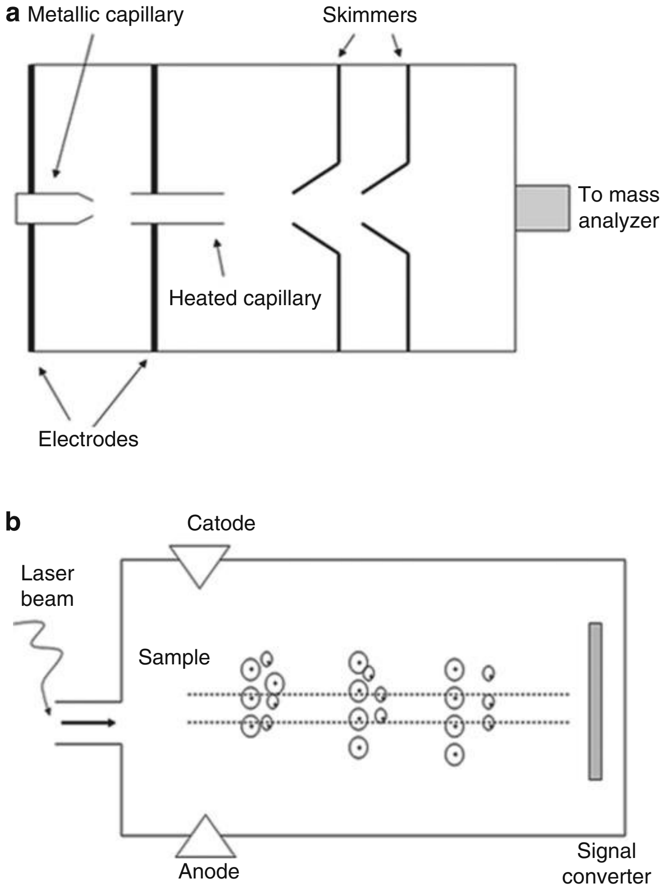


Fig. 6.5 Major types of MS in use. (**a**) ESI-MS schematic set-up representation: The solution is passed through a metallic capillary tube to produce a fine spray of droplets; the droplets cross through a large electric field which ionize them (*dark*, *thick lines* indicate the electrodes). The heated capillary dries the sample. The skimmers are cones with a central orifice designed to intercept the center of the droplet spray so as to sample the central portion of the expansion. (**b**) MALDI-TOF MS schematic set-up representation: The sample has been vaporized by the use of a laser beam, and it is ionized by the use of a electric field; as the ions pass through the analyzing chamber, the different ions are separated according to their sizes

[42]. There are currently two major MS techniques, which rely on different modes of sample ionization (in some examples, both techniques can be combined [42]):

(a) Electrospray ionization (*ESI*)-*MS*: An exhaustive description of the ESI technique can be found in [43, 44]. In this type of MS instrument, ESI is used to vaporize the sample. In ESI-MS, a volatile solvent with the molecule is pushed through a very small charged capillary, creating a spray of charged droplets, which during their evaporation leads to ionization (Fig. 6.5a), creating singly- or multiply-charged ions. The sample is directed towards the analyzer, which, in most cases, is an ion trap. The ion trap uses a combination of magnetic and electric fields to trap ions in vacuum, by variation of the intensity and/or the direction of both fields (the so-called quadrupole trap).

(b) Matrix-assisted laser desorption ionization-time of flight (*MALDI-TOF*)-*MS*: This MS method is based on a soft ionization of the sample. This ionization uses a high-energy laser beam applied to a macromolecule, which is immersed in a protecting matrix (α-cyano-4-hydroxicynnaminic acid). The mixture is added to a plate and is allowed to dry (crystallize) before insertion into the MS. The laser causes the vaporization and ionization of the crystallized mixture (hence the name, matrix-assisted laser desorption ionization) (Fig. 6.5b). The ions produced are accelerated in the electric field in a vacuum long tube. We measure the time they take to reach the detector (their "time of flight"): larger ions reach the detector later than lighter (low m/z ratios) ones. As the acceleration occurs in vacuum, the velocity is wholly dependent on acceleration, which relies on mass, and ions reach the detector within 1 ms. MALDI-TOF-MS does not require sample fragmentation: it can be applied to biomolecules with a large mass, including viral particles; in contrast to ESI-MS, it relies on the production of singly-charged ions. Excellent reviews describing the advantages, applications and development of MALDI-TOF-MS can be found in [45, 46].

6.5 MS to Study Virus Structure

There are several advantages in the use of MS, when compared to other structural or analytical techniques described in this book (see Chaps. 3, 4 and 5, and Sects. 6.2 and 6.3). A first technical advantage is the very small amount of material required (usually two or three orders of magnitude lower than fluorescence or CD, and several orders of magnitude lower than for EM, X-ray crystallography and NMR). The second technical advantage relative to structural techniques is the speed of analysis, which can take only a few minutes. Regarding its uses in structural virology, MS is ideally suited to monitor changes and interactions of macromolecules in real-time: in particular, we can use MS to follow virus particle assembly pathways and conformational changes, obtaining information on the stability, dynamics and stoichiometry.

To study virus particles, MS is usually carried out at near-neutral pH, where ionization does not occur to a large extent, and relatively high ionic strength (close to physiological conditions). The low ionization has two major technical drawbacks: (i) weak signals in MS; and, (ii) high m/z ratios. Furthermore, the presence of salt and buffer ions (either specifically or non-specifically bound to the virus particles) leads to heterogeneous, broad signals, from which the mass cannot be easily and accurately obtained; for structural studies of virions containing nucleic acids the situation is worse, due to cation binding to the nucleic acid. Despite these drawbacks, MS is being increasingly used in varied and interesting applications in the field of virus assembly and structure. Several examples are provided in the three sub-sections below, which are interrelated.

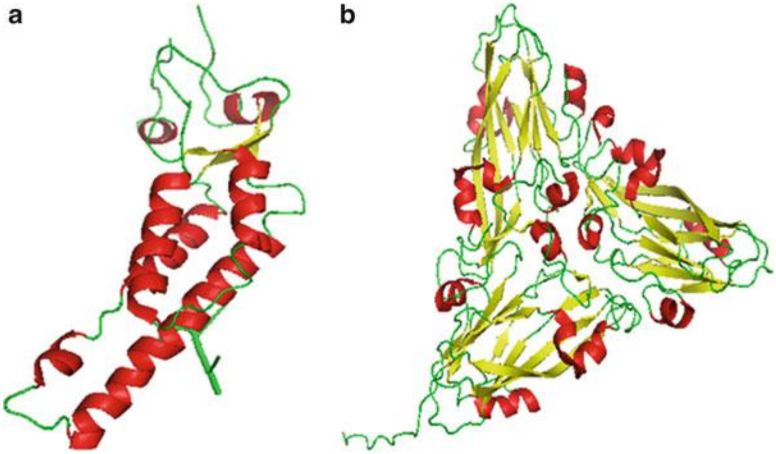


Fig. 6.6 Structures of two virus capsid proteins. (**a**) Structure of the TMV capsid protein (PDB number: 2OM3; [48]). The figure was obtained with PyMol [49]. A nucleotide-binding region is shown in *green* on the right side of the structure. (**b**) Structure of a trimer of the RYMV capsid protein (PDB number: 1F2N; [50]). Helices are shown in *red*, loop regions in *green* and β-strands in *yellow*

6.5.1 Virus Particle Composition and Structure

An early report of the study of an intact virus capsid by MS was described in the mid-nineties [47]. The author ionized rice yellow mottle virus (RYMV) and tobacco mosaic virus (TMV) (Fig. 6.6). Both are single-stranded RNA virus, but whereas RYMV has a 29-nm-diameter icosahedral capsid (with a total molar mass of 6.2 MDa), TMV forms a rod-shaped helical capsid with a diameter of 17.5 nm and a length of 300 nm (with a total mass of 40.5 MDa). The masses determined by MS were within ±15 % from the calculated ones [51]. Usually, in these experiments the samples originate high-charged states, with various m/z rates, which are difficult to assign. The authors showed that the samples used in the MS remained infectious and retained their native morphology; these experiments provided proof-of-concept for the non-invasiveness of MS, and the possibility of studying biological, weak, non-covalent complexes.

Since then, MS has been used to determine the masses of viral particles with different morphologies. For instance, the HBV capsids exhibit two different icosahedral geometries ($T = 3$ and $T = 4$, see Chap. 2), composed of 90 and 120 dimers of the capsid protein, and theoretical masses of ~3 and ~4 MDa, respectively. The experimentally determined masses were obtained by MS, and they were within 0.1 % of the calculated values [52]. These and other studies show that MS can be used to characterize virus particles by precise determination of their molecular weight.

6.5.2 Virus Particle Stability, Conformational Dynamics and Assembly

MS has been used also to explore the factors that determine the correct pathways of assembly, and stable enough assembly intermediates. For instance, the HBV capsid protein has two regions: a core (residues 1–140) and the protamine domain (residues 150–183), connected by a linker, whose exact length (and the conditions where assembly takes place) determines the proportion of $T = 3$ and $T = 4$ capsids that are formed during assembly (see above). MS has allowed to determine the rate of inter-conversion between both capsid forms, and that between those species and the soluble pool of capsid proteins (by using a mixture of ^{15}N-labelled with unlabelled proteins); therefore, the relative stability of both species can be also explored. Both capsids showed a similar temperature-dependent stability, although the $T = 3$ form maintained its integrity better during the MS experiments [52].

The stability and assembly of virus capsids may depend on capsid protein concentration, pH, redox conditions, the presence of disulphide-bridges, ionic-strength, metal ions, scaffolding proteins, the viral nucleic acid and other factors (see Chaps. 1, 10, 11, 12 and 19). A study on Norwalk virus provides an example on the use of MS to investigate the influence of pH on capsid assembly. The fully assembled capsid is formed by 180 copies of capsid protein VP1. MS analysis showed that the capsid is very stable at acidic pHs, but at alkaline pHs intermediate or aberrant structures formed by 2, 4, 6, 18, 40, 60 and 80 copies of VP1 were identified by MS (the latter two being the most abundant) [53]. These species could be reversibly converted into the native capsid (180 VP1 copies) at acidic pHs. The authors proposed a nucleation model of capsid assembly (see Chap. 19) based in the formation of smaller oligomers, mainly dimers, prior to their association into larger 60- and 80-copies intermediate species.

Another study on bacteriophage Qβ provides an example on the use of MS to analyze the influence of disulphide-bridges in capsid assembly [54]. MS analysis of capsid disassembly after disulfide bridge reduction yielded capsid protein dimers. However, disassembly of the intact capsids (without reduction of the disulphides) yielded mainly pentamers; reduction of the disulfide bridges in the pentamers led to monomers. Thus, the assembly pathway of the Qβ capsid may be directed by the intermolecular formation of covalent (disulphide) bonds around the capsid five-fold axes.

MS has been used also to study the influence of the viral nucleic acid on virus particle assembly (see also Chap. 12). Two examples are provided. In the first study, the *in vitro* reassembly of bacteriophage MS2 was triggered by a sequence-specific RNA-protein interaction between the capsid protein dimer and a RNA stem loop (SL) encompassing the Ψ-site (packaging signal) in the viral RNA [55–57]. MS analysis showed that the protein-RNA complexes had a 1:1 stoichiometry. In the second study, the concomitant dimerization and packaging of the HIV-1 genome (also analyzed by NMR, see Chap. 5) was investigated by using MS. The two copies of the genomic RNA are bound by an inter-molecular base-paired region: formed by the self-complementation of a sequence located in the SL of the Ψ-site. The

RNA dimer is stabilized through the interaction with the nucleocapsid protein, and the ESI-MS studies showed that binding-sites are present in two, and not all four, of the SLs [56]. Therefore, the comparison of the assembly pathways in the presence and in the absence of the SL for MS2 has allowed to discern the importance of the different intermediate self-associated capsid protein species to achieve an efficient capsid assembly [57]; in fact, intermediates in which each subunit protein capsid had only a handful of contacts with the nearby subunits would tend to be favoured during capsid assembly, and thus a nucleation mechanism was proposed.

6.5.3 Dissection of Conformational Stability and Dynamics in Virus Particles Using Limited Proteolysis-MS or Hydrogen-Exchange (HX)-MS

Major advances in the study of specific regions of a virus particle in its conformational stability and dynamics have been possible by using limited proteolysis or hydrogen exchange (HX) coupled to MS and a few other techniques. The interested reader can find this topic reviewed in [58, 59].

Limited Proteolysis-MS

Viral capsids are conformationally dynamic assemblies. Even at equilibrium, a capsid may sample different conformations and transiently expose some structural elements that would be otherwise buried (the so-called capsid breathing). Limited proteolysis-MS (and HX-MS, see below) has been used to identify dynamic regions in virus capsids. The limited proteolysis approach is based on the identification by MS of the peptides produced in a time-dependent, limited proteolytic attack on a virus particle. Particle regions that are conformationally more dynamic will, in general, expose cleavage sites more readily and will be cleaved earlier. The size of the fragments of the capsid proteins obtained at different times reveal the sites cleaved in the particle, and thus the dynamic regions [59].

By using this approach, the biologically relevant, transient externalization of (normally internal) terminal segments of the capsid proteins has been detected in picornaviruses (poliovirus, HRV) [60], nodaviruses (flock house virus, FHV) and HBV; in general, the tertiary and the quaternary structure of the virions can be mapped (see [61] and references therein). In addition, by using this approach, local and global breathing have been shown to differ between different states of a viral capsid, for example during maturation of tetraviruses (*Nudaurelia capensis* ω virus). The binding of the viral RNA genome to the capsid inner wall in FHV [62] or of hydrophobic compounds to pockets in the capsid of HRV [63] also impaired local or global breathing. Some of these effects on capsid dynamics have been shown to be important for virus biology (see Chap. 10) or the action of antiviral agents (see Chap. 20): whereas some anti-virals stabilized the capsids against proteolysis, others had the opposite effect (see [61] and references therein).

HX-MS

The basis of hydrogen exchange in biomolecules including proteins and protein complexes was explained in Chap. 5 as a part of the description of the HX-NMR technique. The pattern of exchanged protons with deuterium is dependent on the conformational dynamism of the region considered, and the number of incorporated deuterium nuclei in each region analyzed can be used to estimate its conformational dynamism. Since the mass of deuterium is larger (roughly twice) than that of a proton, the number of incorporated deuterium nuclei can be easily determined with precision by the current mass spectrometers [59, 64]. Several examples of the use of HX-MS to investigate the conformational stability or dynamics of virus particles, and specific regions in them are described next.

HX-MS studies on HRV indicated that intersubunit interfaces are conformationally rather static while regions around the capsid fivefold axes, involved in the translocation of polypeptide segments are more dynamic [65]. Some aspects of the conformational changes of the phage HK97 capsid during maturation (see Chap. 13) have been described by the combined use of protein engineering and HX-MS. During maturation, the HK97 capsid undergoes an expansion of roughly 20 % in diameter. The number of exchanged hydrogens suggested that a pre-expanded particle suffers significant refolding and twisting of the tertiary structure to accommodate the expansion [66, 67]. The differences in solvent-accessibility observed between capsid subunits in different quasi-equivalent conformations disappeared after the conformational transition that led to a first maturation intermediate. However, other quaternary interactions close to the icosahedral three-fold axes were preserved during maturation. These three-fold “staples” are formed by residues Arg, Glu and a metal binding site; mutations at those sites facilitated the conformational transition, as revealed by a larger proportion of protons being exchanged, and a decrease in the *in vitro* expansion rate. Taken together, these and other results suggested that the HK97 capsid, before the expansion, is in a high-energy metastable state.

The dynamics of maturation of the HIV-1 capsid has also been studied by HX-MS; both, the stability of the immature (either non-mutated or carrying a mutation that prevented the maturation cleavage) and the mature capsids have been described [68–70]. Among other results, it was found that the mature capsid is more stable (it exchanges more slowly) than the immature capsid. Changes in conformational dynamism during maturation and expansion of the bacteriophage P22 capsid have also been analyzed by HX-MS [71]. The P22 capsid protein consists of two domains connected by a stable linker. The HX pattern indicated that during maturation the C-terminal region maintained its conformational flexibility (dynamics), whereas the N-terminal region was made conformationally less dynamic. These experiments were consistent with the possibility that maturation increases the number of inter-subunit contacts.

HX-MS has been used also to investigate structural and dynamic aspects of nucleic acid encapsidation in double-stranded RNA or DNA bacteriophages. These phages rely on an active packaging mechanism, in which an ATP-dependent molecular motor forces the nucleic acid into the preformed procapsid (see Chap. 12). These packaging motors consist of at least two regions: the ATPase complex and the connector portal complex; the portal complex replaces the capsid proteins at one of the five-fold vertices of the capsid. The use of HX-MS revealed, for example, intersubunit interactions between the portal and the capsid protein, and the number of subunits forming the portal (in P22, the portal when produced *in vitro* is a mixture of 11- and 12-mers) [70, 72, 73]. Conformational changes related with nucleic acid packaging have also been investigated by HX-MS. For instance, the structure of the packaging motor of the double-stranded RNA bacteriophage Φ8 was not structurally modified during ATP-binding and hydrolysis (as suggested by the absence of large changes in the HX pattern); conversely, RNA-binding caused changes in kinetics that allowed the identification of several nucleotide-binding sites [74]. Finally, the HX-MS results suggested than in the presence of the motor, the capsid is conformationally less dynamic [75].

6.5.4 Virus-Ligand Interactions

Certain MS conditions preserve non-covalently bound complexes including ligand-bound viral particles. Thus, the interactions between virus particles or their components and other biomolecules can be determined using MS. I shall describe several examples in the following paragraphs.

In several studies interactions between lipids or membrane proteins and viral capsids were detected by MS, and in some cases they were confirmed by other techniques such as EM (Chap. 3). For instance, the outer membrane proteins OmpA and OmpC protein from *Shigella*, the bacterial host of phage Sf6, have been detected in complex with the Sf6 capsid [76]. Also, MS in conjunction with enzyme digestion has shown that most of the proteins forming the capsid of iridovirus are able to bind lipids [77].

Some viral capsids recruit proteins from the host cells or other viral proteins required for infection. For instance, when HSV-1 nucleocapsids exit the nucleus, they become surrounded by a primary lipid envelope. The use of MS to detect mecromolecular interactions has allowed the identification of most of the proteins forming those primary enveloped virions: a glycoprotein, a cellular protein and a tegument protein [78, 79].

MS has also proved useful, together with electrophoresis, in analyzing the composition of viral particles and their covalent modifications. In many of these proteomic studies, cross-linking or proteolysis have been used to facilitate protein identification (see [61] for a review). These studies have helped in the determination of the stoichiometry of complexes involving viral proteins and/or nucleic acids. For instance, recently, a map of the interactions between HIV and human proteins complexes has been obtained; as many as 497 interactions, involving 435 human proteins, were identified [80].

6.5.5 *Understanding Viruses: Some Major Contributions of MS*

MS has become a very important technique for different analyses on the structure or properties of virus particles, as proven by the examples mentioned above and others. MS can be used to determine the size and composition of capsid building blocks, assembly intermediates, assembled capsids and virions, and aberrant viral particles. MS can be used also to follow structural changes in response to external factors, during assembly or disassembly of virus particles. MS has proven instrumental in the analysis of viral capsid breathing and in detecting conformationally dynamic regions in viral particles, and in the role of molecules or conditions facilitating or hampering those changes.

6.6 Perspectives

I foresee that CD, fluorescence spectroscopy and MS will hold a critical supporting role in structural virology. Fluorescence and CD are highly valuable for a first or rapid general structural characterization of viral proteins and viral particles; the information provided may prove particularly important in those cases where the protein or viral particle can not be crystallized, or can not be tackled by liquid NMR or solid-state NMR analyses. CD and, especially, fluorescence spectroscopy provide adequate probes for the rapid detection and following of some conformational rearrangements in virus particles. MS has become a critically important technique in several fields including structural virology. I anticipate that its use in this area will become widespread for the identification and study of the conformational dynamics of virus capsids, and for proteomic analyses in virology, including the detection of interactions between viral proteins and virus particles and cellular biomolecules which may participate in virus replication and assembly factories (Chap. 14). The power of the techniques described in this chapter is maximized in those cases where they can be combined (see Chap. 7) with structural techniques including EM (Chap. 3), X-ray crystallography (Chap. 4) and/or NMR (Chap. 5).

Acknowledgements I thank Luis Pérez and Francisco N. Barrera for careful reading of the manuscript, suggestions and new ideas. I apologize to those colleagues whose works were not cited due to lack of space or inadvertent omission. Work in the author's laboratory was supported partially by the Spanish Ministerio de Ciencia e Innovación (CTQ2011-24393 and CSD2008-00005), the FIPSE foundation (2006/0008), and intramural BIFI 2011 projects.

References and Further Reading

1. Lakowicz JR (1999) Principles of fluorescence spectroscopy, 2nd edn. Plenum, New York
2. Albani JR (2007) Principles and application of fluorescence spectroscopy. Blackwell, Oxford
3. Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY (1995) Understanding, improving and using green fluorescent proteins. TIBS 20:448–455

4. Stepanenko OV, Stepanenko OV, Scherbakova DM, Kuznetsova IM, Turoverov KK (2011) Modern fluorescence proteins: from chromophore formation to novel intracellular applications. Biotechniques 51:313–327
5. Morrison RN, Boyd RN (1992) Organic chemistry, 6th edn. Prentice Hall, Englewood Cliffs
6. Wang AHJ, Quigley GJ, Koplpak FJ, Crawford JL, van Boom JH, van der Marel G, Rich A (1979) Molecular structure of a left-handed double helical DNA fragment at atomic resolution. Nature 282:680–686
7. Schmidt W (2005) Optical spectroscopy in chemistry and life sciences. Wiley-VGH, Weinheim
8. Atkins P, de Paula J (2006) Physical chemistry for the life sciences. Oxford University Press, Oxford
9. Tinoco J, Sauer K, Wang Y (1996) Physical Chemistry: principles and applications for the life sciences, 3rd edn. Prentice Hall, Englewood Cliffs
10. Crabbe P (1967) Optical rotatory dispersion and optical circular dichroism in organic chemistry. Top Sterochem 1:93–98
11. Kelly SM, Price NC (2000) The use of circular dichroism in the investigation of protein structure and function. Curr Protein Pept Sci 1:349–384
12. Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res 32: W668–W673
13. Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers 89:392–400
14. Clarke J, Fersht AR (1993) Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. Biochemistry 32:4322–4329
15. Privalov PL (1992) Physical basis of the stability of the folded conformations of proteins. In: Creighton TE (ed) Protein folding. Freeman, New York
16. Wingfield PT, Stahl SJ, Williams RW, Steven AC (1995) Hepatitis core antigen produced in *Escherichia coli*: subunit composition, conformational analysis, and *in vitro* capsid assembly. Biochemistry 34:4919–4932
17. Mateu MG (2002) Conformational stability of monomeric and dimeric forms of the C-terminal domain of human immunodeficiency virus-1 capsid protein. J Mol Biol 318:519–531
18. Gonçalves RB, Mendes YS, Soares MR, Katpally U, Smith TJ, Silva JL, Oliveira AC (2007) VP4 protein from human rinovirus 14 is elapsed by pressure and locked in the capsid by the antiviral compound WIN. J Mol Biol 366:295–306
19. Kneijzlik Z, Ulbrich P, Strohalm M, Lastůvková H, Kodícek M, Sakalian M, Ruml T (2009) Conformational changes of the N-terminal part of the mason-Pfizer monkey virus p12 protein during multimerization. Virology 393:168–176
20. Datta SA, Temeselew LG, Christ RM, Soheilian F, Kamata A, Mirro J, Harvin D, Nagashima K, Cachau RE, Rein A (2011) On the role of SP1 domain in HIV-1 particle assembly: a molecular switch? Virology 85:4111–4121
21. Carreira A, Menéndez M, Reguera J, Almendral JM, Mateu MG (2004) *In vitro* disassembly of a parvovirus capsid and effect on capsid stability of heterologous peptide insertions in surface loops. J Biol Chem 279:6517–6525
22. Katen S, Zlotnick A (2009) The thermodynamics of virus capsid assembly. Methods Enzymol 455:395–417
23. Ceres P, Stray SJ, Zlotnick A (2004) Hepatitis B virus capsid assembly is enhanced by naturally occurring mutation F97L. J Virol 78:9538–9543
24. del Alamo M, Neira JL, Mateu MG (2003) Thermodynamic dissection of a low affinity protein-protein interface involved in human immunodeficiency virus assembly. J Biol Chem 278:27923–27929
25. Green B, King J (1999) Folding and stability of mutant scaffolding proteins defective in P22 capsid assembly. J Biol Chem 274:16141–16146

26. Umahankar M, Murthy MRN, Savithiri HS (2003) Mutation of interfacial residues disrupts subunit folding and particle assembly of *Physalis* mottle tymovirus. J Biol Chem 278:6145–6152
27. Reguera J, Carreira A, Riolobos L, Almendral JM, Mateu MG (2004) Role of interfacial amino acid residues in assembly, stability and conformation of a spherical virus sample. Proc Natl Acad Sci U S A 101:2724–2729
28. Kirkitadze MD, Barlow PN, Price NC, Kelly SM, Boutell CJ, Rixon FJ, McClelland DA (1998) The herpes simplex virus protein, VP23, exists as a molten globule. J Virol 72:10066–10072
29. Lee KK, Gan L, Tsuruta H, Hendrix RW, Duda RL, Johnson JE (2004) Evidence that a local refolding event triggers maturation of HK97 bacteriophage capsid. J Mol Biol 340:419–433
30. Xue B, Willians RW, Oldfield CJ, Goh GK, Dunker AK, Uversky VN (2010) Viral disorder or disordered viruses: do viral proteins posses unique features? Protein Pept Lett 17:932–951
31. Rexroad J, Martin TT, McNeilly D, Godwin S, Middaugh CR (2006) Thermal stability of adenovirus type 2 as a function of pH. J Pharm Sci 95:1469–1479
32. Kang H, Yu J, Jung G (2008) Phosphorylation of hepatitis B virus core C-terminally truncated protein (Cp149) by PKC increases capsid assembly and stability. Biochem J 416:47–54
33. Waldron TT, Murphy KP (2003) Stabilization of proteins by ligand binding: application to drug screening and determination of unfolding energetics. Biochemistry 42:5058–5064
34. Sun J, Yu JS, Jin S, Zha X, Wu Y, Yu Z (2010) Interactions of synthetic HPV-16 capsid peptides with heparin: thermodynamic parameters and binding mechanism. J Phys Chem B 114:9854–9861
35. Teschke CM, King J, Prevelige PE Jr (1993) Inhibition of viral capsid assembly by 1,1′-bis(4-anilinonaphthalene-5-sulfonic acid). Biochemistry 32:10658–10665
36. Zlotnich A, Mukhopadhyay S (2010) Virus assembly, allostery and antivirals. Trends Microbiol 19:14–23
37. Du S, Betts L, Yang R, Shi H, Concel J, Ahn J, Aiken C, Zhang P, Yeh JL (2011) Structure of the HIV-1 full-length capsid protein in a conformational trapped un assembled state induced by small-molecule binding. J Mol Biol 406:371–386
38. Zlotnick A, Lee A, Bourne CR, Johnson JM, Domanico PL, Stray SJ (2007) *In vitro* screening for molecules that affect virus capsid assembly (and other protein association reaction). Nat Protoc 2:490–498
39. Stray SJ, Johnson JM, Kopek BJ, Zlotnick A (2006) An *in vitro* fluorescence screen to identify antivirals that disrupt hepatitis B virus capsid proteins. Nat Biotechnol 24:358–362
40. Johnstone RAW, Rose ME (1996) Mass spectroscopy for chemist and biochemist. Cambridge University Press, Cambridge
41. Gault VA, McClenaghan NH (2009) Understanding bioanalytical chemistry: principles and applications. Wiley and Blackwell, Oxford
42. Ashcroft AE (2005) Recent developments in electrospray ionization mass spectrometry: non-covalently bound protein complexes. Nat Prod Rep 22:452–464
43. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246:64–71
44. Wilm MS, Mann M (1994) Electrospray and Taylor-cone theory, Dole's beam of macromolecules at last? Int J Mass Spectrom Ion Proc 136:167–180
45. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. Anal Chem 60:2299–2301
46. Tanaka K, Waki H, Ido Y, Akita S, Yoshida T (1988) Protein and polymer analysis up to m/z 100000 by laser ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 2:151–153
47. Siuzdak G (1996) Mass spectrometry and viral analysis. Chem Biol 9:707–715
48. Sachse C, Chen JZ, Coureux PD, Stroupe ME, Fandrich M, Grigorieff N (2007) High-resolution electron microscopy of helical specimens: a fresh look at tobacco mosaic virus. J Mol Biol 371:812–835

49. DeLano WL (2002) The PyMOL molecular graphics system. De Lano Scientific, San Carlos
50. Qu C, Liljas L, Opalka N, Brugidou C, Yeager M, Beachy RN, Fauquet CM, Johnson JE, Lin T (2000) 3D domain swapping modulates the stability of members of an icosahedral virus group. Struct Fold Des 8:1095–1103
51. Fuesternau SD, Benner WH, Thomas JJ, Brugidou C, Bothner B, Siuzdak G (2001) Mass spectrometry of an intact virus. Angew Chem Int Ed 40:541–544
52. Uetrecht C, Versluis C, Watts NR, Roos WH, Wuite GJ, Wingfield PT, Steven AC, Heck AJ (2008) High resolution mass spectrometry of viral assemblies: molecular composition and stability of dimorphic hepatitis B virus capsids. Proc Natl Acad Sci U S A 105:9216–9220
53. Shoemaker GK, van Dujin E, Crawford SE, Uetrecht C, Baclayon M, Roos WH, Wuite GJ, Estes MK, Prasad BV, Heck AJ (2010) Norwalk virus assembly and stability monitored by mass spectrometry. Mol Cell Proteomics 9:1742–1751
54. Ashcroft AE, Lago H, Macedo JM, Horn WT, Stonehouse NJ, Stockley PG (2005) Engineering thermal stability in RNA phage capsids *via* disulphide bonds. J Nanosci Nanotechnol 5:2034–2041
55. Stockley PG, Rolfsson O, Thompson GS, Basnak G, Francese S, Stonehouse NJ, Homans SW, Ashcroft AE (2007) A simple, RNA-mediated allosteric switch controls the pathway to formation of a $T = 3$ capsid. J Mol Biol 369:541–552
56. Fabris D, Chaudhari P, Hagan N, Turner K (2007) Functional investigation of retroviral protein ribonucleic acid complexes by nanospray Fourier transform ion cyclotron resonance mass spectrometry. Eur J Mass Spectrom 13:29–33
57. Knapman TW, Morton VL, Stonehouse NJ, Stockley PG, Ashcroft AE (2010) Determining the topology of virus assembly intermediates using ion mobility spectrometry-mass spectrometry. Rapid Commun Mass Spectrom 24:3033–3042
58. Johnson JE (2003) Virus particle dynamics. Adv Protein Chem 64:197–218
59. Bothner B, Hilmer JK (2011) Probing viral capsids in solution. In: Agbandje-McKenna M, McKenna R (eds) Structural virology. RSC Publishing, London, pp 41–61
60. Bothner B, Dong XF, Bibbs L, Johnson JE, Siuzdak G (1998) Evidence of viral capsid dynamics using limited proteolysis and mass spectrometry. J Biol Chem 272:672–676
61. Morton VL, Stockley PG, Stonehouse NJ, Ashcroft AE (2008) Insights into virus capsid assembly from non-covalent mass spectrometry. Mass Spectrom Rev 27:575–595
62. Bothner B, Schneemann A, Marshall D, Reddy V, Johnson JE, Siuzdak G (1999) Crystallographically identical virus capsids display different properties in solution. Nat Struct Biol 6:114–116
63. Lewis JK, Bothner B, Smith TJ, Siuzdak G (1998) Antiviral agent blocks breathing of the common cold virus. Proc Natl Acad Sci U S A 95:6774–6778
64. Tuma R, Coward LU, Kirk MC, Barnes S, Prevelige PE Jr (2001) Hydrogen-deuterium exchange as a probe of folding and assembly in viral capsids. J Mol Biol 306:389–396
65. Wang L, Smith DL (2005) Capsid structure and dynamics of a human rhinovirus probed by hydrogen exchange mass spectrometry. Protein Sci 14:1661–1672
66. Gertsman I, Fu CY, Huang R, Kornives EA, Johnson JE (2010) Critical salt bridges guide capsid assembly, stability and maturation behavior in bacteriophage HK97. Mol Cell Proteomics 9:1752–1763
67. Gertsamn J, Komives EA, Johnson JE (2010) HK97 maturation studies by crystallography and $H/^2H$ exchange reveals the structural basis for exothermic particle transitions. J Mol Biol 397:560–574
68. Monroe EB, Kang S, Kyere SK, Li R, Prevelige PE Jr (2010) Hydrogen/deuterium exchange analysis of HIV-1 capsid assembly and maturation. Structure 18:1483–1491
69. Cortines JR, Monroe EB, Kang S, Prevelige PE Jr (2011) A retroviral chimeric capsid protein reveals the role of the N-terminal β-hairpin in mature core assembly. J Mol Biol 410:641–652
70. Fu CY, Uetrecht C, Kang S, Morais MC, Heck AJ, Walter MR, Prevelige PE Jr (2010) A docking model based on mass spectrometric and biochemical data describes phage packaging motor incorporation. Mol Cell Proteomics 9:1764–1773

71. Kang S, Prevelige PE Jr (2005) Domain study of bacteriophage P22 coat protein and characterization of the capsid lattice transformation by hydrogen/deuterium exchange. J Mol Biol 347:935–948
72. Poliakov A, van Dujin E, Lander G, Fu CY, Johnson JE, Prevelige PE Jr, Heck AJ (2007) Macromolecular mass spectrometry and electron microscopy as complementary tools for investigation of the heterogeneity of the bacteriophage portal assemblies. J Struct Biol 157:371–383
73. Kang S, Poliakov A, Sexton R, Renfrow MB, Prevelige PE Jr (2008) Probing conserved helical modules of portal complexes by mass spectrometry-based hydrogen/deuterium exchange. J Mol Biol 381:772–784
74. Lisal J, Kainov DE, Lam TT, Emmett MR, Wei H, Gottlieb P, Marshall AG, Tuma R (2006) Interaction of packaging motor with the polymerase complex of dsRNA bacteriophage. Virology 351:73–79
75. Lisal J, Lam TT, Kainov DE, Emmett MR, Marshall AG, Tuma R (2005) Functional visualization of viral molecular motor by hydrogen deuterium exchange reveals transient states. Nat Struct Mol Biol 12:460–466
76. Zhao H, Sequeira RD, Galeva NA, Tang L (2011) The host outer membrane proteins OmpA and OmpC are associated with the *Shigella* phage Sf6 virion. Virology 409:319–327
77. Wu J, Chan R, Wenk MR, Hew C-L (2010) Lipidomic study of intracellular Singapore group iridovirus. Virology 399:248–256
78. Padula ME, Sydnor ML, Wilson DW (2009) Isolation and primary characterization of herpes simplex virus 1 primary enveloped virions from the perinuclear space. J Virol 83:4757–4765
79. Harper AL, Meckes DG Jr, Marsh JA, Ward MD, Yeh PC, Baird NL, Wilson CB, Semmes OJ, Wills JW (2010) Interaction domain of the UL16 and Ul21 tegument proteins of herpes simplex virus. J Virol 84:2963–2971
80. Jäger S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, Shales M, Mercenne G, Pache L, Li K et al (2011) Global landscape of HIV-human protein complexes. Nature 481:365–370

Further Reading

Baker M (2010) Mass spectrometry for biologists. Nat Meth 7:157–161
Burlingame AL, Carr SA, Baldwin MA (eds) (2000) Mass spectrometry in biology. Humana Press, Totowa
Kelly SM, Ness TJ, Price NC (2005) How to study proteins by circular dichroism. Biochim Biophys Acta 1751:119–139

Also recommended for further reading are references [1, 2, 7–9, 40, 41, 43] *listed above.*

Chapter 7
Combined Approaches to Study Virus Structures

Daniel Badia-Martinez, Hanna M. Oksanen, David I. Stuart, and Nicola G.A. Abrescia

Abstract A virus particle must work as a safe box for protecting its genome, but at the same time it has to undergo dramatic conformational changes in order to preserve itself by propagating in a cell infection. Thus, viruses are miniaturized wonders whose structural complexity requires them to be investigated by a combination of different techniques that can tackle both static and dynamic processes. In this chapter we will illustrate how major structural techniques such as X-ray crystallography and electron microscopy have been and can be combined with other techniques to determine the structure of complex viruses. The power of these hybrid method approaches are revealed through the various examples provided.

Keywords Hybrid methods • X-ray crystallography • Electron microscopy • Cryo-electron microscopy • Electron tomography • Cryo-electron tomography • Small-angle X-ray scattering • Virus • Bacteriophage • Capsid • Mutagenesis • Dissociation • Crystal structure • Fitting

---

D. Badia-Martinez
Structural Biology Unit, CICbioGUNE, CIBERehd, Bizkaia Technology Park, 48160 Derio, Spain

H.M. Oksanen
Institute of Biotechnology and Department of Biosciences, Viikki Biocenter, University of Helsinki, P.O. Box 56, Viikinkaari 5, 00014 Helsinki, Finland

D.I. Stuart
Division of Structural Biology, The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Headington, Oxford OX3 7BN, UK

Diamond Light Source Ltd, Diamond House, Harwell Science and Innovation Campus, Didcot, UK

N.G.A. Abrescia (✉)
Structural Biology Unit, CICbioGUNE, CIBERehd, Bizkaia Technology Park, 48160 Derio, Spain

Ikerbasque, Basque Foundation for Science, 48011 Bilbao, Spain
e-mail: nabrescia@cicbiogune.es

M.G. Mateu (ed.), *Structure and Physics of Viruses: An Integrated Textbook*,
Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_7,
© Springer Science+Business Media Dordrecht 2013

Abbreviations

2D	Two dimensional
3D	Three dimensional
BMV	Bromegrass mosaic virus
EM	Electron microscopy
ET	Electron tomography
Fab	Antigen-binding antibody fragment
FMDV	Foot-and-mouth disease virus
HCV	Hepatitis C virus
HCV-LP	HCV-like particle
HIV	Human immunodeficiency virus
HRV-16	Human rhinovirus 16
HSV	Herpes simplex virus
MR	Molecular replacement
NCS	Non-crystallographic symmetry
NMR	Nuclear magnetic resonance
RVFV	Rift Valley fever virus
SANS	Small-angle neutron scattering
SAXS	Small-angle X-ray scattering
SBMV	Southern bean mosaic virus
SeMet	Seleno-methionine
SIV	Simian immunodeficiency virus
TBSV	Tomato bushy stunt virus
TEM	Transmission electron microscopy
TMV	Tobacco mosaic virus
TNV	Tobacco necrosis virus
VLP	Virus-like particle
WNV	West Nile virus

7.1 Introduction: The ‘Multi-Disciplinary Approach’ Concept in Structural Virology

7.1.1 Early Structural Studies on Viruses

In previous chapters several independent approaches to study virus structure including electron microscopy (EM, Chap. 3), X-ray crystallography (Chap. 4), nuclear magnetic resonance spectroscopy (NMR, Chap. 5), mass spectrometry (MS) and other spectroscopic methods (Chap. 6) have been described. In this chapter, we will illustrate how some of those methods can be combined to provide further insights into virus structures. To clarify how these initially independent methods have converged, complemented and combined with each other in the study of viruses, we will provide a brief historical perspective and a recapitulation of some of these structural methods as applied to virology (Fig. 7.1).

Structural virology began when Ernst Ruska built the first electron microscope and together with his brother, medical doctor, Helmut Ruska initiated the imaging of viruses and other submicroscopic structures during the years 1938 and 1939. Alongside this, the principles for the small-angle X-ray scattering (diffraction) (SAXS) technique were proposed by André Guinier (see Sect. 7.2.2). In 1935, Wendell Stanley provided the conditions for crystallization of the first virus, the tobacco mosaic virus (TMV), a plant-infecting virus. For his studies in producing pure enzymes and viruses, in 1946 he was awarded with the Nobel Prize in Chemistry. Also Ruska received the Nobel Prize but in Physics for his fundamental contribution in electron optics, incredibly, 55 years after building of his first microscope! In 1937, Bawden and Pirie crystallized another plant virus, tomato bushy stunt virus (TBSV) but it wasn't until 1941 that crystals of TBSV and TMV produced clear X-ray diffraction patterns.

The 1950s saw the beginning of the modern biology with the determination of the three-dimensional (3D) structure of the DNA by James Watson and Francis Crick and the first high-resolution X-ray structures of proteins, myoglobin by John Kendrew and hemoglobin by Max Perutz. Interestingly, in 1956 Watson and Crick also proposed the principles of virus organization. According to their hypothesis based on electron micrographs of plant viruses and their propensity to form crystals and concepts of genetic economy, simple viruses were formed by capsid of multiple copies of relatively small proteins arranged in a symmetrical fashion, providing identical protein contacts among their neighbours. They suggested that a cubic symmetry or a cylindrical shell would satisfy the above requirements whilst allowing a larger volume of nucleic acid to be packaged. The case for icosahedral symmetry of viruses was made by John Finch and Aaron Klug in 1959 in unveiling the organisation of poliomyelitis virus by X-ray diffraction in 1959. Caspar and Klug continued on this idea and formulated the quasi-equivalence theory indicating that certain multiples of 60 subunits could also form viral capsids but the equivalent interactions among neighbour subunits could no longer exist. Klug's thinking on this was inspired by the work of Buckminster Fuller in particular by his book *The Dymaxion World of Buckminster Fuller*. Caspar and Klug based the quasi-equivalent theory on the fact that the capsomers are adaptable molecules that organize in different ways within a highly ordered structure thus introducing the concept of protein flexibility to explain these different arrangements (see Chap. 2). Methods development for 3D reconstruction from electron micrographs began in 1964 when Aaron Klug and Jacob Berger introduced the use of Fourier transform in micrograph analysis. This analysis was extended to object with rotational symmetry, helical symmetry (for rod-shaped viruses) and finally to icosahedral specimens (for an historical account see 'Further reading' at the end of this chapter). In the same period Walter Hoppe, who contributed to the development of methods and instrumentation in X-ray analysis moved to the electron microscopy field transferring the same momentum and know-how in the advances of electron microscopy methods (see the ptychography technique in Sect. 7.7.2).

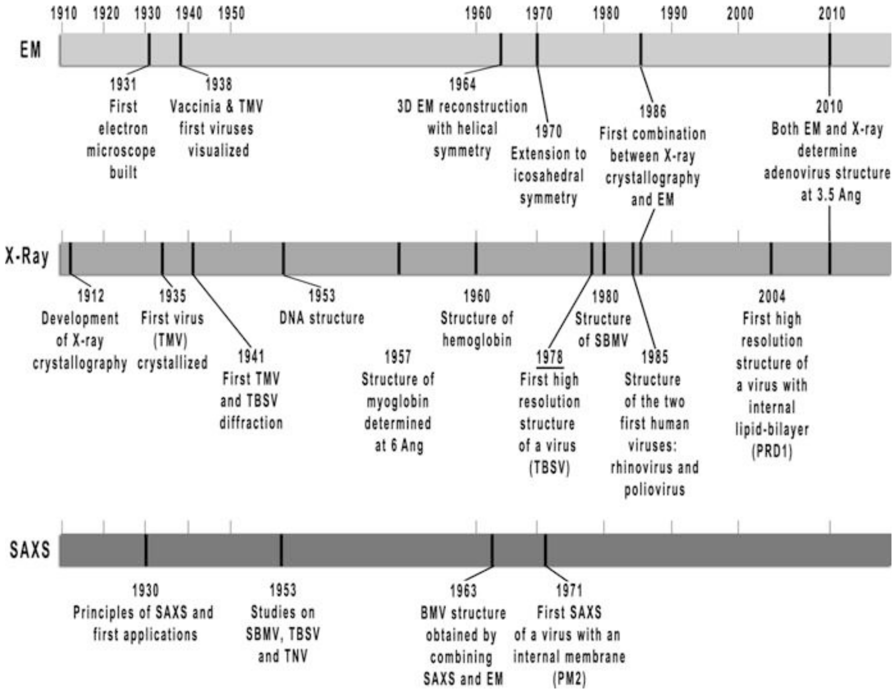


Fig. 7.1 Timeline of the major achievements in structural virology by EM, X-ray crystallography and SAXS techniques for the determination of virus structures. Due to their relevance the DNA and myo-hemoglobin structures have been included. Tobacco mosaic virus (*TMV*); Tobacco necrosis virus (*TNV*); tomato bushy stunt virus (*TBSV*); lipid-containing bacteriophage PRD1 (*PRD1*); southern bean mosaic virus (*SBMV*); bromegrass mosaic virus (*BMV*); lipid-containing bacteriophage PM2 (*PM2*)

Together with X-ray and EM, also small-angle scattering (using both X-rays [SAXS], and neutrons, [small-angle neutron scattering, or SANS]) was contributing to the analysis of virus structures. However in the years that followed, the impact of SAXS as a primary method became weaker leaving the leading role in structural virology to the other two structural techniques (Fig. 7.1). In parallel to the above developments, efforts were made to improve the diffraction of virus crystals and to devise ways to analyse symmetric structures that ultimately culminated into the structure of TBSV at 2.9-Å resolution in 1978 [1] and two years later into the structure of southern bean mosaic virus (SBMV) at 2.8-Å resolution [2]. Since then, both virus crystallography and electron microscopy have been fundamental techniques in providing structural insights in complex viruses. The techniques are complementary and well suited to hybridisation.

7.1.2 The Emerging Concept of Hybrid Methods

To grasp the concept of hybrid methods in structural virology one has to go back to the early culture of structural biology. The methods were developed to answer fundamental biological questions but the X-ray diffractometers and microscopes were designed, built and maintained by engineers, physicists and biophysicists. The interdisciplinary expertises gathered in pioneering laboratories, in, among others, Birbeck and King's College in London (JD Bernal and Rosalind Franklin), Oxford University (Dorothy Hodgkin), the Royal Institution London (Lawrence Bragg and David Phillips) and the MRC Unit for the Study of the Molecular Structure of Biological Systems at Cambridge University (Max Perutz) around the middle of the twentieth century forged the spirit and the ambition to tackle the unexplored complexity of biology by complementary approaches (see Max Perutz's video interview and commentary on: http://www.youtube.com/watch?v=5I4eq63-3_I).

Thus, from one side, complex biological systems were reduced to smaller pieces (proteins or protein domains), in a reductionist (or "top-to-bottom") approach, studied separately and then, possibly, put back again in an overall framework (see also Chap. 5 regarding this approach in virus structure studies using NMR). On the other side, the complexity of larger macromolecular assemblies such as viruses could be looked at head-on thanks to the diverse expertise in those laboratories. So, hybrid methods and expertises were from the start supporting each other, validating the results obtained by the individual techniques.

Over the past three decades the multi-disciplinary approach has moved on from its original rationale towards the use of hybrid methods in predicting biological interactions, in uncovering dynamic processes and in bridging different resolution scales in molecular and cellular processes. The study of human adenovirus, a complex non-membrane icosahedral dsDNA virus (see below and Chap. 11) provides a good example of this. Indeed, this is the first case of combining X-ray crystallography and electron microscopy data, where the crystal structure of the viral capsid protein, assembled as stable capsomeric hexamers (hexons), was used to build an atomic model of the entire virion by exploiting the capsomer arrangement revealed by electron microscopy [3]. Since then the individual methods have become more powerful and the combination of (cryo-)EM and X-ray crystallography is a method of choice not just for many viruses but generally within structural biology.

In this chapter we will focus on describing how hybrid methods, with emphasis on the combination of X-ray crystallography and EM, have shaped our understanding of virus structures. We will provide examples of their contributions to the structure solution of very large and complex viruses and briefly discuss the challenges ahead, highlighting the power and vigour of methodological interactions in structural virology.

7.2 Some Classical Methods in Structural Virology: A Brief Overview

In this section we will recapitulate the three major structural techniques that have primarily contributed to the understanding of virus structures: EM, SAXS/SANS and X-ray crystallography. EM and X-ray crystallography as applied to viruses have been described extensively in Chaps. 3 and 4, respectively. Here we simply underline those concepts that link the methods.

7.2.1 *Electron Microscopy*

Transmission electron microscopy (TEM) relies on an electron beam passing through a virus sample either negatively stained to enhance the contrast or (in cryo-EM) vitrified at <100 K to ameliorate the radiation damage (vitrification is usually through simple flash freezing, but freezing rates may be enhanced by a simultaneous pressure wave; see Chap. 3). Electrons scattered elastically *via* the Coulomb potential of the atoms composing the virus (or any biological/material sample) are combined *via* magnetic lenses to generate a magnified image of the specimen. Depending on the TEM set-up data can be recorded in real space (images) or a reciprocal space (diffraction). In mathematical terms what allows the scattering process to be formalized in one or the other space is the Fourier transform (see also Chaps. 3 and 4). Since the energy of the electrons is very high their equivalent wavelength is much shorter than that for the X-rays used in crystallography. As a consequence in the case of image-mode, the result is a two-dimensional (2D) projection (real space) conceptually equivalent to the optical image obtained by cameras and visible light microscopes. In reciprocal space this corresponds to a plane of information which passes through the origin of reciprocal space (a central section) and information on both amplitude and phase of the transform of the elastically scattered electrons can be derived (Fig. 7.2). In the diffraction-mode only the amplitudes can be directly calculated from the measured intensity of diffraction spots (reciprocal space) but with higher precision and resolution than the corresponding values extracted from real-space images, which provide the less precise phase information (Fig. 7.2).

The dual-use of TEM in image- and diffraction- modes on the same sample constitutes the field of 2D electron crystallography. However, to collect electron diffraction images it is fundamental that the specimen forms a 2D ordered lattice, a requirement not needed for imaging mode. By combining these modes, 2D-electron crystallography circumvents the ‘phase-problem’ that X-ray crystallography must inevitably face (see Chap. 4). Each EM snapshot is equivalent to a ‘still’ X-ray image. Interestingly, it has proved much easier to generate 3D crystals than 2D (perhaps especially so for viruses which are often isometric), so the field of 2D crystallography remains much less developed.

Viruses have been investigated by EM since its early days due to their large and regular shape and relative ease of purification. Indeed, since the visualization of the first viruses in 1938–39 [4], a head-to-head race began in the technological advances of each of the major structural techniques at the time: EM, SAXS and X-ray crystallography (Figs. 7.1 and 7.2). Electron microscopy remains a workhorse method in several different aspects of virology from virus detection and identification to the 3D structure determination and analysis of virus-cell interactions. Continuing developments mean that nowadays electron microscopy can provide near-atomic resolution structures of viruses where atoms though not resolved, can be inferred [5] (see Chap. 3). The highest resolution analyses come from combining many projections of the virion collected from different particles thrown down at random;

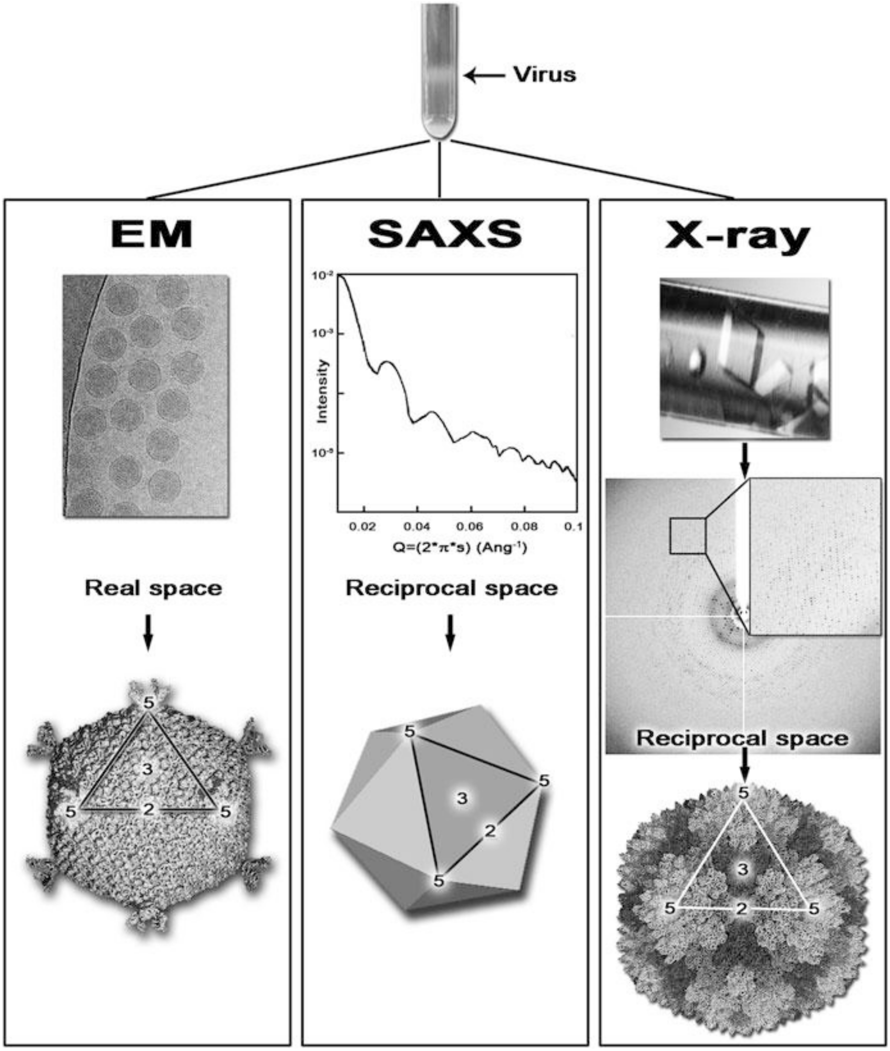


Fig. 7.2 Three major structural techniques at glance. Schematic representation of the workflow of virus structure determination in EM, SAXS and X-ray. All three techniques require purified virus; at the top, a *black arrow* indicates virus band resulting from ultracentrifugation. In EM the purified virus is placed on a microscopy grid, flash-frozen and then the 2D projections of the virus are visualized in an electron microscope producing real space images (*top*); post-processing of these images provides the virus structure (*bottom*). In SAXS the purified virus in solution is irradiated by an X-ray beam producing a low-angle scattering curve (*top*) containing raw data in reciprocal space; post-processing of the data produces the virus overall shape and molecular architecture. In X-ray crystallography the purified virus is used to obtain virus crystals (*top*) that are then diffracted producing diffraction images (*centre*) containing raw data in reciprocal space; once the data processing has been completed and the phase-problem solved, the virus structure is obtained at atomic resolution (*bottom*). The outlined region in the virus models in EM, SAXS and X-ray panels delineates delineates one of the 20 viral facets, and numbers indicate icosahedral symmetry axes (two, 2-fold; three, 3-fold; five, 5-fold). Each technique provides structural information in a different resolution range; in some cases these ranges overlap (see Fig. 7.4)

however in recent years it has become possible to get useful but lower resolution (perhaps 30-Å) data from an individual virus particle by collecting a full tilt series and doing a tomographic reconstruction (electron tomography, ET) (Chap. 3). Furthermore the resolution of both of these methods (EM and ET) can be enhanced for symmetrical virus particles by applying symmetry constraints (see Sect. 7.5.2). X-ray and electron methods begin to almost fully overlap when we consider X-ray microscopy, an emerging technique which is beyond the scope of this review to discuss, but it can simply be seen as a lower resolution, higher contrast version of ET capable of visualising specimens as thick as a eukaryotic cell [6]. Together this armoury of methods allows, for instance, the visualization of viruses and/or replication factories in-situ within the cell (see below and Chap. 14).

7.2.2 Small-Angle X-ray Scattering

SAXS is a technique that provides a structural characterization of a protein, protein complex or virus by using the elastic X-ray scattering from a solution containing the system of interest. The mathematical principles of small angle scattering were laid down in 1931 by André Guinier (Fig. 7.1) [for a review see [7]]. As used for the analysis of virus structure it is usual to work in solution phase and to use an X-ray (SAXS, wavelength usually ~1.5-Å) or neutron beam (SANS) onto a homogenous solution containing the biological specimen and to record the one-dimensional plot of scattered beam as a function of scattering angle (Q in the terminology of this method) close to the undeflected beam leaving the sample solution (hence “small-angle”). This one-dimensional scattering profile is the average of all scattering profiles deriving from the individual molecules (Fig. 7.2). Thus, in comparison to EM, where individual slices of the virus particle transform are assembled to produce as complete as possible version of the 3D transform, in small angle scattering the virus particle transform is spherically averaged to produce the linear scattering profile. Despite this massive loss of data, it is possible to extract information that relates to the volume, radius of gyration and the shape of the biological object. Modern synchrotron radiation sources are capable of producing highly acute curves which can support precise interpretation. Due to the development of sophisticated software which uses the massive computational power available today, it is, remarkably, possible to gain *ab-initio* low-resolution structural information by the use of the randomly distributed dummy atom modelling approach that has replaced the original trial-end-error shape assignment [8]. In brief, a large number of differently shaped molecules are generated and the scattering patterns simulated by analytical calculation are correlated and compared with the experimentally measured profile. Additional and/or external knowledge on the sample, for example disulfide-bonds presence, oligomerization state etc., will put constraints on the generation of the different models thereby increasing the reliability of the final shape assigned by the *ab-initio* procedure. Nonetheless, it is not possible to reconstruct a detailed 3D structure of a biological macromolecule from solely SAXS data due to the limited dimensionality of the raw data recorded (Fig. 7.2).

SAXS has been used in the characterization of the shape and composition of viruses since the early days of structural virology. Leonard and co-workers in 1953 carried out SAXS studies on SBMV, TBSV and TNV, in the same year the structure of the DNA was deduced from X-ray diffraction images (Fig. 7.1). At that point no virus structure was solved at high-resolution, but TBSV and TMV had already been crystallized and showed striking diffraction patterns. It is worth mentioning along this line that though not strictly SAXS, wide-angle diffraction and the interpretation of X-ray scattering from a partially ordered specimen, often referred to fibre diffraction have been pivotal in the elucidation not only of the DNA structure but also of virus structures such as TMV or phage Pf1 (for an exhaustive review on fiber diffraction studies of filamentous viruses see [9]). The use of such ordered fibres for structure determination has declined and nowadays is a technique almost unused.

In the early days turnip yellow mosaic virus, cucumber mosaic virus and bromegrass mosaic virus (BMV) were also investigated by SAXS. Interestingly, in the BMV study SAXS data were used alongside negative-stain images of the virus obtained by EM to reveal the virus dimensions and the presence of a hollow structure in the centre of the particles. The small set of plant viruses already mentioned were the initial common targets for the different structural techniques. A deviation from this theme was the structural study of a marine bacteriophage, PM2, the first phage to have been described possessing an internal lipid-bilayer [10]. This painstaking work combined both SAXS and negative-stain EM methods to describe in detail the architectural organization and assembly of proteins and lipids within the phage. Finally, fourty years later the *quasi*-atomic model of the entire PM2 virion was revealed by X-ray crystallography [11]. In retrospect the value of the method is evident – where specific questions can be asked, such as the radius or particle shell thickness and it is possible to construct a unique model, the method can give highly accurate answers and distinguish between competing models. This is exemplified by the results from neutron diffraction.

As with SAXS, SANS has provided insightful information of virus structures exploiting the differential contrast generated by the different interaction of neutrons with the nuclei of water (^1H_2O) and deuterated water (^2H_2O) to spatially map the genome and lipids from the protein mass [7]. Indeed, the advantages of this systematic contrast matching lies in the fact that since the intensity pattern measured (see Fig. 7.2) is the spherically averaged Patterson function defined as the Fourier transform of the structure factors describing the particle embedded in the solvent (for more details on Patterson function see general references on X-ray crystallography provided in Chap. 4), whole sets of Patterson vectors are eliminated by matching out certain components – disturbing markedly the diffraction.

In recent years SAXS has been re-emerging as an important structural tool to study biological macromolecules thanks to synchrotron radiation but also to automation in sample preparation and new high-capabilities of the DECTRIS-Pilatus detectors (see for example some of the European initiatives at http://www.saxier.org/ and at http://www.biostruct-x.eu/). However, from the virus point of view major insights using SAXS should be expected in the study of kinetic and time-resolved processes of virus assembly and disassembly which have started to appear [12].

7.2.3 X-ray Crystallography

In contrast to TEM and SAXS, virus X-ray crystallography (Chap. 4) relies on viruses ordering themselves in a 3D lattice, a crystal (Fig. 7.2). Illuminating the crystal by X-rays then allows one to record the Fourier transform of the unique portion of the lattice (unit cell) and hence the transform of the virus particle. The transform can only be sampled at those points specified by the Bragg condition, however this still allows a fully detailed reconstruction to be produced if all of the unique Bragg spots are measured and phases can be determined (see Chap. 4). Radiation damage is ameliorated since it is spread across all of the particles in the illuminated portion of the X-ray beam – usually in excess of a billion particles. All viruses crystallized so far possess icosahedral symmetry [icosahedron: a platonic solid with 20 identical equilateral triangular faces, 30 edges and 12 vertices; see Chap. 2 for details]. Due to this regular isometric shape, crystallization is often not the bottleneck in a virus crystallographic project and this is particularly true for relatively small and biochemically stable viruses. Indeed, the stability of virus crystals decreases with the increase of the virus size and it is approximately proportional to the inverse of the square of the virus radius [13]. Limitations in the amount or heterogeneity of the purified virus are the most restrictive factors in a successful project. However, recent advances in nano-crystallization techniques and efficient data collection strategies have allowed the determination of virus structures using only approximately 30 μg of virus as in the case of equine rhinitis A virus [14]. Once the crystals are available and show useful X-ray diffraction patterns, data recording and processing, phase determination and refinement with subsequent model building and refinement lead to the final atomic virus structure (for details of the entire protocol see Chap. 4 and [13]).

Thus, X-ray crystallography allows the determination of the virus structure at the atomic level and has contributed and still contributes to a very large extent to the understanding of virus biology. Milestones in structural virology (Fig. 7.1) have been the TBSV structure, the first virus solved at 2.9-Å resolution by Harrison's group in 1978 [1] and the structure of SMBV solved at 2.8-Å resolution by Rossmann's group in 1980 [2]. Five years later the structure of two viruses infecting humans, poliovirus and the common cold virus were solved by X-ray crystallography [15, 16] (see Chap. 4). Since then more virus structures have been elucidated by this technique as witnessed by the number of entries in the Protein Data Bank (PDB). The maximum number of virus structures was deposited in 2000. After this, the rate has decreased. In spite of this trend, in 2004, a significant contribution in the field came from the study of bacteriophage PRD1. PRD1 is the first and the only virus with an internal lipid-bilayer so far solved by X-ray crystallography in atomic detail, revealing a general principle of virus assembly scalable to very large viruses [17, 18].

Finally, in 2010, after several years of parallel advances, both X-ray crystallography and electron microscopy stunningly converged in terms of resolution achieved (~3.5-Å) with the elucidation of the structure of adenovirus (~900-Å diameter), which is so far the largest structure to be successfully tackled by X-ray crystallography [19, 20]. Meanwhile the methods for the determination of protein structure by X-ray crystallography have become much more routine, in part thanks to efforts to systematise all aspects of the process by structural genomics/proteomics activities [21]. Methodological improvements have also impacted whole virus studies for instance data collection directly from crystallisation plates should make structure determination more routine and safer [22, 23] and continued developments at synchrotrons, including microbeams [24] and pixel array detectors indicate the viability of the method. An area currently unexplored is the use of nanocrystals at highly powerful pulsed X-ray sources (the so-called X-ray free electron lasers). Recent work suggests that structure analysis might be possible from much smaller arrays of viruses than has been possible with more conventional synchrotron sources (see Sect. 7.6.1 and Chap. 4). Different public databases [see *e.g.* PDB http://www.rcsb.org/pdb/home/home.do, http://www.ebi.ac.uk/pdbe/emdb/ and http://viperdb.scripps.edu/] bring together all structural information available on studied viruses.

7.3 Combining X-ray Crystallography and Electron Microscopy

7.3.1 General Approach: The Beginning of Quasi-Atomic Models

Since early times structural virologists have tried to merge information and techniques to provide reliable insights into virus structures. Prior to 1978 (when X-ray virus crystallography first revealed the atomic details of a complete virus [1]), SAXS, EM and X-ray diffraction were usually combined to cross-validate the low resolution information gathered by each individual technique. Increasingly powerful electron microscopes and the definitive establishment of protein and virus X-ray crystallography as a high-resolution technique have made it possible to combine structural information at different resolution levels yielding *quasi*-atomic virus models capable of predicting molecular interactions within the virus or between virus particles and ligands, including virus-host cell recognition processes and virus-antibody interactions.

EM and cryo-EM allow visualization and 3D reconstruction of whole viruses and virus-ligand complexes that cannot be studied by X-ray crystallography. At the same time X-ray crystallography can frequently provide atomic snapshots of the composite pieces of the virus and, as in a 3D jigsaw, structural virologists collate the different pieces to produce the most accurate and precise representation of the virus and its interaction with other (macro)molecules of the outside world. This process defines the procedure of fitting X-ray protein structures, either viral proteins or ligands of the virus into density maps of viruses and virus-ligand complexes obtained by EM and cryo-EM. Despite the fundamental differences in the way electrons and X-rays interact with matter, the end result of 3D analyses in each is a ‘density’ map which looks surprisingly similar between the two methods and can therefore be interpreted and analysed in a similar way, using similar software. It is important to mention that NMR spectroscopy, although it cannot tackle a whole virus structure, has provided structures of viral proteins that have been subsequently used for fitting into EM maps (see Chap. 5). NMR is also increasing its role as hybrid method providing constraints on possible dynamic events or conformational changes of virus structural components (see Chap. 5).

After an initial manual docking of the entire viral and/or ligand protein structure into the corresponding volume of the EM reconstruction aided by graphic computer programs, improvement and refinement of the fitting carried out in either real space or reciprocal space, can be performed using a correlation coefficient-criterion and/or a least-square minimization [25]. As a result, this rigid-body refinement protocol aims to determine a total of six parameters, a rotation matrix and a translation vector to optimize the fitting to the virus map (Fig. 7.3a). It goes without saying that for a successful fitting the scale of the EM derived density map has to match that of the atomic model.

It has been estimated that this procedure can reach an accuracy of 2.2-Å in a virus map of 22-Å resolution and of 4-Å in a 24-Å map [27], the higher the resolution of the map the better the docking. Programs available for this task include the General Averaging Program (GAP; D.I. Stuart and J.M. Grimes, unpublished software available upon request from authors), EMfit (http://bilbo.bio.purdue.edu/~viruswww/Rossmann_home/softwares/emfit.php), Situs (http://situs.biomachina.org/index.html) and UROX (http://sites.google.com/site/xsiebert/urox). Recently, the concept of flexible fitting has also been introduced that allows distinct domains of the fitted structure to undergo conformational changes to match the volume restrictions imposed by the low-resolution EM map, cf. Norma (http://www.igs.cnrs-mrs.fr/elnemo/NORMA/) or Flex-EM (http://salilab.org/Flex-EM/). One of the critical points in flexible fitting is how to define what is fixed and what is flexible. This can take advantage either of prior knowledge or of normal mode analysis (*e.g*. Norma) that mathematically assigns directions of motion to structural elements of the structure to be fitted. Either approach to the definition of the rigid bodies has its own strengths and weaknesses, for example a user might tend to be over conservative and not reflect the dynamic properties of the structure shown by the EM reconstruction whilst with normal modes the assignment of too many motions can lead to over interpretation.

However, if the resolution of the virus map reaches 7-Å then it becomes possible to identify the fold of the most ordered part of the virus, cf. of major capsid proteins, and to model it using templates for β-barrel or α-helical motifs [28]. It is also possible to adopt template substitution methods, such as those provided by the Rosetta software [29] for modelling into low resolution X-ray and EM derived density maps (4–7 Å).

Nevertheless a major problem when combining different levels of information and specifically in generating (virus) quasi-atomic model is how to finely balance prior knowledge relative to the resolution achieved. In Fig. 7.4 we show the number of effective data points *vs* parameters for the different methods allowing the reader

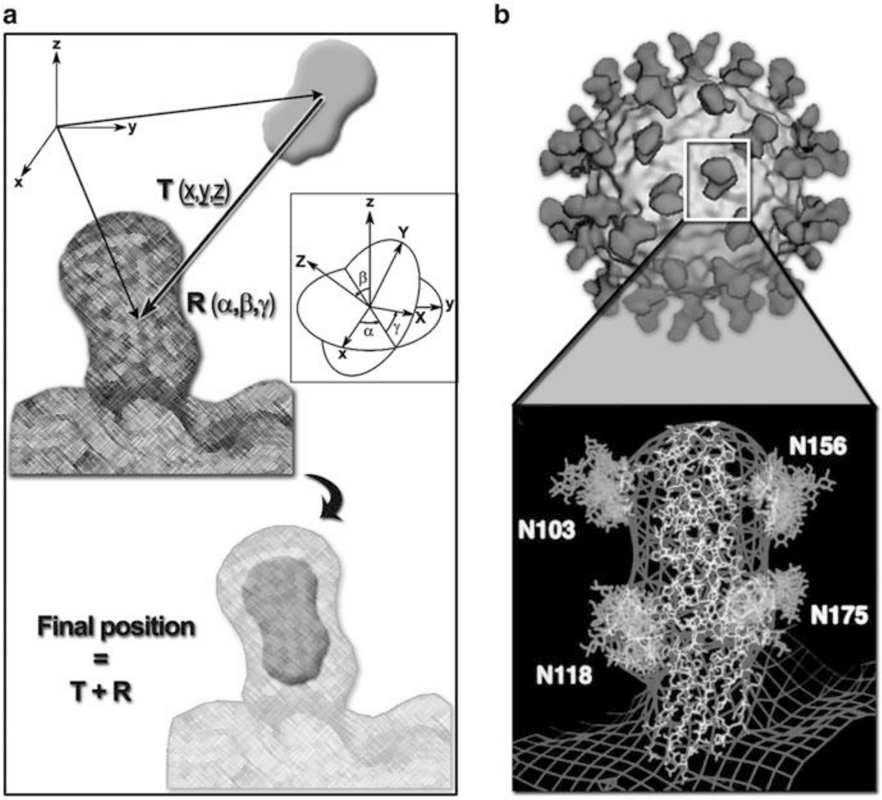


Fig. 7.3 Fitting X-ray data into EM maps. (**a**) Schematic of the fitting procedure; the crystal structure of a virus-binding molecule (ligand, represented here as a *grey* two-lobe object) is moved coarsely from its initial position into the corresponding electron density map of the virus-ligand complex obtained by EM (*dark-grey mesh*) by a translation (*T*) and a rotation (*R*); the final position of the object within the density is achieved by the refinement of six parameters [$\underline{x}$, $\underline{y}$, $\underline{z}$ and α, β, γ (Euler angles; inset)] using either real- or reciprocal-space refinement procedures (see [25]). (**b**) *Top*, EM reconstruction of rhinovirus 16 (*light-grey*) with CD155 cellular receptor molecules bound (*dark-grey*); *bottom*, the CD155 X-ray structure (*white* stick model) fitted into the virus-receptor EM density (*dark-grey* mesh); the N-glycosidic groups in the CD155 structure are labelled according to the asparagine residue to which each is bound (Adapted from [26]. With permission)

to grasp what is achievable by each technique and by their combination. In principle we need an objective way of assessing the information content of each of the sources of data and producing a model that satisfies the data to the level of accuracy of the data, but otherwise assumes nothing. We return to discuss this challenge below.

As mentioned earlier, one of the first examples of structural methods combination can be found in the study of adenovirus by Roger Burnett and his laboratory in 1986 [3]. Adenovirus is a large dsDNA virus that infects vertebrates, including

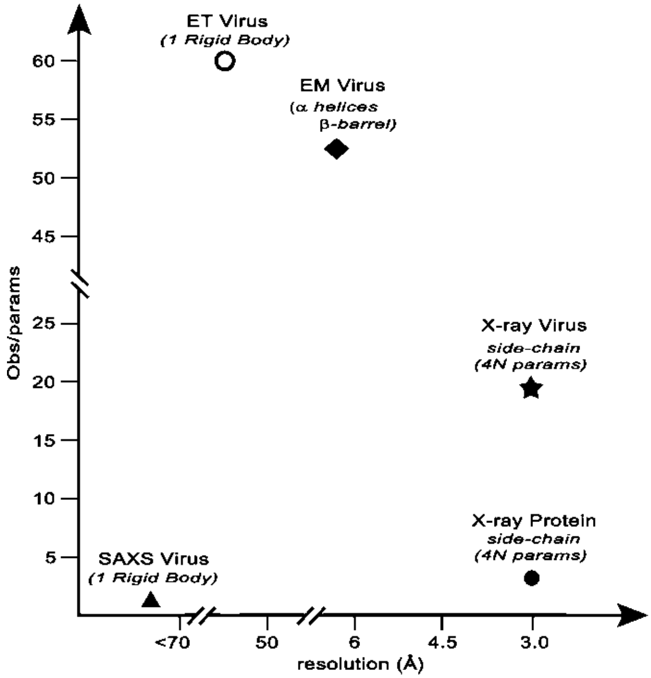


Fig. 7.4 Simplified ratio *per* technique. The data points were obtained by estimating the number of observations for a virus of a determined size and achievable resolution divided the number of structural elements (parameters) to be fitted. For the determination of the virus structure by X-ray crystallography virus we considered the lowest non-crystallographic symmetry scenario in a virus crystal

humans. As passionately stated by Roger Burnett (http://www.cicnetwork.es/secciones/interviews/11_burnett.html) his interest in adenovirus stemmed from beautiful electron microscopy images collected in the ‘70s by Nick Wrigley of hexamers of the major capsid protein, named hexons that recapitulated architectural elements of the entire virion. However, X-ray structural studies on the hexon structure had already started [30] prefiguring the approach that finally led to the interpretation of protein-protein interactions in the assembling of the large icosahedral adenovirus particle. By solving the X-ray structure of the hexon at 2.9-Å resolution and by placing it into the adenovirus EM reconstruction, Roberts and co-workers [3] opened the way to the hybrid methods approach in structural virology. This has been followed by many other studies for example of the binding of the human cellular receptor ICAM-1 to human rhinovirus 16 (HRV-16) [26], or neutralizing antibody bound to foot-mouth-disease virus (FMDV) [31].

More recently *quasi*-atomic models have been generated for enveloped and membrane-containing viruses. In the former case, a lipid-bilayer decorated by glycoproteins surrounds a nucleocapsid whilst in the latter case the vesicle is engulfed by the proteinaceous capsid shell. One of the best structurally described enveloped virus is human pathogenic dengue virus (*Flaviviridae* family). Years of investigation on the entire virion by EM and of its glycoproteins by X-ray crystallography have finally produced a clear picture of the dengue virus maturation, assembly pathway and cell-entry mechanism [32, 33]. Also, by combining the crystal structure of an antigen-binding antibody fragment (Fab) in complex with West Nile virus (WNV) E glycoprotein and the dengue and WNV structures in complex with the Fab by EM, Cherrier and co-workers [34] have been able to establish the structural basis for antibody neutralization of infectious immature particles. Likewise *quasi*-atomic models for other biomedical relevant enveloped viruses have been recently generated with a similar 'fitting X-ray structures into cryo-EM maps' approach, for instance chikungunya virus (*Alphavirus* genus, *Togaviridae* family) and vesicular stomatitis virus (*Rhabdoviridae* family).

In the case of lipid-containing viruses, the original example in hybrid methods is PRD1, a bacteriophage that infects gram-negative bacteria [35] and whose *quasi*-atomic model was proposed in 2001 [36]. PRD1 is indeed the model system for this type of viruses having been intensively studied biochemically, genetically and structurally. Its architecture, major capsid protein fold and principle of assembly strikingly recapitulate those elucidated years earlier for adenovirus [3, 17, 18, 37] leading to the concept of viral lineages and virus common ancestry [38] (see below, Sect. 7.8). Another illustrative example, although not strictly a case of fitting X-ray structures into cryo-EM map, is presented by the process that led to the X-ray structure solution of the marine lipid-containing bacteriophage PM2. The moderate 7-Å resolution achieved with the PM2 virus crystals (in comparison to the ~4-Å resolution obtained for PRD1 virus) prompted the determination of the structures of the individual components at higher resolution [11]. This "divide et impera" approach led to high resolution structures of the two major viral capsid components that were then docked into the low-resolution X-ray map. As a result, key aspects of viral evolution, capsid assembly and membrane morphogenesis were elucidated [11].

Furthermore, it is worth underlining that the protocol of fitting of atomic structures into electron density maps can be also extended to volumes obtained by sub-tomogram averaging (see below and Chap. 3; sub-tomogram averaging algorithms [39–41] allow averaging equivalent structures thus improving the signal-to-noise ratio). This can lead to plausible fitting of X-ray crystal structures as in the case of the capsid protein domain of the Gag polyprotein in HIV [42].

Finally, the combination of methods in structural virology has pioneered an equivalent approach nowadays emerging in biology and defined as 'integrative modelling'. Indeed, this approach aims to combine different available sources of information – structural, biochemical, biophysical and proteomics – to obtain a molecular picture or a unified structural view of the system in study. Nice examples of this are embodied in the 'Integrative Modelling Platform' (IMP) (http://salilab.org/imp/) and Haddock (http://www.nmr.chem.uu.nl/haddock2.1/), with a recent example being the analysis of the nuclear pore [43]. How and with what weights the different sources of information are combined remain to be fully and routinely implemented in a suitable statistical framework (Bayesian, Maximum-Likelihood, etc.) that would provide a validation and statistical significance assessment of the resultant model. Although this work is ongoing and is still very crude in many cases, the vision is already set, thanks to the combination of methods used to study virus structures.

7.3.2 Validating the Quasi-Atomic Models

Validation of the proposed *quasi*-atomic model is fundamental to the reliable interpretation of molecular interactions and their usefulness. Although several programs (see Sect. 7.3.1) that carry out positional refinement of the initial docked structure into low resolution maps provide correlation coefficients (for real space analysis), or metrics equivalent to the crystallographic R-factor [44] (for reciprocal space analysis), the inclusion of such values in the final publications is still left to the discretion of the authors. It must be said however that there is no unique criteria that describes the accuracy of the model and a multi-parameter assessment (atomic clashes, stereochemistry, biological meaningfulness) together with the statistical measures is the best solution at present. Furthermore, if additional biochemical or biophysical information is available, for instance a proved interaction between specific residues of the ligand and the virus, then it should be taken into account for example by imposing restraints on the movements of the fitted X-ray structure during refinement.

In 2010, the EM community attempted to standardize practices during the inaugural meeting of the Electron Microscopy Validation Task Force (http://vtf.emdatabank.org/). A series of guidelines and recommendations were suggested with the ultimate goal of increasing the impact and the significance of EM in biology and biomedicine. This reflected the earlier successful efforts of the X-ray community through the X-ray Validation Task Force (http://www.wwpdb.org/workshop/2011/index.html), where the broad principles are now accepted and required for publication. Given the limitations on the validation criteria it is advisable when combining X-ray with EM, to keep a conservative attitude in the biological interpretation of the resulting *quasi*-atomic model. Also, remember that the data-to-parameter ratio generally available during this hybridization process is unfavourable (see Fig. 7.4) – better safe than sorry!

7.3.3 Phase Information "Interplay"

A major problem in crystallography, as previously stated, is the so-called phase problem, a methodological 'inconvenience' that prevents the retrieval of atomic positions directly from the raw data (see Chap. 4). EM images embody the phases of the object, by visualizing it in real space and thus are a potential source of phase information for diffraction methods. When highly symmetrical particles such as icosahedral viruses are investigated structurally, it is possible to exploit the redundancy of information for phase refinement (see Chaps. 3 and 4). In crystallography, in the best-case scenario when none of the symmetry axes of the virus align with the crystallographic symmetry axes of the crystal, the full 60-fold non-crystallographic symmetry (NCS) can be iteratively used to improve the starting phase assignment [in the worst-case scenario only 5-fold NCS]. It has been shown that the availability of 60-fold NCS can allow convergence of phase refinement starting from phases derived from a spherical object [45, 46] (for an overview on *ab-initio* phasing see [47]). However, for this protocol to succeed it is paramount to know (*i*) the exact location of the icosahedral symmetry axis relative to the asymmetric unit (AU) of the unit cell, (*ii*) the radius of the starting model [45, 46] and (*iii*) the protein region for the generation of appropriate envelopes used in the averaging and solvent flattening procedures.

In virus crystallography, apart from the first virus structures that were solved using heavy atom methods [1, 2], it has been and still is a common practice to use as a starting phasing model a virus previously solved [unrelated or even a sphere (see above)] or, if available, a low-resolution derived EM map located within the asymmetric unit of the unit cell of the crystal through the molecular replacement (MR) technique [13, 48] (see Chap. 4). It is worth noting that for the phasing of FMDV [49], a hybrid virus model was built from existing viruses weighting-up the conserved structural pieces and used as a searching model in MR, an approach that is now implemented in some of the standard protein MR protocols.

Advances in EM (see Chap. 3) have propelled the application of this phasing strategy to X-ray data obtained from crystals of biological macromolecules smaller than viruses and often less symmetrical [50–53]. Although there are exciting perspectives of this alternative approach for circumventing the crystallographic phase problem, further requirements need to be fulfilled for its successful application: resolution range overlap between the 3D electron microscopy data and those obtained by X-ray crystallography, *i.e.* it is important to collect the low-resolution terms, and the presence of non-crystallographic symmetry. The former condition allows the object to be located within the asymmetric unit of the unit cell (by MR), while the latter boosts the phase refinement and phase extension procedures. To be more precise the power of phase extension depends on the degree of oversampling of the transform of the unique portion of the object and non-crystallographic symmetry delivers this directly, however a very high solvent content does so as well (for instance a thin virus shell with disordered interior).

One of the first examples in which the use of low-resolution models was successful in phasing X-ray data of a small protein (38 kDa) was inspired by studies on Mengo and FMDV viruses [50]. Villeret and co-workers were not only able to build the correct geometry of the oligomeric state (12-fold) using a monomeric X-ray structure of a related protein and generate low-resolution phases at 8-Å but they were also able to phase-extend those initial phases to the full 3-Å resolution of the X-ray data (see also ‘Further reading’).

In principle also SAXS could provide initial low-resolution phasing as it has been described in the case study of nitrite reductase, a 34.4 kDa protein that forms homotrimers. However, in this case once the SAXS envelope was correctly located in the asymmetric unit of the unit cell, only phases to 20-Å were reliable and no further phase extension was attempted [54].

A recent enlightening case of phase interplay between EM and X-ray crystallography comes from the structure of the marine bacteriophage PM2 at 7-Å resolution [11] (Fig. 7.5). The initial phases for solving the entire virus X-ray crystal were derived from the cryo-EM based PM2 structure at 8.4-Å resolution [55]. In spite of the apparently minimal ~1.4-Å improvement in resolution of the X-ray over the EM data, the quality of the resulting X-ray map thanks to the accuracy of the crystallographically measured structure factor amplitudes, allowed a clearer interpretation of the different, 60-fold symmetrical arranged structural components, in particular those pivotal elements located between the capsid and the lipid vesicle [11]. However, the 7-Å X-ray map could not provide the unequivocal determination of the fold of the major capsid protein P2 (32 kDa). So, an X-ray crystallography study of isolated P2 was undertaken [57]. The failure of MR using supposedly related X-ray structures as searching models, and the difficulty of getting heavy-atom soaked crystals or suitable selenomethionine (SeMet)-derivatized X-ray data (see Sect. 7.4) required the use of the low-resolution electron density corresponding to a P2-trimer extracted from the X-ray averaged map of PM2 as searching model. The location of the two P2-trimers within the asymmetric unit of the cell by molecular replacement led to the determination of the initial phases at 7.6-Å resolution then extended to 2.5-Å by means of non-crystallographic symmetry and solvent flattening [56].

In conclusion, although the examples mentioned above do not yet represent routine procedures, the experience gained in virus phasing provides an additional structure determination tool for protein crystallography.

7.4 Dissecting Virus Structures by Combining Biochemical, Genetic and Biophysical Tools

Virus structures can be tackled by complementing not only the structural techniques mentioned above but also methods that involve the biochemical and genetic manipulation of the virion. It is clear that the prerequisite for high-quality structural data is the quality of the original biological material. The traditional ways to purify viruses are based on their sedimentation behaviours (rate zonal centrifugation) or their particle densities (equilibrium centrifugation). These can then be combined with differential centrifugation, concentration methods (*e.g*. filter devices) or precipitation (*e.g*. using different polyethylene glycols). The addition of biochemical and genetic modifications makes it possible to produce and to study different kinds of virus material from wild type and mutant virus infections, perform quantitative virus particle dissociations experiments, or use recombinant protein expression

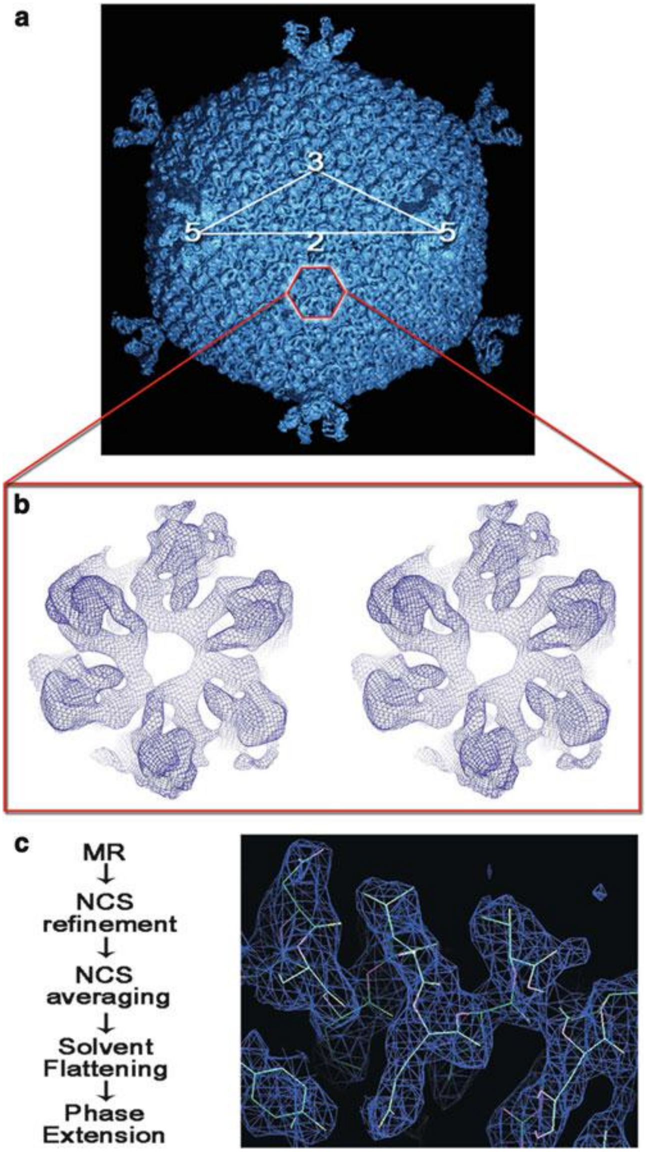


Fig. 7.5 Phasing with low-resolution phases. (**a**) X-ray-derived map of the lipid-containing bacteriophage PM2 [11] (~600-Å in diameter), which was phased using the corresponding EM map [55]; icosahedral symmetry axes in the capsid are labelled; a trimeric capsomer of the major capsid protein P2 is outlined in *red*. (**b**) Enlarged stereo-view of the extracted electron density at 7-Å corresponding to the P2 trimer used in the MR protocol. (**c**) *Left*, workflow for phasing X-ray data using the low-resolution density map in (**b**) as a searching model; *right*, experimental electron density map at 2.5-Å for the P2 major capsid protein resulting from MR, followed by phase refinement, extension and solvent flattening (Reproduced from [56]. With permission)

systems. In the following sections we will illustrate some examples where the tailored use of biochemical, genetic and biophysical techniques has provided insightful views into virus structure.

7.4.1 Classical Antibody Labeling Methods

Immuno-EM can be used as a powerful tool to localize specific proteins also asymmetrically distributed in the virion. The method relies on the specific recognition, by primary antibody, of a target protein (antigen) exposed on the virion surface. A secondary molecule recognizing the primary antibody is then added, for instance, gold-conjugated protein A, which binds with high affinity to rabbit-IgG molecules. The gold particles produce high contrast under TEM (see Sect. 7.2.1) allowing the localization of specific proteins in the virus structure, but detailed structural information is lost. Using secondary binders conjugated to gold particles of different sizes, different antigens can be co-localized within the same sample. Prior to the labelling, the virus particles are usually fixed with cross linkers, glutaraldehyde being one of the most popular choices. Recently the toolbox of cross-linkers and chemical probing protocols available has been expanded [58] allowing the exploration of novel avenues in virus biochemical manipulation. Furthermore, virus particles can be loosened *e.g.* by mild treatment with non-ionic detergent such as Triton X-100 during the sample-grid preparation to increase the accessibility of the antigens to the primary antibodies; see for example [59].

Typical problems of immune-EM, such as unspecific labeling or background can be overcome by having appropriate controls and adjusting the amount of primary antibody. The most reliable conclusions, however, can be made when immunogold labeling EM data is combined with other approaches, as demonstrated for the dsDNA bacteriophage PRD1. Combination of immunolabeling [59] with the work done with PRD1 mutants [60] verified the presence of the unique vertex structure used for genome packaging.

7.4.2 Mutagenesis Studies

Classical mutagenesis using shotgun mutagens such as ultraviolet light (UV) or N-methyl-N-nitro-N- nitrosoguanidine (NTG) and, especially, site-directed mutagenesis are powerful techniques to create virus mutants. To identify proteins and investigate their functions specific virus mutants can be created either by generating a stop codon inside the target gene (nonsense mutants) or by deleting the entire protein coding region (deletion mutants). If the mutations are lethal, rescue systems are needed in a form of suppression or complementation, which limit the usage of virus mutants to those viruses infecting a host for which a genetic system has been established. Genetic engineering of obscure micro-organisms isolated from the environment is demanding [61]. Bacteriophage PRD1 infects gram–negative bacteria such as *Escherichia coli* and *Salmonella enterica* enabling the use of well-established genetic tools. PRD1 is one example for which numerous suppressor-sensitive mutants have been obtained and have proved invaluable in assigning functions to the corresponding viral proteins. Furthermore, the genetic removal of the PRD1 flexible receptor-binding proteins from the virion vertices was crucial in obtaining diffracting crystals leading to the determination of the X-ray virus structure at 4-Å resolution [17, 18].

Moreover mutagenesis has also allowed the insertion of green-fluorescence protein (GFP) into large and complex viruses such as herpes or vaccinia viruses – providing an important tool for studying cell interactions or virus localization within the cell (see Chap. 14). We expect that careful selection of multiple tags will also provide structural information, *e.g.* from fluorescence resonance energy transfer (FRET) and correlation spectroscopy methods. An early example comes from work on influenza virus polymerase revealing interactions in the cell [62].

Fine structure-function relationships in virus particles at the amino acid residue level can be carried out by individually replacing specific residues based on high-resolution structural information. Classically residues are replaced by alanine, removing the side chain (beyond the Cβ atom) and the intra- or intermolecular interactions it establishes with other residues, without introducing additional groups and non-native interactions, and with the minimum probability of altering the main-chain conformation of the viral proteins. For example, systematic alanine-scanning mutational analyses on the structural and functional relevance of protein-protein and protein-nucleic acid interactions in a virus particle have been carried out for FMDV and the parvovirus minute virus of mice (see Chaps. 1 and 7). Also, resolution of the crystal structure of some designed single-amino acid mutants is providing high-resolution insights into the structural bases of physical properties, assembly, conformational dynamics, and disassembly of virus particles, for example in rhinovirus and poliovirus.

7.4.3 Quantitative Dissociation Studies

The controlled removal of specific proteins from virus particles is very valuable for structural and structure-function studies (Fig. 7.6a). Using this biochemical quantitative dissociation approach several viral proteins have been assigned to their right positions in the virion. The interactions between the capsid proteins and the underlying membrane component can be disrupted by *e.g.* urea or guanidine hydrochloride treatments, as exemplified by studies on bacteriophages PM2 and PRD1, respectively, which revealed the virus membrane associated proteins and released the major capsid proteins [64, 65]. The released proteins can be further purified by standard biochemical methods for protein crystallography [37, 57, 65]. Quantitative protease treatments of the virus particles might identify structures

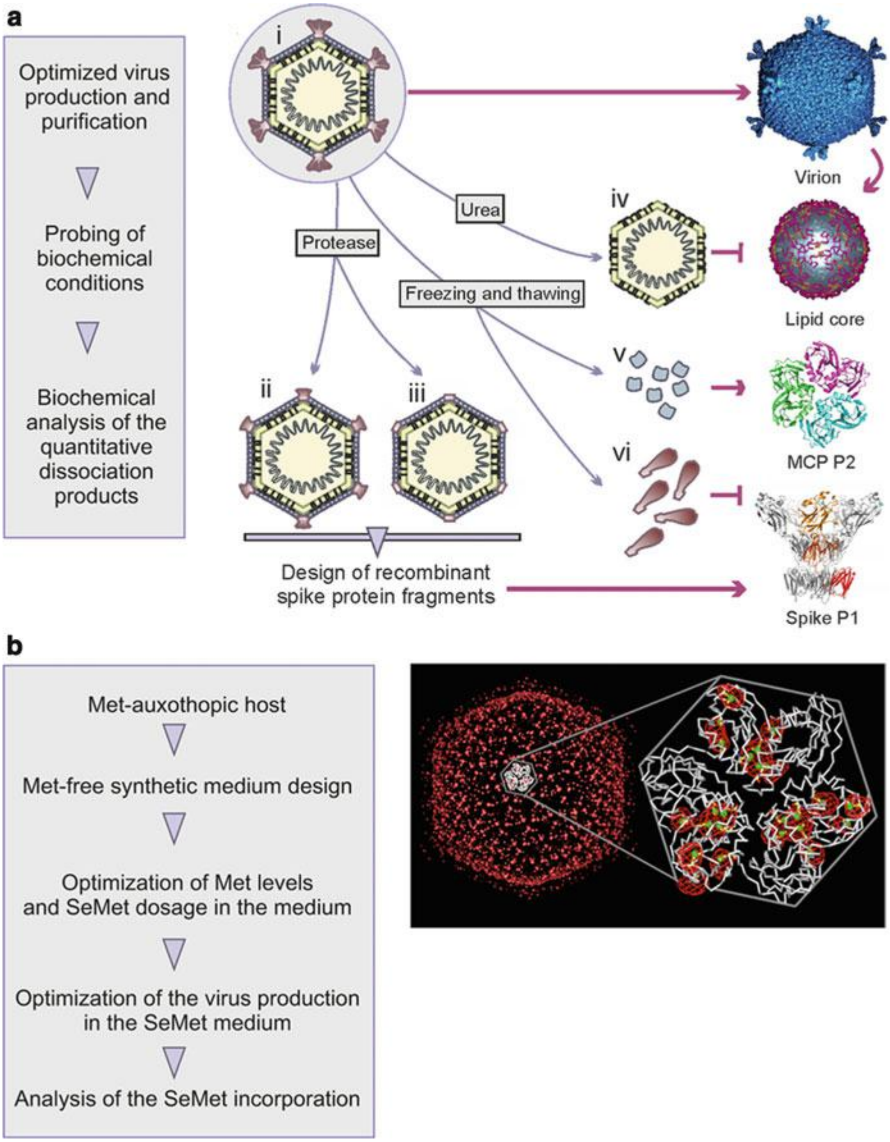


Fig. 7.6 Schematic workflow for biochemical probing. (**a**) *Left*, workflow in a study on the quantitative dissociation of virus particles. *Right*, schematic representation of the controlled dissociation of bacteriophage PM2 [63] used to solve the structures of virion components. (i) Highly purified and infectious virions used to solve the virion structure by cryo-EM [55] and X-ray crystallography [11]; (ii) proteinase K-treated virus particles devoid of the distal receptor binding domain of the spike protein P1; (iii) bromelain-treated virus particles devoid of the two outermost domains of the spike protein P1; (iv) urea treatment expose the internal lipid core particle (membrane proteins, lipids and the genome); (v) and (vi) freezing and thawing of the virions release (v) the trimeric major capsid protein (MCP), and (vi) the monomeric spike proteins.

protruding from the virion surface [55, 65]. For the analysis of membrane-containing viruses, membrane extraction with different detergent treatments can be used to selectively remove different membrane-associated proteins from the virus particle. In case of the internal-membrane containing PRD1 virus integral and peripheral membrane proteins have been separated from one another [66].

One example of a quantitative dissociation analysis is marine icosahedral dsDNA bacteriophage PM2 (Fig. 7.6b). For PM2, several unique subviral particles have been distinguished by their sedimentation properties in a sucrose gradient [55, 65]. A combination of biochemical, cryo-electron microscopic [55] and X-ray crystallographic approaches [11, 56, 57] has visualized the entire PM2 virion structure at subnanometric resolution and led to the determination of its major structural protein structures (Fig. 7.6b). Also archaeal virus SH1 [67], archaeal pleomorphic viruses [68] and *Thermus* bacteriophage P23-77 are examples of viruses for which quantitative dissociation analysis has provided useful information and significantly helped in the structural analysis [69].

7.4.4 Selenomethionine Incorporation as a Labelling Tool

Efficient selenomethionine labelling of protein species within bacteria is nowadays a standard technique but less so for the labelling of proteins expressed in mammalian cells [70]. Even more challenging is the selenomethionine labelling within viruses that when combined with X-ray crystallography is a valuable method in the structural virology. To date, only two efficient methods for selenomethione labeling of complex viruses have been reported [63, 71] and their roles in the interpretation of the electron density maps of PRD1 and PM2 were significant [11, 17, 63]. The PM2 labeling method relied on methionine auxotrophic host (Figs. 7.6c and d; [63]), whereas successful PRD1 selenemothionine incorporation was obtained with methionine prototrophic host grown in minimal medium and pulsed with selenomethione at a late stage of infection [71]. In addition to obtaining phase information in X-ray crystallography [72], selenomethionine labeling combined with difference Fourier maps can facilitate the interpretation of protein folds and even the identification of the protein species [11]. Using selenomethionated and native virus X-ray data it was possible to assign electron density to all ordered protein components in the virion based on their amino acid sequence [11, 63].

Fig. 7.6 (continued) *Far right, top to bottom*: X-ray crystallographic structure of a virion and its lipid core particle, MCP P2 trimer and the pentameric vertex-spike complex of protein P1 (Adapted from [11]. With permission) (**b**) *Left*, workflow of the selenomethionine (SeMet) labeling of bacteriophage PM2. *Right*, SeMet difference Fourier map for the entire SeMet-labeled PM2 virion contoured at 4σ (*red* mesh) with a single trimer of the MCP P2 outlined in *white*; an enlarged view of the P2 Cα backbone is shown, with the methionine residues depicted as balls-and-sticks (the incorporated selenium atoms are shown as *green spheres*) (Reproduced from [63]. With permission)

7.4.5 *Some Old and New Biophysical Techniques to Study Viruses*

Together with the above techniques other biophysical methods for studying virus particles in solution, including fluorescence spectroscopy, circular dichroism (CD), hydrogen exchange or limited proteolysis followed by MS (see Chap. 6) and differential scanning calorimetry (DSC) and isothermal scanning calorimetry (ITC) may be used to study different aspects of virus particle stability and conformational dynamics, and relate these properties to virus structure. DSC and ITC respectively are useful in assessing the thermal stability and conformational transitions of virus particles and in the study of virus/cell binding, entry processes and genome release. Several studies have been carried out using DSC and ITC on viruses and phages, cf. phage λ [73]. Changes in secondary structure of the capsid proteins in viral particles have been followed by far-UV CD; intrinsic fluorescence of viral particles or extrinsic fluorescence by binding dyes to the viral nucleic acid exposed on capsid disassembly have been used to follow capsid conformational rearrangements and determine the thermal or chemical stability of viral particles against dissociation (see Chap. 6).

Recently an application of the thermofluor assay used to assess protein stability before crystallization has been adapted for high-throughput screening of thermal stability of viruses [74]. This new application uses two distinct fluorescent dyes simultaneously, one with affinity to hydrophobic residues and the other to the genome. Upon heating, hydrophobic residues becoming exposed and genome release are simultaneously monitored by the fluorescence intensity and virus stability rapidly assessed.

Finally, recent developments in MS techniques have shown that it is possible to investigate very large protein complexes such as viruses by native MS (see Chap. 6). This formidable advance [75] opens the way to the analysis of virus disassembly processes traditionally considered out-of-range for this method.

7.5 Combining Electron Microscopy and Electron Tomography

7.5.1 *General Approach*

The understanding of the architectural and molecular organization of enveloped and complex viruses is a current challenge. Traditional approaches such as EM and more stringently X-ray crystallography require sample homogeneity in terms of structural organization and size, a premise that fails when studying those viruses that do not possess a regular shape or composition, the so-called pleomorphic viruses. Indeed such viruses often come in several sizes and shapes and the lack, until recent years, of a structural technique that could tackle such viruses had left them at the margin of structural investigation.

The advent of ET and cryo-ET [76] (Chap. 3) has allowed these difficult viruses to be tackled [77]. With this technique (for details and some examples see Chap. 3) viruses can be studied as individual particles whilst those structural elements conserved and redundant within the virus structure can be averaged, thus improving the signal-to-noise ratio and overcoming typical problems inherent to the technique such as anisotropy and incompleteness of data due to the missing wedge [76].

Thus, if a virus possesses distinct structural levels some with no symmetry (*e.g.* envelope) and others with symmetry (*e.g.* nucleocapsid) it is in principle possible to combine the cryo-ET 3D reconstruction with the other symmetry-imposed and/or averaged structural elements. These latter elements are usually at higher resolution and can be either obtained by sub-tomogram averaging within the same tomographic data collection (see for example [39, 41] and websites http://www.dynamo-em.org; http://www.biochem.mpg.de/en/rg/foerster/Content_Software/PyTom/index.html; http://www.opic.ox.ac.uk/mediawiki/index.php/Jsubtomo) or by other structural methods such as EM or X-ray crystallography. This combination will picture the architectural complexity of the virus at different resolution levels. ET is also very powerful in studying viruses in the cellular context, in particular during viral entry and replication (see also Chaps. 3 and 14). In the following section we will give a few examples of the combination of ET and EM techniques and we will illustrate the smart use of tomography in assessing particle symmetry [78].

7.5.2 Case Studies

One of the most elegant examples of combination of EM techniques in picturing a complex virus is the cryo-ET study of herpes simplex virus (HSV) [79]. In this work the regular icosahedral nucleocapsid was separated from the enclosing pleomorphic envelope in turn decorated by the tegument and glycoproteins, both indispensable for infection. Nucleocapsids were extracted from the tomograms, 60-fold averaged (as in standard icosahedral 3D reconstruction approach) and replaced in the whole virus 3D reconstructed tomogram (Fig. 7.7a), thus revealing contacts between the nucleocapsid and the tegument not seen in individual tomographic reconstructions.

Combination of ET and EM is not limited to viruses with an intrinsic symmetry mismatch. Several situations during the infection cycle induce a symmetry loss even in highly symmetric viruses. Picornaviruses are small, icosahedral, non-enveloped viruses that have been studied for several viral processes such as receptor attachment or RNA translocation. The use of decorated liposomes with poliovirus receptor CD155 allowed the imaging of the receptor bound to poliovirus both by single-particle reconstruction and electron cryo-tomography [82].

For retroviruses, cryo-ET has provided structural information of the human and simian immunodeficiency viruses (HIV/SIV) envelope spike proteins gp41/gp120 (Env) responsible for binding to the cellular receptor. The gp120/gp41 complex is randomly distributed onto the virion surface making it difficult to study [83, 84].

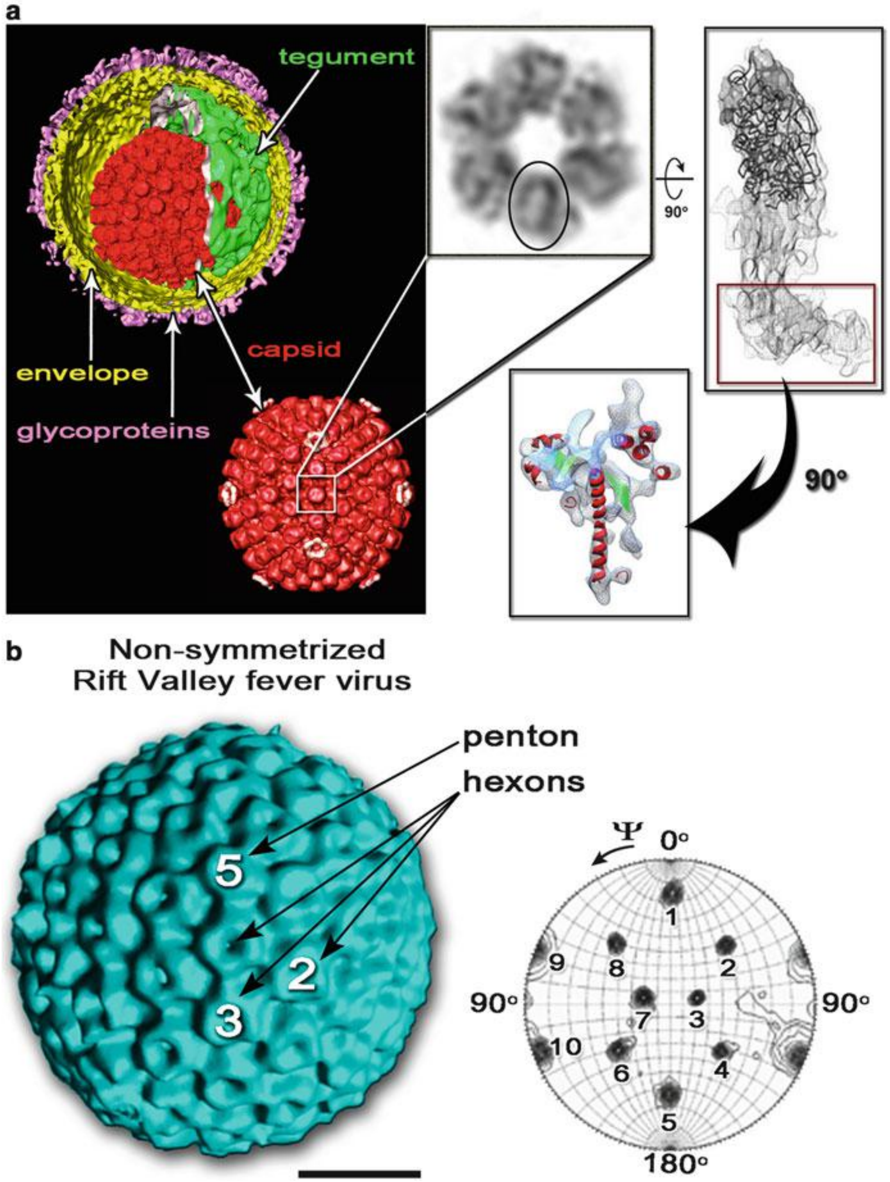


Fig. 7.7 Using X-ray tools in electron tomography. (**a**) *Top left*, herpes virus electron cryo-tomography reconstruction with the different structural elements [79]; below, herpes virus nucleocapsid reconstruction obtained by sub-tomogram averaging combined with icosahedral averaging; outlined one capsomer sitting on a 2-fold icosahedral axis (Courtesy of K. Grünewald, University of Oxford); *top center* cryo-EM derived density corresponding to this capsomer which is composed by six copies of VP5 protein; *top right*, a VP5 monomer composed of three domains (upper, middle and floor domain); the crystal structure of the upper domain is shown as a Cα ribbon model (*dark-grey*) fitted into the corresponding cryo-EM density; no X-ray structure exists for the

Subtomogram averaging of thousands of individual spikes and the imposition of the threefold symmetry gave a 3D map of sufficient quality to dock in the structure of SIV-gp120 core previously solved by X-ray crystallography [85]. Additional studies with bound receptor and antibody have elucidated key aspects of the spatial arrangement during cell and antibody recognition and conformational tuning [86]. Moreover beneath the HIV/SIV envelope, during maturation and after cleavage of the Gag polyprotein, the capsid domain forms the cone-shaped core shell organized as a hexagonal lattice. Tomographic images of this lattice followed by subtomogram averaging of the hexagonal building block allowed individual protein domains to be identified and fitted with the corresponding X-ray crystal structure [42].

Cryo-ET has also provided a 3D picture of influenza virus that displays an evident pleomorphy [87]. The two different surface proteins could be distinguished directly from the tomograms and molecular shape comparison with the crystal structures of hemagglutinin (HA) or neuraminidase (NA) allowed mapping of the individual spikes in the tomograms.

Preliminary analysis of negatively-stained samples of several bunyaviruses showed a large diversity in size, morphogenesis, structure and surface. One member of this virus family is Rift Valley fever virus (RVFV), a major pathogen of animals and humans in Africa. Cryo-ET provided the first useful 3D reconstruction of the RVFV [78], achieved by reconstructing 46 individual particles with similar overall size and shape and then averaging them together. The averaged map was Fourier transformed and the derived structure factor amplitudes used to calculate a self-rotation function [a routine procedure used in X-ray crystallography to detect the redundant presence of structural elements within the asymmetric unit of the crystal (see Chap. 4), Fig. 7.7b]. This identified five-, three- and twofold axis and allowed the unambiguous assignment of icosahedral symmetry to this virus. In turn, imposition of the 60-fold icosahedral symmetry further improved the resolution of the 3D reconstruction from 75-Å to 61-Å, revealing the capsomers arrangement and the symmetry of the lattice (triangulation number *pseudo*-T = 12). This laid the basis for structural studies at higher resolution that was achieved by the systematic single-particle sub-tomogram averaging of 267 particles, leading to a virus map at 20-Å resolution. This dramatic improvement allowed the analysis of the structure of the heterodimer formed by the G_N and G_C glycoproteins which decorate the envelope [88].

Fig. 7.7 (continued) middle domain; the floor domain is delineated by a *red rectangle* (Figure adapted from [80]. With permission). *Bottom right*, cryo-EM density corresponding to the floor domain; secondary structure templates have been fitted to suggest structural homology with the gp5 capsid protein of bacteriophage HK97 (Figure adapted from [81]. With permission). (**b**) *Left*, electron density reconstruction of RVFV by single particle electron tomography [78]; capsomers and icosahedral 2-, 3- and 5-fold symmetry axes are marked; *right*, stereographic projection of the Kappa = 120° section for the self-rotation function of the tomographic reconstruction, revealing the presence of 10 3-fold axes, and thus demonstrating the icosahedral nature of the RVFV; scale bar represents 25 nm (original images courtesy of SJ Watowich, University of Texas Medical Branch)

7.6 From Viruses in Solution to Virus in Cells

7.6.1 Virus Structures in the Cellular Context

Viruses have been traditionally observed within infected cells using either fluorescent labelled viruses and light-microscopy (resolution worse than 200-nm) or using cell-sectioning techniques for subsequent imaging in an electron microscope [4] (see Chap. 14). In this latter case different cell and sample preparation techniques have been put in place and optimized to preserve close-to-native conditions of the virus and cell. These efforts have been coupled to advances that have made the acquisition of tomographic data routine. Tomography thus allows adding depth information at higher resolution than the serial addition of ultrathin 2D sections (for details and examples see Chaps. 3 and 14).

One of the most recent additions to the available tools to investigate virus structures has been provided by synchrotron facilities through the generation of soft X-rays in the water window (wavelength 2.3–4.4 nm) that due the differential absorption of oxygen and carbon atoms (ten times lower for O than C) provide contrast for cellular tomography imaging. Virus X-ray tomography has seen pioneering applications using as model systems the virus life cycle of vaccinia virus [89] and herpesvirus [6]. These studies demonstrate that soft X-rays can provide valuable 3D structural insights into virus maturation events at the sub-nanometer level (50–30 nm resolution) for reasonably thick (5–8 μm) cells and compartments, including the cell nucleus.

Interestingly, virus crystallography is also aiming towards *in vitro* structural analysis. Indeed crystals of large viruses are extremely fragile and there are attempts to grow intracellular crystals for structure determination – a grand challenge for crystallography (see also Chap. 4). Inspirational of this new crystallogenesis strategy has been the observation that not only does protein crystallization occur *in vivo* [90, 91] but also paracrystalline arrays of virus particles often form within infected cells [4, 92]. However further developments in synchrotron radiation technology in particular in delivering high-brilliance synchrotron light, are needed to make X-ray data collection possible from these *in vivo* crystals and hopes are raised by the preliminary tests of the X-ray free electron laser (XFEL) on viruses [93, 94], that although technically successful provided very low resolution information (<200-Å) due to the limited number of incident photons in each short pulse of X-rays.

7.6.2 Methods for Cellular Landscape Scanning

One of the major difficulties in the structural interpretation of virus-infected cells is the confusion deriving from macromolecular crowding. In order to be able to provide an atlas of the different steps of a viral infection, following how the viruses enter the cell, replicate and assemble, and finally leave the cell, different timeline points are structurally investigated during infection (see Sect. 7.7.1). For each timepoint tomograms are collected and combined computational techniques, from segmentation to template matching, are employed to detect recurrent cellular or virus structures [95, 96]. Sub-tomogram averaging of equivalent structures increases the signal-to-noise ratio and allows the improved structure to be inserted back into the cellular landscape [97, 98]. This approach can also use virus structures obtained at higher resolution by other methods such as cryo-EM or X-ray crystallography, thus providing more accurate information about the geometrical arrangement and spacing between virus particles and cellular structures. This hybrid approach fills-in information inevitably missing from grid-based ET methods (see Chaps. 3 and 14) and bridges the resolution gap between 3D cellular tomography and other methods such as single particle EM and X-ray crystallography [92].

7.7 Emerging Hybrid Methods

7.7.1 Correlative Microscopy

The analysis of biological functions from whole organisms to cells and subsequently to single molecules has been a dream for structural scientists for decades. Structural biology and virology has recently become increasingly interested in dynamic cellular and virus-cell interaction processes This requires a time-resolved approach to the visualization at the different size levels and the correlation in time and space between the different methods remains a challenge.

Correlative microscopy allows the integration of data across the microns to nanometers range, using fluorescence light microscopy and electron microscopy [6, 99, 100] (see Chap. 14). However, from the point of view of combined structural strategies it is important to underline that the aim is to identify cellular regions of interests by the use of specific chromophore labels in the light microscope and track biological events before stopping the clock and imaging the same region at much higher-resolution using an electron microscope. In this way virus-cell interactions are reconstructed in 3D and framed in time. A critical step is finding the same cellular region once the cell is within the electron microscope. To this end a common reference system between the light- and electron-microscopes must exist. Technological advances in robotics and microscopy cryo-holders have made possible data collection and a general work-flow for the procedure, pioneered by Plitzko [99] (Fig. 7.8). While correlative microscopy for 3D tomographic reconstruction is still in its infancy the potential benefits are enormous in studying virus-infected cells. Recent important contributions come, for example from the study of retroviruses where it has been possible to identify HIV particles attached to cell membranes [101] and to show the release of mutant HIV cores into the cytoplasm of host cells [102].

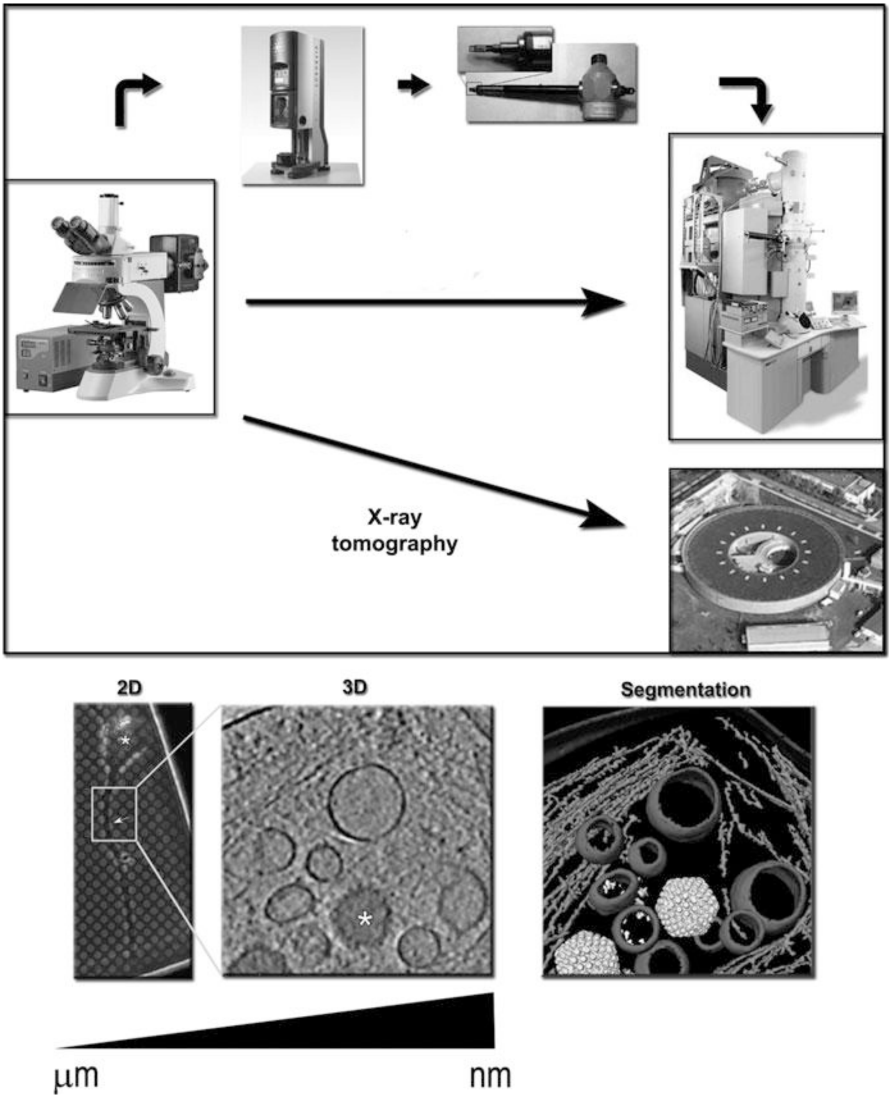


Fig. 7.8 Workflow in correlative microscopy. *Top*, Schematic of the methodological steps involved in performing correlative light and electron microscopy (CLEM). Cellular structures are sequentially identified with different microscopic techniques with increasing resolving power. Targeted processes in cells are imaged using a fluorescence microscope (*centre left*); then, the cells supported on a grid, are vitrified using a plunging freezing device (*top left*) and the grid transferred using a suitable cryo-holder (*top right*) in a field-emission-gun transmission electron microscope (*centre right*) for visualization of sub-cellular features by electron tomography; nowadays also X-ray tomography on entire cells can be performed at synchrotron facilities. *Bottom*, *left to right*: detection of the area of interest in 2D (using light microscopy; micrometer scale); visualization and 3D reconstruction by ET techniques (nanometer scale); segmentation of the different structural elements

The advent of correlative microscopy will benefit the study of those viruses that are characterized by low replication and infection efficiencies and have therefore escaped direct visualization. A clear example of this is hepatitis C virus, whose low-level replication *in vitro* has hampered the visualization of viral morphological processes by traditional ET technique.

7.7.2 Developing 3D Imaging Techniques

One of the most recent (and contentious) method developments in 3D image reconstruction algorithms that could, in principle, be used for initial low-resolution phasing is represented by ankylography [103]. Briefly, it consists in collecting on a curved detector the diffraction pattern generated by the scattering of a coherent soft X-ray laser hitting the object of interest (not necessarily crystalline). The proposed advantage of ankylography over XFEL crystallography or X-ray tomography is that, unlike these techniques, it uses a 2D spherical diffraction pattern to reconstruct the full 3D structure of the object. A proof-of-principle of this reconstruction method was the determination of poliovirus at 20–30 Å resolution [103]. So, while it might appear that ankylography might have an important impact, there is a fundamental problem in that capturing a thin spherical shell of data in reciprocal space is ultimately a very poor approximation to measuring the full 3D volume needed to produce a high resolution reconstruction [104].

As a variant on coherent X-ray scattering, ptychography, an imaging technique that was proposed in the late ‘60s by Hoppe [105] and aiming to solve the phase-problem by using an interference phenomenon has been recently re-emerging as an EM technique that could provide potentially superresolution (~5-Å) [106]. The principles behind ptychography have been also employed in X-ray tomography in life science experiments [107]. It will be interesting to see the ultimate scope of this method, however in its present form it is slow (even at a synchrotron) and requires a high radiation dose to achieve a 3D reconstruction at modest resolution.

7.8 The Biology Behind the Combined Methods

7.8.1 Implications for Drug-Design and Vaccine Development

Current methods for drug-discovery rely mainly on: (i) ligand-based design and (ii) structure-based design. In the first case the search for new compounds against a target molecule is based on previously known drugs; in the second case the search is based on the availability of the 3D structure of the target molecule (see Chap. 20 for a comprehensive view on the translational applications of structural virology

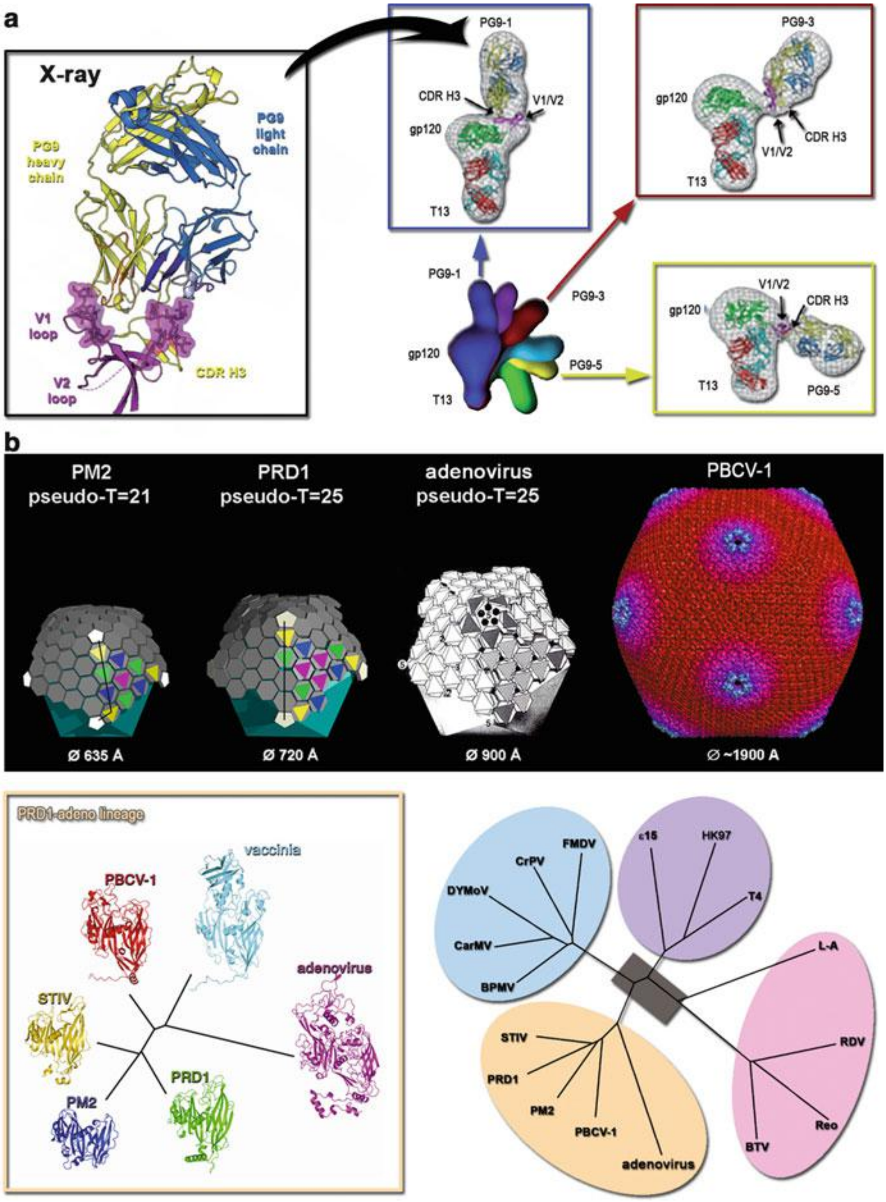


Fig. 7.9 Impact of the combination of structural methods. (**a**) *Left*, X-ray structure of the HIV-1 gp120 V1/V2 domain (in *magenta*) complexed with neutralizing antibody PG9 (*yellow*, heavy chain; *blue*, light-chain) represented as ribbons; *right*, negative stain EM reconstruction classes (six in total, colour-coded from *blue* to *green*) of ternary complexes composed by gp120, PG9 and CD4-binding-site antibody T13 with the corresponding fitting of the gp120 V1/V2-PG9 crystal structure in three of them (insets) (Adapted from [109]. With permission). (**b**) *Top left*, schematic representation of the architecture of bacteriophages PM2 and PRD1; *grey* hexagons represent the pseudo-hexameric morphology displayed by viral major capsid proteins forming trimers; the trimers composing the icosahedral asymmetric unit are represented as *colour-coded triangles*;

in antiviral research). Both screening strategies are computationally intensive. Structure-based design uses atomic or quasi atomic models to rationalize the interaction of the potential candidate drug with the target biomolecule, for example to block the binding of the virus to the cellular receptor. One of the first examples where the interdisciplinary approach of virus EM with X-ray crystallography has provided clues for the development of antiviral drugs is represented by the docking of ICAM I to human rhinovirus [26]. Indeed this quasi-atomic model (Fig. 7.3b) contributed to the development of pleconaril, a drug that inhibited the uncoating of rhinovirus upon infection by binding to the major capsid protein VP1 [108]. Although this drug was never licensed, failing in phase-clinic III studies [108], the whole process reinforced the utility of the combined approach in virus structure determination.

Undoubtedly, the interaction of viruses with their cellular receptors is one of the major structural targets in antiviral research facilitating the rational design of novel drugs aiming to halt cell entry. Similarly, studies aimed at understanding the mechanism of viral evasion of antibody neutralization, for example for HIV-1 gp120 [109], have made use of crystal structures and negative stain electron microscopy (Fig. 7.9a). Nevertheless there are other steps during virus morphogenesis that can be tackled by combining EM information and X-ray crystallography, for example targeting the ribonucleoprotein complex of influenza virus responsible for viral RNA synthesis [111]. Furthermore the structural elucidation of virus-like-particles (VLPs) as transporters of foreign antigenic epitopes (see Chap. 21) by combining cryo-EM and crystal structures of the component proteins allow the localization and investigation of the modifications introduced in the virus capsid structure. The knowledge of the conformation of the inserted epitope in the VLPs and their structural relationship with the rest of the particles directs possible improvements of construct design; examples include TBSV and Norovirus which have been used as a platform for polyvalent display of antigenic epitopes and vaccine design. Indeed structural virology has a huge potential for the structure-based development of next generation vaccines.

7.8.2 *Novel Concepts in Virus Evolution*

In the last decade the accumulation of virus and viral protein structures determined mainly by X-ray crystallography and cryo-EM or the combination of these two

Fig. 7.9 (continued) black-lines on the capsomers indicates the T numbers). *Top right*, adenovirus (Adapted from [3]), and the cryo-EM derived density of Paramecium bursaria chlorella virus-1 (PBCV-1) (image taken from http://viperdb.scripps.edu/info_page.php?VDB=1m4x). All four represented viruses are members of the PRD1-adeno viral lineage. *Bottom left*, structure-based phylogenetic tree of viral capsid protein members of the PRD1-adeno viral lineage (Reproduced from [110]. With permission). *Bottom right*, the four viral lineages so far grouped using the phylogenetic structural approach; each lineage includes some virus/phage representatives (PRD1-adeno viral lineage, *ocre*; BTV-like lineage, *pink*; Picorna-like lineage, *light-blue*; HK97-lineage, *violet*) (Reproduced with permission from Abrescia et al. 2010)

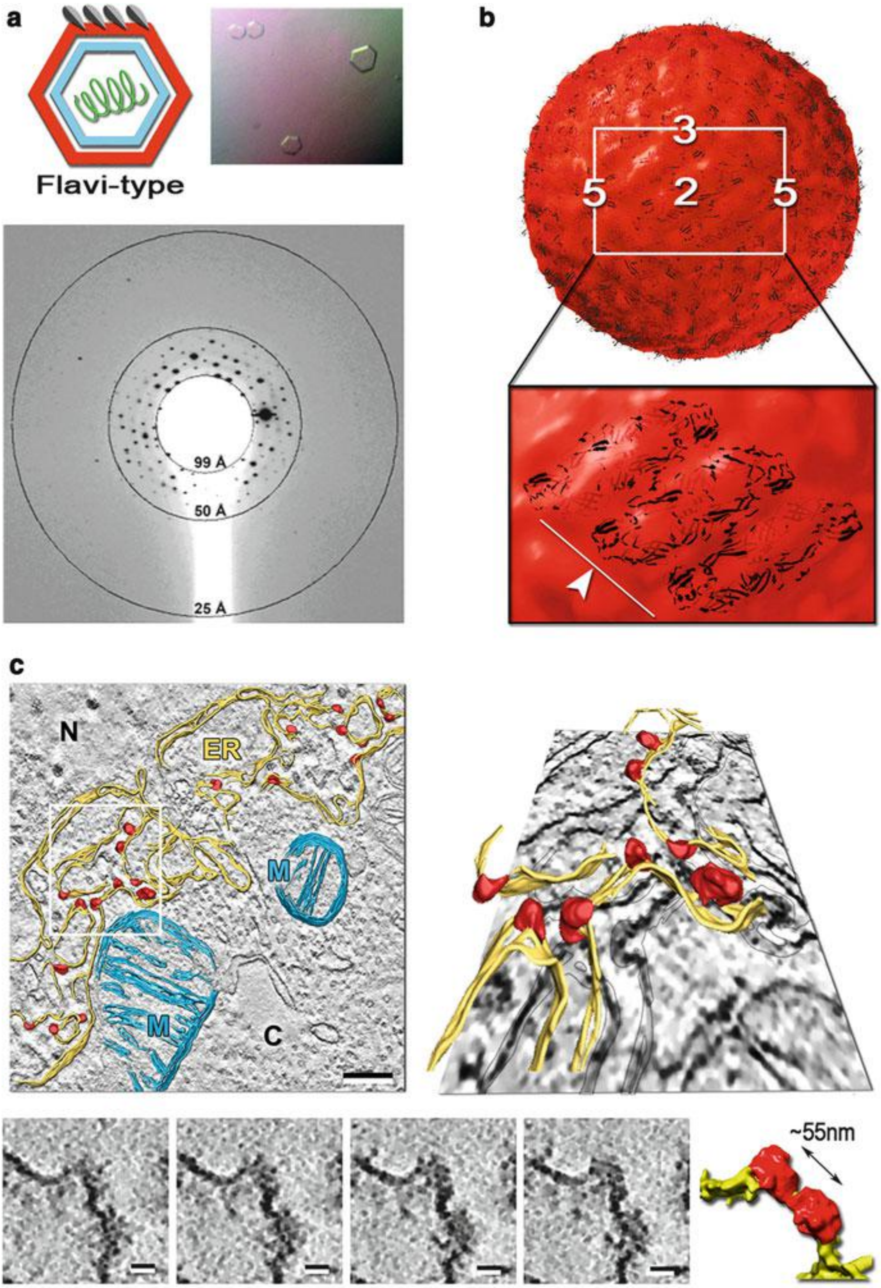


Fig. 7.10 Some of the challenges ahead. (**a**) *Top left*, scheme of a flavivirus; the envelope is shown in *red*, the glycoproteins in *grey* or *light-grey*, the capsid in *cyan* and the genomic RNA in *green*. *Top right*, one of the different morphologies of WNV crystals. *Bottom*, diffraction pattern at ~25-Å resolution corresponding to the crystals shown on top (Adapted from [113]. With permission). (**b**) *Top*, three-dimensional cryo-EM reconstruction of HCV-like-particles (HCV-LPs)

techniques has allowed us to look at the virus world (virosphere) on the basis of their structural relationships. This has led to the emergence of a novel concept in virus evolution that uses specific structural markers, the so-called virus-self elements (for example the fold of the capsid protein) to infer common ancestry [38]. Although the observation that viruses infecting different species (cf. SBMV vs. rhinovirus) displayed a similar structural fold in their capsid protein (single jelly-roll) goes back to 1985 [16], it was only in 2002 [38] that the capsid-centered view in virus classification was formalized as a precise hypothesis and through the years validated by the systematic experimental approach of structure comparisons [112]. This novel taxonomic approach allows us to simplify the organization of the virosphere and pinpoint relationships between viruses that infect different hosts across the three domains of life. A clear example of this is represented by adenovirus and PRD1, which infect vertebrate and bacteria, respectively. Their capsid proteins share a double jelly-roll fold and the capsids share common architectural principles [17] (Fig. 7.9b). The power of this structural classification is apparent in the fact that almost half of the virus families currently described, many of which do not belong to any order according to the International Committee on Taxonomy of Viruses (ICTV) classification scheme (see Chap. 1), can be traced back to one of the four viral lineages so far described: (*i*) PRD1/adenovirus-like, (*ii*) Blue-tongue-virus-like, (*iii*) HK97-like and (*iv*) Picorna-like [112] (Fig. 7.9b). It is unclear how many more virus lineages will be unveiled by the structure-based classification method since this depends on the discovery of new viral protein folds. Finally, this structure-based taxonomy of viruses opens the way to conceptualize anti-viral strategies that bridge viruses that were considered unrelated until now.

7.8.3 Perspectives and Challenges Ahead

Much progress has been made in the understanding virus biology thanks to the combination of biochemical, genetic and structural techniques. EM and X-ray crystallography have played (and still do) a dominant role in describing and

Fig. 7.10 (continued) (~500-Å in diameter) at 30-Å resolution; the icosahedral symmetry axes are indicated; *Bottom*, close-up of a HCV-LP capsid region; dengue virus glycoproteins (*black* cartoon models, indicated with a *white arrow-head* and *line*) have been fitted within the HCV-LP cryo-EM density (Adapted from [115]. With permission). (**c**) *Top*, electron tomographic reconstruction of Sf9 cells showing the cellular ER landscape when expressing hepatitis C virus-like-particles (HCV-LPs), and segmentation of its major structural elements. *Top left*, slice through the center of a denoised tomogram with superimposed rendering of the convoluted ER membranes in *gold*, budding HCV-LPs in *red* and mitochondria in *slate-blue*. Scale-bar = 200 nm. *Top right*, perspective view of the region outlined by a white rectangle. N = nucleus, C = cytoplasm, M = mitochondria, ER = endoplasmic reticulum. *Bottom*, four consecutive tomographic slices showing coalescence of HCV-LPs buds together with their isosurface representation and segmentation on the right. Scale-bar = 20 nm (Reproduced from [116]. With permission) providing high-resolution viral structures, whilst their hybridization has provided insights into the mechanisms of virus-host recognition, antibody neutralization, antigen display, virus assembly and maturation and other properties and functions of virus particles (see Part III of this book, Chaps. 10, 11, 12, 13, 14, 15, 16, 17, 18 and 19). Nevertheless, efforts in the past have focused mainly on virus particles that display icosahedral morphology (about half the virus families known), which are easier to analyse, as described above.

Structural virology has still several challenges ahead, including the structural characterization at atomic resolution of enveloped viruses, such as members of the *Flaviviridae*, *Alphaviridae* and *Bunyaviridae* families, and even more complex viruses, such as the poxviruses.

Attempts to crystallize WNV, a member of the *Flaviviridae* family, have been carried out [113] and indeed crystals were obtained and diffracted to 25-Å resolution – not competitive with the 14-Å resolution achieved by cryo-EM [114] (Fig. 7.10a). Another structural challenge is the biomedically very important hepatitis C virus (HCV) whose architecture, structure of its glycoproteins and assembly are not understood. A wealth of biochemical information has been gathered on the HCV life-cycle in past and recent years but only very recently, have efforts focused on the structural characterization of the virus [115–117] (Fig. 7.10b).

Technological developments with synchrotron radiation with the advent of XFEL and the *in-situ* diffraction [118], or the routine use in EM of the phase-plate contrast technology [119], the new direct detection cameras and correlative microscopy will undoubtedly shape our investigation on virus structures in the near future.

To sum up, the hybridization of approaches in structural virology is a powerful method in generating knowledge and tools serving in the fight against viral infections. These contributions are fundamental to the health of humans and animals, and will help to create a "Systems Virology" vision capturing the complex biology of viruses.

Acknowledgements DB acknowledges support from the CIC bioGUNE and HMO thanks the University of Helsinki for funding to EU ESFRI Instruct Centre for Virus Production and Purification. DIS is supported by the UK MRC. This work was enabled by the Spanish Ministerio de Ciencia y Innovacion (BFU2009-08123), Spanish Ministerio de Economia y Competitividad (BFU2012-33947) and the Basque Government (PI2010-20) grants to NGAA.

References and Further Reading

1. Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G (1978) Tomato bushy stunt virus at 2.9 Å resolution. Nature 276:368–373
2. Abad-Zapatero C, Abdel-Meguid SS, Johnson JE, Leslie AG, Rayment I, Rossmann MG, Suck D, Tsukihara T (1980) Structure of southern bean mosaic virus at 2.8Å resolution. Nature 286:33–39
3. Roberts MM, White JL, Grutter MG, Burnett RM (1986) Three-dimensional structure of the adenovirus major coat protein hexon. Science 232:1148–1151

4. Goldsmith CS, Miller SE (2009) Modern uses of electron microscopy for detection of viruses. Clin Microbiol Rev 22:552–563
5. Jiang W, Baker ML, Jakana J, Weigele PR, King J, Chiu W (2008) Backbone structure of the infectious epsilon15 virus capsid revealed by electron cryomicroscopy. Nature 451:1130–1134
6. Hagen C, Guttmann P, Klupp B, Werner S, Rehbein S, Mettenleiter TC, Schneider G, Grünewald K (2012) Correlative VIS-fluorescence and soft X-ray cryo-microscopy/tomography of adherent cells. J Struct Biol 177:193–201
7. Svergun DI, Koch MH (2003) Small-angle scattering studies of biological macromolecules in solution. Rep Prog Phys 66:1735–1782
8. Svergun DI, Koch MH (2002) Advances in structure analysis using small-angle scattering in solution. Curr Opin Struct Biol 12:654–660
9. Stubbs G (2001) Fibre diffraction studies of filamentous viruses. Rep Prog Phys 64:1389–1425
10. Harrison SC, Caspar DL, Camerini-Otero RD, Franklin RM (1971) Lipid and protein arrangement in bacteriophage PM2. Nat New Biol 229:197–201
11. Abrescia NG, Grimes JM, Kivela HM, Assenberg R, Sutton GC, Butcher SJ, Bamford JK, Bamford DH, Stuart DI (2008) Insights into virus evolution and membrane biogenesis from the structure of the marine lipid-containing bacteriophage PM2. Mol Cell 31:749–761
12. Canady MA, Tsuruta H, Johnson JE (2001) Analysis of rapid, large-scale protein quaternary structural changes: time-resolved X-ray solution scattering of Nudaurelia capensis omega virus (NomegaV) maturation. J Mol Biol 311:803–814
13. Fry EE, Abrescia NGA, Stuart DI (2007) Virus crystallography. In: Sanderson M, Skelly J (eds) Macromolecular crystallography: conventional and high-throughput methods – a practical approach. Oxford University Press, Oxford
14. Tuthill TJ, Harlos K, Walter TS, Knowles NJ, Groppelli E, Rowlands DJ, Stuart DI, Fry EE (2009) Equine rhinitis A virus and its low pH empty particle: clues towards an aphthovirus entry mechanism? PLoS Pathog 5:e1000620
15. Hogle JM, Chow M, Filman DJ (1985) Three-dimensional structure of poliovirus at 2.9 A resolution. Science 229:1358–1365
16. Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht H, Johnson JJ, Kamer G, Luo M, Mosser AG, Rueckert RR, Sherry B, Vriend G (1985) Structure of a human common cold virus and functional relationship to other picornaviruses. Nature 317:145–153
17. Abrescia NG, Cockburn JJ, Grimes JM, Sutton GC, Diprose JM, Butcher SJ, Fuller SD, San Martín C, Burnett RM, Stuart DI, Bamford DH, Bamford JK (2004) Insights into assembly from structural analysis of bacteriophage PRD1. Nature 432:68–74
18. Cockburn JJ, Abrescia NG, Grimes JM, Sutton GC, Diprose JM, Benevides JM, Thomas GJ Jr, Bamford JK, Bamford DH, Stuart DI (2004) Membrane structure and interactions with protein and DNA in bacteriophage PRD1. Nature 432:122–125
19. Liu H, Jin L, Koh SB, Atanasov I, Schein S, Wu L, Zhou ZH (2010) Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. Science 329:1038–1043
20. Reddy VS, Natchiar SK, Stewart PL, Nemerow GR (2010) Crystal structure of human adenovirus at 3.5 A resolution. Science 329:1071–1075
21. Perrakis A, Daenke S, Stuart DI, Sussman JL (2011) From SPINE to SPINE-2 complexes and beyond. J Struct Biol 175:105
22. Axford D, Owen RL, Aishima J, Foadi J, Morgan AW, Robinson JI, Nettleship JE, Owens RJ, Moraes I, Fry EE, Grimes JM, Harlos K, Kotecha A, Ren J, Sutton G, Walter TS, Stuart DI, Evans G (2012) In situ macromolecular crystallography using microbeams. Acta Crystallogr D Biol Crystallogr 68:592–600
23. Wang X, Peng W, Ren J, Hu Z, Xu J, Lou Z, Li X, Yin W, Shen X, Porta C, Walter TS, Evans G, Axford D, Owen R, Rowlands DJ, Wang J, Stuart DI, Fry EE, Rao Z (2012) A sensor-

adaptor mechanism for enterovirus uncoating from structures of EV71. Nat Struct Mol Biol 19:424–429

24. Evans G, Axford D, Owen RL (2011) The design of macromolecular crystallography diffraction experiments. Acta Crystallogr D Biol Crystallogr 67:261–270
25. Rossmann MG (2000) Fitting atomic models into electron-microscopy maps. Acta Crystallogr D Biol Crystallogr 56:1341–1349
26. Kolatkar PR, Bella J, Olson NH, Bator CM, Baker TS, Rossmann MG (1999) Structural studies of two rhinovirus serotypes complexed with fragments of their cellular receptor. EMBO J 18:6249–6259
27. Gilbert RJ, Grimes JM, Stuart DI (2003) Hybrid vigor: hybrid methods in viral structure determination. Adv Protein Chem 64:37–91
28. Topf M, Sali A (2005) Combining electron microscopy and comparative protein structure modeling. Curr Opin Struct Biol 15:578–585
29. Kaufmann KW, Lemmon GH, Deluca SL, Sheehan JH, Meiler J (2010) Practically useful: what the Rosetta protein modeling suite can do for you. Biochemistry 49:2987–2998
30. Franklin RE, Harrison SC, Pettersson U, Philipson L, Branden CJ, Werner P (1971) Structural studies on the adenovirus hexon. Cold Spring Harb Symp Quant Biol 36:503–510
31. Hewat EA, Verdaguer N, Fita I, Blakemore W, Brookes S, King A, Newman J, Domingo E, Mateu MG, Stuart DI (1997) Structure of the complex of an Fab fragment of a neutralizing antibody with foot-and-mouth disease virus: positioning of a highly mobile antigenic loop. EMBO J 16:1492–1500
32. Yu IM, Zhang W, Holdaway HA, Li L, Kostyuchenko VA, Chipman PR, Kuhn RJ, Rossmann MG, Chen J (2008) Structure of the immature dengue virus at low pH primes proteolytic maturation. Science 319:1834–1837
33. Li L, Lok SM, Yu IM, Zhang Y, Kuhn RJ, Chen J, Rossmann MG (2008) The flavivirus precursor membrane-envelope protein complex: structure and maturation. Science 319:1830–1834
34. Cherrier MV, Kaufmann B, Nybakken GE, Lok SM, Warren JT, Chen BR, Nelson CA, Kostyuchenko VA, Holdaway HA, Chipman PR, Kuhn RJ, Diamond MS, Rossmann MG, Fremont DH (2009) Structural basis for the preferential recognition of immature flaviviruses by a fusion-loop antibody. EMBO J 28:3269–3276
35. Bamford DH, Caldentey J, Bamford JK (1995) Bacteriophage PRD1: a broad host range dsDNA tectivirus with an internal membrane. Adv Virus Res 45:281–319
36. San Martin C, Burnett RM, de Haas F, Heinkel R, Rutten T, Fuller SD, Butcher SJ, Bamford DH (2001) Combined EM/X-Ray imaging yields a quasi-atomic model of the adenovirus-related bacteriophage PRD1 and shows key capsid and membrane interactions. Structure 9:917–930
37. Benson SD, Bamford JK, Bamford DH, Burnett RM (1999) Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. Cell 98:825–833
38. Bamford DH, Burnett RM, Stuart DI (2002) Evolution of viral structure. Theor Popul Biol 61:461–470
39. Nickell S, Forster F, Linaroudis A, Net WD, Beck F, Hegerl R, Baumeister W, Plitzko JM (2005) TOM software toolbox: acquisition and analysis for electron tomography. J Struct Biol 149:227–234
40. Scheres SH, Melero R, Valle M, Carazo JM (2009) Averaging of electron subtomograms and random conical tilt reconstructions through likelihood optimization. Structure 17:1563–1572
41. Castaño-Diez D, Kudryashev M, Arheit M, Stahlberg H (2012) Dynamo: a flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. J Struct Biol 178:139–151
42. Briggs JA, Riches JD, Glass B, Bartonova V, Zanetti G, Kräusslich HG (2009) Structure and assembly of immature HIV. Proc Natl Acad Sci U S A 106:11090–11095

43. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT, Sali A, Rout MP (2007) The molecular architecture of the nuclear pore complex. Nature 450:695–701
44. Brunger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355:472–475
45. Tsao J, Chapman MS, Rossmann MG (1992) Ab initio phase determination for viruses with high symmetry: a feasibility study. Acta Crystallogr A 48(Pt 3):293–301
46. Chapman MS, Tsao J, Rossmann MG (1992) Ab initio phase determination for spherical viruses: parameter determination for spherical-shell models. Acta Crystallogr A 48(Pt 3): 301–312
47. Rossmann MG (1995) Ab initio phase determination and phase extension using non-crystallographic symmetry. Curr Opin Struct Biol 5:650–655
48. Rossmann MG (1990) The molecular replacement method. Acta Crystallogr A 46(Pt 2):73–82
49. Acharya R, Fry E, Stuart D, Fox G, Rowlands D, Brown F (1989) The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution. Nature 337:709–716
50. Villeret V, Tricot C, Stalon V, Dideberg O (1995) Crystal structure of Pseudomonas aeruginosa catabolic ornithine transcarbamoylase at 3.0 Å resolution: a different oligomeric organization in the transcarbamoylase family. Proc Natl Acad Sci U S A 92:10762–10766
51. Dodson EJ (2001) Using electron-microscopy images as a model for molecular replacement. Acta Crystallogr D Biol Crystallogr 57:1405–1409
52. Navaza J (2008) Combining X-ray and electron-microscopy data to solve crystal structures. Acta Crystallogr D Biol Crystallogr 64:70–75
53. Trapani S, Schoehn G, Navaza J, Abergel C (2010) Macromolecular crystal data phased by negative-stained electron-microscopy reconstructions. Acta Crystallogr D Biol Crystallogr 66:514–521
54. Hao Q, Dodd FE, Grossmann JG, Hasnain SS (1999) Ab initio phasing using molecular envelope from solution X-ray scattering. Acta Crystallogr D Biol Crystallogr 55:243–246
55. Huiskonen JT, Kivelä HM, Bamford DH, Butcher SJ (2004) The PM2 virion has a novel organization with an internal membrane and pentameric receptor binding spikes. Nat Struct Mol Biol 11:850–856
56. Abrescia NG, Grimes JM, Oksanen HM, Bamford JK, Bamford DH, Stuart DI (2011) The use of low-resolution phasing followed by phase extension from 7.6 to 2.5 Å resolution with noncrystallographic symmetry to solve the structure of a bacteriophage capsid protein. Acta Crystallogr D Biol Crystallogr 67:228–232
57. Abrescia NG, Kivelä HM, Grimes JM, Bamford JK, Bamford DH, Stuart DI (2005) Preliminary crystallographic analysis of the major capsid protein P2 of the lipid-containing bacteriophage PM2. Acta Crystallogr Sect F Struct Biol Cryst Commun 61:762–765
58. Leitner A, Reischl R, Walzthoeni T, Herzog F, Bohn S, Förster F, Aebersold R (2012) Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. Mol Cell Proteomics 11(M111):014126
59. Gowen B, Bamford JK, Bamford DH, Fuller SD (2003) The tailless icosahedral membrane virus PRD1 localizes the proteins involved in genome packaging and injection at a unique vertex. J Virol 77:7863–7871
60. Strömsten NJ, Bamford DH, Bamford JK (2003) The unique vertex of bacterial virus PRD1 is connected to the viral internal membrane. J Virol 77:6314–6321
61. Kivelä HM, Madonna S, Krupovic M, Tutino ML, Bamford JK (2008) Genetics for Pseudoalteromonas provides tools to manipulate marine bacterial virus PM2. J Bacteriol 190:1298–1307
62. Avilov SV, Moisy D, Munier S, Schraidt O, Naffakh N, Cusack S (2011) Replication-competent influenza A virus that encodes a split-green fluorescent protein-tagged PB2 polymerase subunit allows live-cell imaging of the virus life cycle. J Virol 86:1433–1448

63. Kivelä HM, Abrescia NG, Bamford JK, Grimes JM, Stuart DI, Bamford DH (2008) Selenomethionine labeling of large biological macromolecular complexes: probing the structure of marine bacterial virus PM2. J Struct Biol 161:204–210
64. Bamford D, Mindich L (1982) Structure of the lipid-containing bacteriophage PRD1: disruption of wild-type and nonsense mutant phage particles with guanidine hydrochloride. J Virol 44:1031–1038
65. Kivelä HM, Kalkkinen N, Bamford DH (2002) Bacteriophage PM2 has a protein capsid surrounding a spherical proteinaceous lipid core. J Virol 76:8169–8178
66. Caldentey J, Luo C, Bamford DH (1993) Dissociation of the lipid-containing bacteriophage PRD1: effects of heat, pH, and sodium dodecyl sulfate. Virology 194:557–563
67. Kivelä HM, Roine E, Kukkaro P, Laurinavicius S, Somerharju P, Bamford DH (2006) Quantitative dissociation of archaeal virus SH1 reveals distinct capsid proteins and a lipid core. Virology 356:4–11
68. Pietilä MK, Atanasova NS, Manole V, Liljeroos L, Butcher SJ, Oksanen HM, Bamford DH (2012) Virion architecture unifies globally distributed pleolipoviruses infecting halophilic archaea. J Virol 86:5067–5079
69. Rissanen I, Pawlowski A, Harlos K, Grimes JM, Stuart DI, Bamford JKH (2012) Crystallization and preliminary crystallographic analysis of the major capsid proteins VP16 and VP17 of bacteriophage P23-77. Acta Crystallogr F68:580–583
70. Aricescu AR, Lu W, Jones EY (2006) A time- and cost-efficient system for high-level protein production in mammalian cells. Acta Crystallogr D Biol Crystallogr 62:1243–1250
71. Benson SD, Bamford JK, Bamford DH, Burnett RM (2002) The X-ray crystal structure of P3, the major coat protein of the lipid-containing bacteriophage PRD1, at 1.65 A resolution. Acta Crystallogr D Biol Crystallogr 58:39–59
72. Hendrickson WA (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science 254:51–58
73. Jeembaeva M, Jonsson B, Castelnovo M, Evilevitch A (2010) DNA heats up: energetics of genome ejection from phage revealed by isothermal titration calorimetry. J Mol Biol 395:1079–1087
74. Walter TS, Ren J, Tuthill TJ, Rowlands DJ, Stuart DI, Fry EE (2012) A plate-based high throughput assay for virus stability and vaccine formulation. J Virol Methods 185(1):166–170
75. Shoemaker GK, van Duijn E, Crawford SE, Uetrecht C, Baclayon M, Roos WH, Wuite GJ, Estes MK, Prasad BV, Heck AJ (2010) Norwalk virus assembly and stability monitored by mass spectrometry. Mol Cell Proteomics 9:1742–1751
76. Lucic V, Forster F, Baumeister W (2005) Structural studies by electron tomography: from cells to molecules. Annu Rev Biochem 74:833–865
77. Subramaniam S, Bartesaghi A, Liu J, Bennett AE, Sougrat R (2007) Electron tomography of viruses. Curr Opin Struct Biol 17:596–602
78. Freiberg AN, Sherman MB, Morais MC, Holbrook MR, Watowich SJ (2008) Three-dimensional organization of Rift Valley fever virus revealed by cryoelectron tomography. J Virol 82:10341–10348
79. Grünewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, Baumeister W, Steven AC (2003) Three-dimensional structure of herpes simplex virus from cryo-electron tomography. Science 302:1396–1398
80. Bowman BR, Baker ML, Rixon FJ, Chiu W, Quiocho FA (2003) Structure of the herpesvirus major capsid protein. EMBO J 22:757–765
81. Baker ML, Jiang W, Rixon FJ, Chiu W (2005) Common ancestry of herpesviruses and tailed DNA bacteriophages. J Virol 79:14967–14970
82. Bostina M, Bubeck D, Schwartz C, Nicastro D, Filman DJ, Hogle JM (2007) Single particle cryoelectron tomography characterization of the structure and structural variability of poliovirus-receptor-membrane complex at 30 A resolution. J Struct Biol 160:200–210
83. Zhu P, Liu J, Bess J Jr, Chertova E, Lifson JD, Grisé H, Ofek GA, Taylor KA, Roux KH (2006) Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature 441:847–852

84. Zanetti G, Briggs JA, Grünewald K, Sattentau QJ, Fuller SD (2006) Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. PLoS Pathog 2:e83
85. Chen B, Vogan EM, Gong H, Skehel JJ, Wiley DC, Harrison SC (2005) Structure of an unliganded simian immunodeficiency virus gp120 core. Nature 433:834–841
86. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. Nature 455:109–113
87. Harris A, Cardone G, Winkler DC, Heymann JB, Brecher M, White JM, Steven AC (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. Proc Natl Acad Sci U S A 103:19123–19127
88. Huiskonen JT, Overby AK, Weber F, Grünewald K (2009) Electron cryo-microscopy and single-particle averaging of Rift Valley fever virus: evidence for GN-GC glycoprotein heterodimers. J Virol 83:3762–3769
89. Carrascosa JL, Chichon FJ, Pereiro E, Rodriguez MJ, Fernandez JJ, Esteban M, Heim S, Guttmann P, Schneider G (2009) Cryo-X-ray tomography of vaccinia virus membranes and inner compartments. J Struct Biol 168:234–239
90. Ji X, Sutton G, Evans G, Axford D, Owen R, Stuart DI (2009) How baculovirus polyhedra fit square pegs into round holes to robustly package viruses. EMBO J 29:505–514
91. Koopmann R, Cupelli K, Redecke L, Nass K, Deponte DP, White TA, Stellato F, Rehders D, Liang M, Andreasson J, Aquila A, Bajt S, Barthelmess M, Barty A, Bogan MJ, Bostedt C, Boutet S, Bozek JD, Caleman C, Coppola N, Davidsson J, Doak RB, Ekeberg T, Epp SW, Erk B, Fleckenstein H, Foucar L, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann A, Hartmann R, Hauser G, Hirsemann H, Holl P, Hunter MS, Kassemeyer S, Kirian RA, Lomb L, Maia FR, Kimmel N, Martin AV, Messerschmidt M, Reich C, Rolles D, Rudek B, Rudenko A, Schlichting I, Schulz J, Seibert MM, Shoeman RL, Sierra RG, Soltau H, Stern S, Strüder L, Timneanu N, Ullrich J, Wang X, Weidenspointner G, Weierstall U, Williams GJ, Wunderer CB, Fromme P, Spence JC, Stehle T, Chapman HN, Betzel C, Duszenko M (2012) *In vivo* protein crystallization opens new routes in structural biology. Nat Methods 9:259–262
92. Iwasaki K, Omura T (2010) Electron tomography of the supramolecular structure of virus-infected cells. Curr Opin Struct Biol 20:632–639
93. Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia FR, Martin AV, Schlichting I, Lomb L, Coppola N, Shoeman RL, Epp SW, Hartmann R, Rolles D, Rudenko A, Foucar L, Kimmel N, Weidenspointner G, Holl P, Liang M, Barthelmess M, Caleman C, Boutet S, Bogan MJ, Krzywinski J, Bostedt C, Bajt S, Gumprecht L, Rudek B, Erk B, Schmidt C, Hömke A, Reich C, Pietschner D, Strüder L, Hauser G, Gorke H, Ullrich J, Herrmann S, Schaller G, Schopper F, Soltau H, Kühnel KU, Messerschmidt M, Bozek JD, Hau-Riege SP, Frank M, Hampton CY, Sierra RG, Starodub D, Williams GJ, Hajdu J, Timneanu N, Seibert MM, Andreasson J, Rocker A, Jönsson O, Svenda M, Stern S, Nass K, Andritschke R, Schröter CD, Krasniqi F, Bott M, Schmidt KE, Wang X, Grotjohann I, Holton JM, Barends TR, Neutze R, Marchesini S, Fromme R, Schorb S, Rupp D, Adolph M, Gorkhover T, Andersson I, Hirsemann H, Potdevin G, Graafsma H, Nilsson B, Spence JC (2011) Femtosecond X-ray protein nanocrystallography. Nature 470:73–77
94. Seibert MM, Ekeberg T, Maia FR, Svenda M, Andreasson J, Jönsson O, Odić D, Iwan B, Rocker A, Westphal D, Hantke M, DePonte DP, Barty A, Schulz J, Gumprecht L, Coppola N, Aquila A, Liang M, White TA, Martin A, Caleman C, Stern S, Abergel C, Seltzer V, Claverie JM, Bostedt C, Bozek JD, Boutet S, Miahnahri AA, Messerschmidt M, Krzywinski J, Williams G, Hodgson KO, Bogan MJ, Hampton CY, Sierra RG, Starodub D, Andersson I, Bajt S, Barthelmess M, Spence JC, Fromme P, Weierstall U, Kirian R, Hunter M, Doak RB, Marchesini S, Hau-Riege SP, Frank M, Shoeman RL, Lomb L, Epp SW, Hartmann R, Rolles D, Rudenko A, Schmidt C, Foucar L, Kimmel N, Holl P, Rudek B, Erk B, Hömke A, Reich C, Pietschner D, Weidenspointner G, Strüder L, Hauser G, Gorke H, Ullrich J, Schlichting I, Herrmann S, Schaller G, Schopper F, Soltau H, Kühnel KU, Andritschke R, Schröter CD,

Krasniqi F, Bott M, Schorb S, Rupp D, Adolph M, Gorkhover T, Hirsemann H, Potdevin G, Graafsma H, Nilsson B, Chapman HN, Hajdu J (2011) Single mimivirus particles intercepted and imaged with an X-ray laser. Nature 470:78–81

95. Lebbink MN, Geerts WJ, van der Krift TP, Bouwhuis M, Hertzberger LO, Verkleij AJ, Koster AJ (2007) Template matching as a tool for annotation of tomograms of stained biological structures. J Struct Biol 158:327–335
96. Pintilie GD, Zhang J, Goddard TD, Chiu W, Gossard DC (2010) Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. J Struct Biol 170:427–438
97. Liljeroos L, Huiskonen JT, Ora A, Susi P, Butcher SJ (2011) Electron cryotomography of measles virus reveals how matrix protein coats the ribonucleocapsid within intact virions. Proc Natl Acad Sci U S A 108:18085–18090
98. Brandt F, Carlson LA, Hartl FU, Baumeister W, Grünewald K (2010) The three-dimensional organization of polyribosomes in intact human cells. Mol Cell 39:560–569
99. Plitzko JM, Rigort A, Leis A (2009) Correlative cryo-light microscopy and cryo-electron tomography: from cellular territories to molecular landscapes. Curr Opin Biotechnol 20:83–89
100. van Driel LF, Valentijn JA, Valentijn KM, Koning RI, Koster AJ (2009) Tools for correlative cryo-fluorescence microscopy and cryo-electron tomography applied to whole mitochondria in human endothelial cells. Eur J Cell Biol 88:669–684
101. Kukulski W, Schorb M, Welsch S, Picco A, Kaksonen M, Briggs JA (2011) Correlated fluorescence and 3D electron microscopy with high sensitivity and spatial precision. J Cell Biol 192:111–119
102. Jun S, Ke D, Debiec K, Zhao G, Meng X, Ambrose Z, Gibson GA, Watkins SC, Zhang P (2011) Direct visualization of HIV-1 with correlative live-cell microscopy and cryo-electron tomography. Structure 19:1573–1581
103. Raines KS, Salha S, Sandberg RL, Jiang H, Rodriguez JA, Fahimian BP, Kapteyn HC, Du J, Miao J (2010) Three-dimensional structure determination from a single view. Nature 463:214–217
104. Reich ES (2011) Three-dimensional technique on trial. Nature 480:303
105. Hoppe W, Langer R, Frank J, Feltynowski A (1969) Image differentiation procedures in electron microscopy. Naturwissenschaften 56:267–272
106. Maiden AM, Humphry MJ, Zhang F, Rodenburg JM (2011) Superresolution imaging *via* ptychography. J Opt Soc Am A Opt Image Sci Vis 28:604–612
107. Dierolf M, Menzel A, Thibault P, Schneider P, Kewish CM, Wepf R, Bunk O, Pfeiffer F (2010) Ptychographic X-ray computed tomography at the nanoscale. Nature 467:436–439
108. Rossmann MG (2012) Crystallography, evolution, and the structure of viruses. J Biol Chem 287:9552–9559
109. McLellan JS, Pancera M, Carrico C, Gorman J, Julien JP, Khayat R, Louder R, Pejchal R, Sastry M, Dai K, O'Dell S, Patel N, Shahzad-ul-Hussan S, Yang Y, Zhang B, Zhou T, Zhu J, Boyington JC, Chuang GY, Diwanji D, Georgiev I, Kwon YD, Lee D, Louder MK, Moquin S, Schmidt SD, Yang ZY, Bonsignori M, Crump JA, Kapiga SH, Sam NE, Haynes BF, Burton DR, Koff WC, Walker LM, Phogat S, Wyatt R, Orwenyo J, Wang LX, Arthos J, Bewley CA, Mascola JR, Nabel GJ, Schief WR, Ward AB, Wilson IA, Kwong PD (2011) Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature 480:336–343
110. Bahar MW, Graham SC, Stuart DI, Grimes JM (2011) Insights into the evolution of a complex virus from the crystal structure of vaccinia virus d13. Structure 19:1011–1020
111. Coloma R, Valpuesta JM, Arranz R, Carrascosa JL, Ortin J, Martín-Benito J (2009) The structure of a biologically active influenza virus ribonucleoprotein complex. PLoS Pathog 5: e1000491
112. Abrescia NG, Bamford DH, Grimes JM, Stuart DI (2012) Structure unifies the viral universe. Annu Rev Biochem 81:795–822

113. Kaufmann B, Plevka P, Kuhn RJ, Rossmann MG (2010) Crystallization and preliminary X-ray diffraction analysis of West Nile virus. Acta Crystallogr Sect F Struct Biol Cryst Commun 66:558–562
114. Kaufmann B, Nybakken GE, Chipman PR, Zhang W, Diamond MS, Fremont DH, Kuhn RJ, Rossmann MG (2006) West Nile virus in complex with the Fab fragment of a neutralizing monoclonal antibody. Proc Natl Acad Sci U S A 103:12400–12404
115. Yu X, Qiao M, Atanasov I, Hu Z, Kato T, Liang TJ, Zhou ZH (2007) Cryo-electron microscopy and three-dimensional reconstructions of hepatitis C virus particles. Virology 367:126–134
116. Badia-Martinez D, Peralta B, Andrés G, Guerra M, Gil-Carton D, Abrescia NG (2012) Three-dimensional visualization of forming Hepatitis C virus-like particles by electron-tomography. Virology 430:120–126
117. Gastaminza P, Dryden KA, Boyd B, Wood MR, Law M, Yeager M, Chisari FV (2010) Ultrastructural and biophysical characterization of hepatitis C virus particles produced in cell culture. J Virol 84:10999–11009
118. Owen RL, Axford D, Nettleship JE, Owens RJ, Robinson JI, Morgan AW, Doré AS, Lebon G, Tate CG, Fry EE, Ren J, Stuart DI, Evans G (2012) Outrunning free radicals in room-temperature macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 68:810–818
119. Murata K, Liu X, Danev R, Jakana J, Schmid MF, King J, Nagayama K, Chiu W (2010) Zernike phase contrast cryo-electron microscopy and tomography for structure determination at nanometer and subnanometer resolutions. Structure 18:903–912

Further Reading

Abeyrathne PD, Chami M, Pantelic RS, Goldie KN, Stahlberg H (2010) Preparation of 2D crystals of membrane proteins for high-resolution electron crystallography data collection. Methods Enzymol 481:25–43

Abrescia NG, Grimes JM, Fry EE, Ravantti JJ, Bamford DH, Stuart DI (2010) What does it take to make a virus: the concept of the viral “self”. In: Twarock R, Stockley PG (eds) Emerging topics in physical virology. Imperial College Press, London, pp 35–58

Derosier D (2010) 3D reconstruction from electron micrographs a personal account of its development. Methods Enzymol 481:1–24

Henderson R, Sali A, Baker ML, Carragher B, Devkota B, Downing KH, Egelman EH, Feng Z, Frank J, Grigorieff N, Jiang W, Ludtke SJ, Medalia O, Penczek PA, Rosenthal PB, Rossmann MG, Schmid MF, Schröder GF, Steven AC, Stokes DL, Westbrook JD, Wriggers W, Yang H, Young J, Berman HM, Chiu W, Kleywegt GJ, Lawson CL (2012) Outcome of the first electron microscopy validation task force meeting. Structure 20:205–214

Johnson JE (2008) Multi-disciplinary studies of viruses: the role of structure in shaping the questions and answers. J Struct Biol 163:246–253

Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North AC (1960) Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. Resolution, obtained by X-ray analysis. Nature 185:416–422

Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q Rev Biophys 40:191–285

Read RJ, Adams PD, Arendall WB III, Brunger AT, Emsley P, Joosten RP, Kleywegt GJ, Krissinel EB, Lütteke T, Otwinowski Z, Perrakis A, Richardson JS, Sheffler WH, Smith JL, Tickle IJ, Vriend G, Zwart PH (2011) A new generation of crystallographic validation tools for the protein data bank. Structure 19:1395–1412

Steven AC, Baumeister W (2008) The future is hybrid. J Struct Biol 163:186–195
Vellieux FM, Read RJ (1997) Noncrystallographic symmetry averaging in phase refinement and extension. Methods Enzymol 277:18–53

Also especially recommended for further reading are references [7, 9, 13, 25, 27, 99, 112] *listed above*.

Chapter 8
Atomic Force Microscopy of Viruses

Pedro J. de Pablo

Abstract Atomic force microscopy (AFM) is a helpful tool to acquire nanometric-resolution images, and also to perform a certain physical characterization of specimens, including their stiffness and mechanical resilience. Besides of the wide range of applications, from materials science to biology, this technique works in a variety of conditions as long as the sample is supported on a solid surface, in air, ultra high vacuum or, most importantly for virus research, in liquids. The adaptability of this technique is also fostered by the variety of sizes of the specimens that it can dealt with, such as atoms, molecules, molecular complexes including viruses and cells, and the possibility to observe dynamic processes in real time. Indeed, AFM facilitates single molecule experiments enabling not only to *see* but also to *touch* the material under study (*i.e*., to undertake mechanical manipulations), and constitutes a fundamental source of information for material characterization. In particular, the study of the mechanical properties at the nano-scale of viruses and other biomolecular aggregates, is providing an important set of data which help to elaborate mechano-chemical structure/function models of molecular biomachines, expanding and complementing the information obtained by other structural techniques.

Keywords Atomic force microscopy • Force • Beam deflection • Tip • Cantilever • Stylus • Topography • Mechanics • Liquids • Tip-sample dilation • Adsorption • Virus • Capsid • Virion • Deformation • Nanoindentation • Force curve • Disruption • Breakage

---

P.J. de Pablo (✉)
Department of Physics of the Condensed Matter, C03, Facultad de Ciencias,
Universidad Autónoma de Madrid, Campus de Cantoblanco, 28049 Madrid, Spain
e-mail: p.j.depablo@uam.es

M.G. Mateu (ed.), *Structure and Physics of Viruses: An Integrated Textbook*,
Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_8,
© Springer Science+Business Media Dordrecht 2013

Abbreviations

AFM	Atomic force microscopy
CCD	Charge-coupled device
CM	Contact mode
DM	Dynamic mode
EDL	Electrostatic double layer
EM	Electron microscopy
HOPG	Highly oriented pyrolytic graphite
JM	Jumping mode
MVM	Minute virus of mice
STM	Scanning tunneling microscopy
UHV	Ultra-high vacuum

8.1 Introduction

The word *microscope* suggests an optical device that uses light to achieve a magnified image of a small object. Consequently, the question that arises when seeing an atomic force microscope (AFM) for the first time is: where do I have to look to see the object? In a conventional microscope (Fig. 8.1), a source emits particles, such as electrons or photons, that interact with the specimen. A detector registers and communicates this interaction to an analyzer, which processes the information received to make it comprehensible. Optical and electron microscopes fit into this source-specimen-detector-analyzer scheme.

In the optical microscope, the photons emitted by the incandescent lamp are conveniently manipulated by a system of lenses, both before and after the interaction with the specimen, finally arriving to the eyepiece where the detector (*i.e*. the eyes, film or charge-coupled device, CCD) collects the information. A typical optical microscope using visible light can reach a resolution of $\lambda/2 \sim 200$ nm (λ being the wavelength of the light).

In the case of electron microscopy (EM) (see Chap. 3), a filament emits electrons and a system of electromagnetic lenses are used to manipulate and focus the electron beam both before and after the interaction with the specimen. Afterwards the electrons are collected by a phosphorescent screen, film or CCD. In 1933 the German engineering Ernst Ruska constructed an electron microscope that exceeded the maximum resolution of an optical microscope, reaching about 1 nm.

AFM is not based in registering interactions between probe particles, such as photons and electrons, and the sample. In AFM the probe is a nanometric tip at the end of a microcantilever which mechanically palpates the object under study like a blind person uses a walking stick. In this way AFM enables the construction of a topographic image based upon recording the deflection angle of the microcantilever at each point of the studied sample. In this chapter, I describe the conceptual and

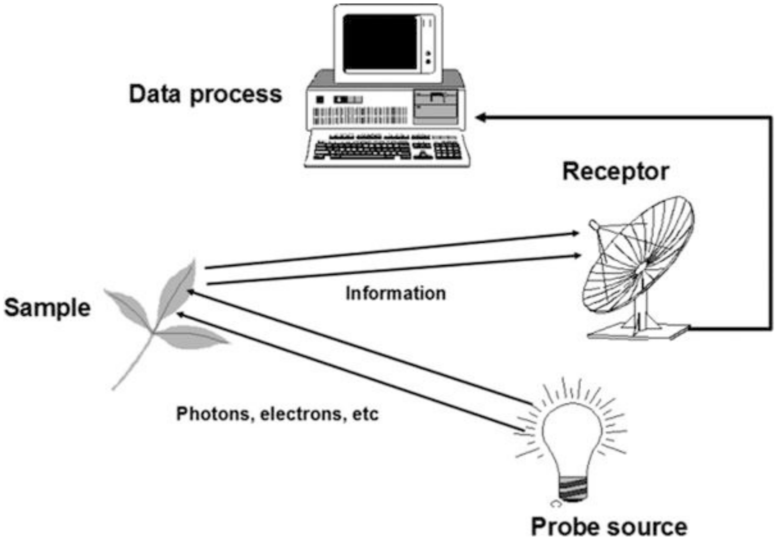


Fig. 8.1 Concept of microscopy. Any microscope requires probe source, sample, receptor and data analysis

technical bases of AFM. In particular, I refer to working methods in liquids and its applications for imaging virus particles and viral components (Fig. 8.2) In a related Chapter (Chap. 18), the application of AFM to determine the mechanical properties of virus particles is described.

8.2 Basic Concepts

In scanning probe microscopy a sharp stylus of a few nanometers in diameter, which can be considered as a probe, approaches the surface of the sample. Binnig and Rohrer [7] invented the Scanning Tunneling Microscope (STM) and received the Nobel Prize for Physics along with Ruska in 1986. The STM is based on a quantum effect (tunneling) that occurs when a sharp metallic tip is brought to a distance (z) of less than 1 nm from a conductive surface. This effect involves the flow of an electronic current (I) between the surface and the tip according to the formula $I \propto \exp(-\sqrt{\phi}z)$, where ϕ is the work-function of the metallic surface [8]. The strong dependence on the tip-surface distance can be used to obtain topographic and electronic maps of the sample by moving, (*i.e.* scanning) the tip on the surface while keeping the tip-sample distance constant through a feedback algorithm. Although this tool provides true atomic resolution in ultra-high vacuum (UHV) conditions, a mandatory requisite is that both the tip and sample should be conductive. Therefore, STM is not suitable for biological samples since these are mainly insulators, which need to be covered with a metallic layer [9].

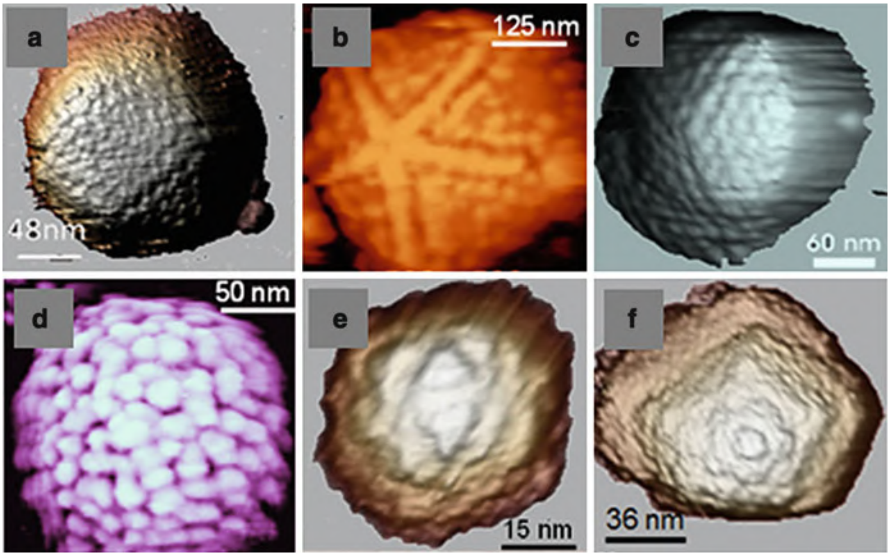


Fig. 8.2 A gallery of virus images obtained by AFM. (**a**) Human adenovirus, taken from Ref. [1]. (**b**) Giant mimivirus, taken from Ref. [2]. (**c**) Herpes simplex virus, taken from Ref. [3]. (**d**) Moloney murine leukemia virus, after taken from [4]. (**e**) MVM, taken from [5]. (**f**) bacteriophage T7, taken from [6]. All images are reproduced with permission

In 1986 Binnig, Quate and Gerber invented the AFM [10] combining the principles of both the STM and the so-called stylus profilometer [11]. In an AFM (Fig. 8.3) a sharp tip (approximately tens of nm in radius) attached to the end of a microcantilever is approached to the surface of an object by means of a piezoelectric device. As a consequence a force appears between the tip and surface that can be attractive or repulsive (see below) causing the cantilever to bend. A feedback algorithm acting on the piezoelectric device controls the relative tip-surface distance, and a topographic map of the object is obtained by scanning the surface in a plane perpendicular to the tip. The original article by Binnig et al. [10] presents the topographic profiles obtained for an insulator (ceramic). The fact that both tip and sample may be insulators is one of the main advantages of AFM, opening the door to a whole range of possibilities such as the study of biological samples including proteins, nucleic acids, membranes, cells, and biomolecular complexes such as viruses.

8.3 AFM Implementation

Although there are a variety of routes to control the deflection of the cantilever in AFM, here we will focus on the beam deflection method [12] since it is commonly employed when working with biological samples. The beam deflection system

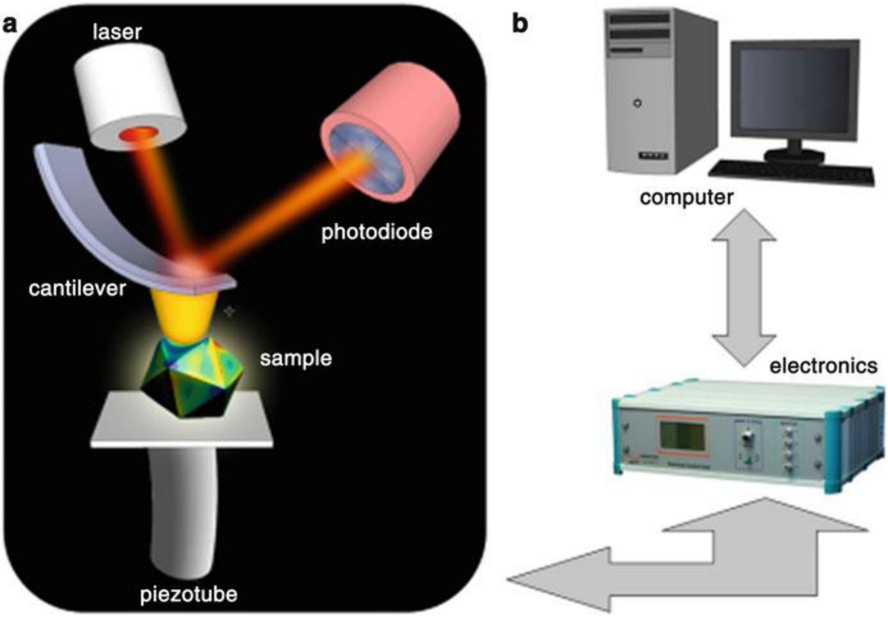


Fig. 8.3 Principle of atomic force microscopy. The software running on the computer controls the electronics (**b**), which gives and receives information to and from the AFM head (**a**)

involves focusing a laser beam on the end of the cantilever and collecting the reflected light with a photodiode (Fig. 8.3a). As a consequence, any bending of the cantilever will affect the position of the reflected laser spot on the photodiode. A normal bending will originate a so-called normal force signal, $F_{n,}$ on the photodiode sectors, whereas a lateral torsion will result in the so-called lateral force, F_l.

The core of an AFM is the head (Fig. 8.3a), where the beam deflection system is integrated along with the piezoelectric device (piezo-tube) that moves the sample in all three directions (x, y, z). In Fig. 8.3a an AFM head configuration is shown where the tip is fixed and the sample is moved by the piezo-tube. Another configuration known as "stand alone", fixes the sample and the tip makes the scanning movement controlled by a piezo-tube that carries the cantilever. The electronic components receive the signals coming from the photodiode, mainly F_n and F_l, and provide high voltages (up to hundreds of volts) to move the piezo-tube to which the sample (or cantilever) is attached. The computer is in charge for managing the data and calculating all the parameters required to move the piezo-tube in a convenient way by using the electronics (Fig. 8.3b). Details on the different components of a typical AFM and AFM operation are given in the following subsections.

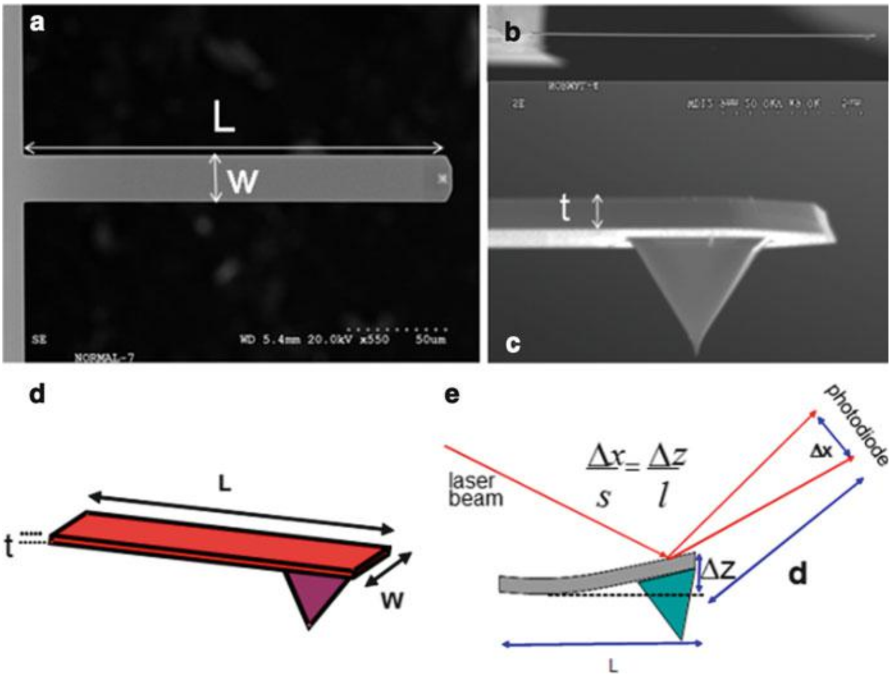


Fig. 8.4 AFM cantilevers. (**a**) and (**b**) respectively show top and side views of a rectangular cantilever of 150 μm in length (L) and 25 μm in width (w) attached to a chip. (**c**) presents a close-up image of a pyramidal tip, and a cantilever with a thickness of 150 nm. (**d**) sketches the geometry of a rectangular cantilever. The cartoon in (**e**) indicates the vertical resolution of the bending cantilever as a function of several geometric parameters

8.3.1 Cantilevers

Integrated tip and cantilever assemblies can be fabricated from silicon or silicon nitride using photolithographic techniques. More than 1,000 tip and cantilever assemblies can be produced on a single silicon wafer. The cantilevers can be rectangular (Fig. 8.4a) or V-shaped, and they typically range from 60 to 200 μm in length, 10–40 μm in width, and 0.3–2 μm in thickness. The usual tip radius is about 20 nm although lower diameters can be obtained (Fig. 8.4b). The mechanical stiffness of the cantilever (see Chap. 18) is given by its spring constant k, which normally ranges between 0.03 N/m and 40 N/m, and strongly depends on the cantilever's dimensions. For example, for a rectangular cantilever $k = \frac{EW}{4}\left(\frac{t}{L}\right)^3$, where E is the Young Modulus (a measurement of the intrinsic elasticity of a solid material, see Chap. 18) of the cantilever, and W, t and L are the cantilever width, thickness and length respectively (Fig. 8.4d). While W and L can be fairly precisely known, t is always difficult to measure. As a consequence, the manufacturers normally provide the cantilever spring constant with an error of 10–30 % and therefore the user should calibrate each cantilever [13]. The spring constant k is

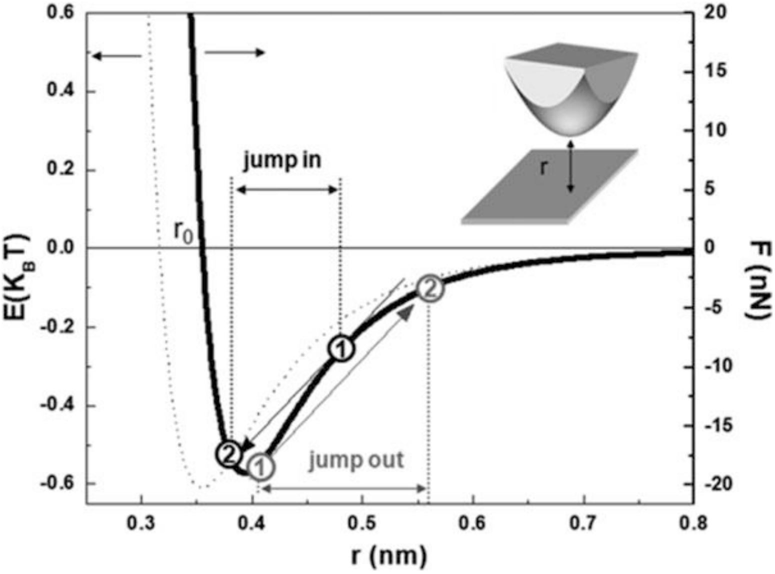


Fig. 8.5 Tip-surface interaction. *The dotted line* represents a Lennard-Jones potential as a function of the distance between atoms, *r*, mimicking the tip-surface interaction (represented in the *inset*). The *solid line* is the interaction force obtained from the potential. The numbers *inside the circles* indicate the cantilever instabilities (see text for a complete description)

used to calculate the force applied on the cantilever as a function of bending Δz using Hooke's law, *i.e.*, $F = k \times \Delta z$ (Fig. 8.4e).

The vertical resolution of the cantilever, Δz, strongly depends on the noise of the photodiode, since it defines the minimum significant displacement Δx, on the detector (Fig. 8.4e). A typical cantilever of 100 μm length has about 0.1 Å resolution assuming a signal-to-noise ratio around 1 [13].

Another important parameter that can have influence on the vertical resolution is the thermal noise [14]: the cantilever oscillates at the resonance frequency with an amplitude $\Delta z = \sqrt{\frac{k_B T}{k}}$, where k_B is the Boltzman constant, T is the absolute temperature and k the cantilever spring constant. For example, at room temperature, there is a noise of about 5 Å for a cantilever of spring constant 0.02 N/m.

8.3.2 Interaction Between the AFM-Tip and the Sample Surface

In order to understand the interaction between tip and sample we shall refer to potentials rather than forces. Physicists prefer this approach as potentials are scalar and therefore easier to deal with compared to vectors such as forces. For this purpose let us graphically depict the Lennard-Jones potential in Jules (J) between two atoms separated by r meters (m) in vacuum [15] (dotted curve in Fig. 8.5).

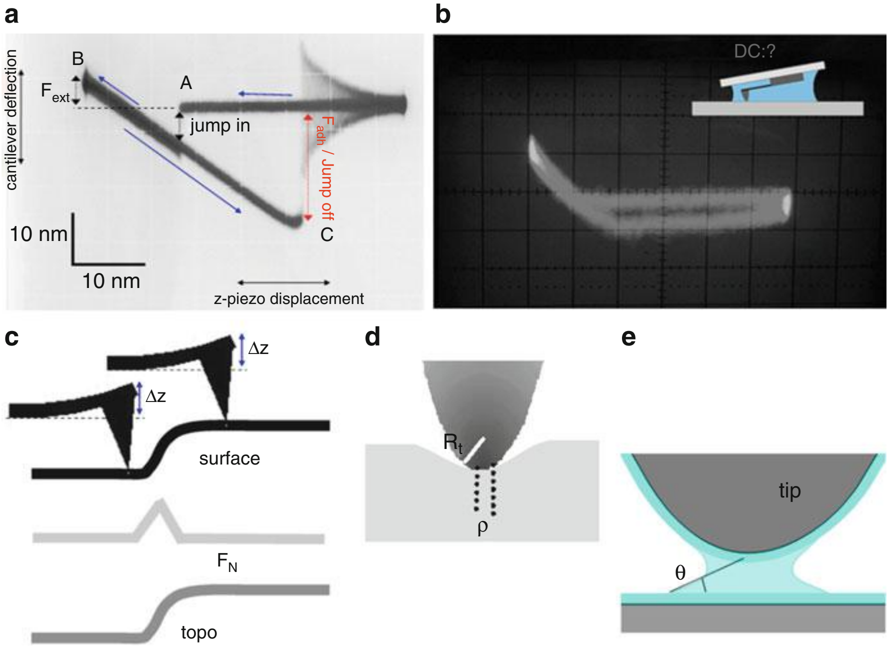


Fig. 8.6 Force vs. distance experiments, feedback and lateral resolution. (**a**) and (**b**) show oscilloscope snapshots of force vs. *z*-piezo displacement curves in air and water, respectively. (**c**) presents the contact mode showing the variation of F_n (*light gray*), and *z*-piezo tube voltage (topography, *dark gray*) as a function of a step (*black*). In (**d**) the geometric features of the tip-surface contact are depicted. (**e**) shows a cartoon illustrating the capillary forces between tip and surface (see text for a complete description)

$U(r) = -A/r^6 + B/r^{12}$, where r is the tip-sample distance, $A = 10^{-77}\ J\ m^6$ and $B = 10^{-134}\ J\ m^{12}$.

A rough approximation to the AFM tip-sample interaction is to consider the approach of two such atoms where, as depicted in the inset in Fig. 8.5: the lower part shows a solid surface and the upper part shows the apex of a sharp tip is moving downwards. An interaction force $F(r) = -\frac{dU}{dr}$ will occur and it can be seen that this force is attractive ($F < 0$) when $r > r_0$ or repulsive ($F > 0$) when $r < r_0$, (solid line curve in Fig. 8.5). These regions define the attractive and the repulsive regimes of operation, respectively.

Now let us consider the role of these regimes in a force vs. distance (FZ) AFM experiment that involves the tip approaching to the surface (Figs. 8.5 and 8.6). The experiment starts with the tip situated far from the surface in the attractive regime. As the tip is approaching and as soon as the gradient (*i.e.*, the slope) of the force equals the cantilever spring constant, the tip jumps to the surface from black point 1 to black point 2 in Fig. 8.5 (both connected by the slope). This effect is seen in the FZ of Fig. 8.6a at point A like a sudden jump of the cantilever deflection (vertical scale). In this way the tip establishes mechanical contact with the surface and it rapidly enters the repulsive regime ($F_N > 0$). The *z*-piezo (horizontal scale) is elongated until a given F_N or deflection value is reached and then stops (point B in Fig. 8.6a). The external loading force F_{ext} can be calculated as the difference between the zero deflection position (*i.e.* before the jump to contact in point A in Fig. 8.6a) and the deflection at point B in Fig. 8.6a. Since the vertical scale is about 10 nm per division and the cantilever spring constant $k = 0.1\ N/m$ we have $F_{ext} \sim 0.5div \times 10nm/div \times 0.1nN/nm = 0.5nN$.

Subsequently, the z-piezo retrace cycle starts, and the tip is released from the surface at point C in Fig. 8.6a (*i.e.*, once more where the derivative of the tip-surface force equals the cantilever spring constant) jumping from the grey point 1 to grey point 2 in Fig. 8.5, following the grey arrow. The cantilever deflection jumps-off to zero with a dampened oscillation. This jump-off is known as the adhesion force F_{adh} (in Fig. 8.6a $F_{adh} \sim 2div \times 10\ nm/div \times 0.1nN/nm = 2\ nN$). The total force at point B in Fig. 8.6a is the sum of F_{ext} and F_{adh}, (*i.e.*, 2.5 nN). It is interesting to stress that no matter how small F_{ext} is, the total force applied to the surface will always be at least F_{adh}.

8.3.3 Contact Mode (CM)

CM is the simplest operational method used for AFM and it was the first to be developed [10]. Here the tip is brought into contact with the surface until a given deflection in the cantilever (F_n) is reached; the tip then scans a square area of the surface to obtain a topographic map. By elongating or retracting the z-piezo, the feedback algorithm tries to maintain the cantilever deflection constant by comparing the F_n signal with a set point reference value established by the user. Topographic data are obtained by recording the z-piezo voltage that the feedback algorithm is applying to correct the cantilever deflection at each position on the surface. Since the z-piezo is calibrated, voltages are transformed into heights and a topographic map is obtained. Let us consider a simple example where the tip is scanning a step (black line in Fig. 8.6c) with $F_n = k \times \Delta z$. When the cantilever moves to the upper part of the step, it undergoes a deflection greater than Δz. Therefore the feedback algorithm retracts the z-piezo in order to achieve the same deflection Δz as when the tip was in the lower part of the step and, as a consequence, a topographic profile of the step is obtained (blue line in Fig. 8.6c). On the other hand F_n varies at the step that is corrected by the feedback, which can be observed as a peak in the deflection signal (red line in Fig. 8.6c). The latter is known as the constant deflection mode or constant height mode, since the z-piezo elongation is not modified and a map of the changes in F_n is obtained. The reader is encouraged to reproduce the profiles of Fig. 8.6c when the tip goes down the step.

Let us consider now the lateral resolution that can be achieved in contact mode. We can make an estimate of this parameter by applying the Hertz theory [16], which accounts for the deformation of solids in contact. Once the tip is in contact with the sample, the radius ρ of the tip-surface contact area is given by (Fig. 8.6d):

$$\rho = \left(\frac{3 \cdot F \cdot R}{4 \cdot E^{*}}\right)^{1/3}$$

Where F is the applied force *i.e.* F_n; E^* is the effective Young modulus expressed by

$$\frac{1}{E^{*}} = \frac{1-\nu_t^2}{E_t} + \frac{1-\nu_s^2}{E_s}$$

With E_t, ν_t and E_s, ν_s being the Young Modulus and Poisson ratio for the tip and sample respectively. R is the effective radius expressed as a combination of the tip R_t and sample radius R_s,

$$\frac{1}{R} = \frac{1}{R_{tip}} + \frac{1}{R_{sample}}$$

It is worth to stress that the main condition for Hertz model applicability is that the radius of contact ρ has to be much less than the effective radius R: ρ<<R. This approximation roughly implies that the deformation is smaller than the dimensions of the tip-sample system [16]. It is important to recall that this assumption does not limit Hertz model to deformations of flat samples. In the case of metals $E = 100$ *GPa*, $\nu \sim 0.5$ and the mechanical thermal noise of the cantilever is 10 pN, with $R_t \sim 20\ nm$, the radius of contact ρ is about 0.15 nm, which implies atomic resolution. However, the adhesion force in air (see below) is about 5 nN, increasing ρ to about 1 nm. Atomic resolution can be achieved by either working in liquids [17] or in UHV conditions [18], where even individual atoms can be chemically identified [19].

CM is not very frequently used to image isolated biological materials such as virus particles. The very high lateral forces that are applied to the surface [20], which can damage the sample, present an important problem when using this mode. This is especially important for single mechanically delicate biomolecules adsorbed onto a surface.

8.3.4 Dynamic Modes (DM)

In DM operation [21] the cantilever is oscillated near to or at its resonance frequency ($\omega_0 \propto \sqrt{E}t/L^2$ for a rectangular cantilever). As the tip approaches the sample the oscillating amplitude decreases until it establishes contact with the sample (Fig. 8.7a), following a similar cycle to that in Fig. 8.6a but now with oscillation. Therefore the feedback loop involves the amplitude rather than the F_n and, by keeping the oscillation amplitude constant, a topographical map of

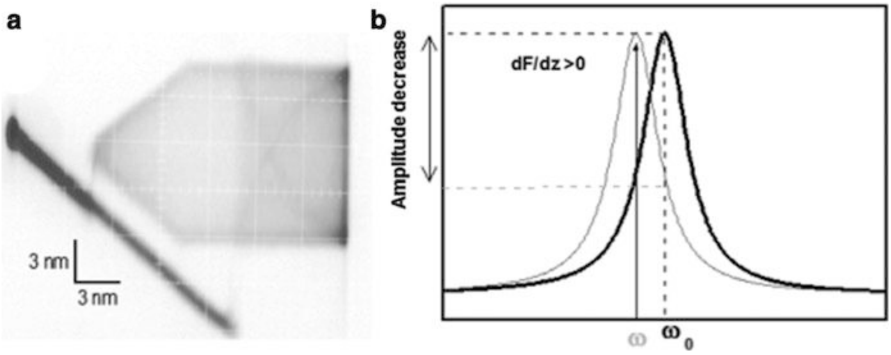


Fig. 8.7 Dynamic modes. (**a**) depicts the force vs. Piezo-tube z displacement curve of an oscillating cantilever. (**b**) illustrates the reduction of the resonance frequency of the cantilever from ω_0 to ω as the tip approaches to the surface (see text)

the object can be obtained. The amplitude is reduced because the resonance frequency (ω) changes with the tip-sample distance (z) and with the tip-sample interaction force F_{ts} according to $\frac{\Delta\omega}{\omega_0} \propto \left(\frac{1}{2k}\right)\frac{dF_{ts}}{dr}$ [22], thereby decreasing with attractive forces. The new resonance frequency ω, is lower than ω_0 and, since the cantilever is still oscillating at ω_0, the cantilever amplitude decreases (Fig. 8.7b).

While CM can damage isolated biological samples on surfaces, DM operated in non-contact mode [22] does not apply high dragging forces and is commonly used to image weakly attached molecules, including delicate biomolecules, in air. In addition to topographic maps, DM can be used to obtain other information on the sample, including a phase map (the time difference between the excitation and the response of the cantilever measured for each point in the sample), which in air carries information on the different materials composing of the sample [22].

When DM is used in liquid the situation changes completely, as the viscosity of water reduces the resonance frequency about fourfold and the quality factor (a measure of the cantilever damping) of the oscillation is reduced from ~100 to 10. When oscillating the cantilever in liquid a mechanical contact between tip and sample is established, thereby applying lateral and normal forces that may damage the specimen, especially when delicate biological objects are imaged [23]. Thus, while DM can be advantageously used to image virus particles in air, its use to image them in more physiological (liquid) conditions may be limited, unless fixation agents such as glutaraldehyde are used [24].

8.3.5 Jumping Mode (JM) or Pulse Force Mode

JM, also called pulse force mode [25, 26] is a variation of CM where lateral tip displacement occurs when the tip and sample are not in mechanical contact, thereby avoiding to a large extent shear forces and the corresponding damage to the tip-sample system. JM performs a FZ curve (Fig. 8.6a) at every point of the sample (pixel in the image), moving the tip to the next point at the end of each cycle when the tip and sample loose contact. Feedback is engaged at point B in Fig. 8.6a, moving the piezo-tube along the z-axis in a convenient way to maintain a constant deflection or loading force F_{ext}.

JM routinely produces adhesion force maps, which provides compositional or geometrical information of the surface [27]. The adhesion force between the tip and the surface can be described as $F_{adh} = 4\pi R\gamma_L \cos\theta + 1.5\pi\Delta\gamma R$, where γ_L is the water surface tension and θ the angle of the water meniscus present between the tip and surface; R is the effective radius (described above) and $\Delta\gamma$ the tip-surface energy difference. Since the first term accounts for the water meniscus between tip and surface (Fig. 8.6e), it mainly informs about the hydrophobicity of the sample. A rough estimation of adhesion force in air at room temperature results in $\sim$7 nN for a tip radius of about 20 nm. The second term depends mainly on the tip-surface geometry. We denote the importance of the first term by comparing the FZ curves taken from glass in both air (Fig. 8.6a) and water (Fig. 8.6b): in liquid the adhesion force is almost absent, since there is not water meniscus between the tip and sample, although some hysteresis appears in the FZ due to the dragging of the water on the cantilever. In our view, JM is a mode of choice for analyzing virus particles in liquid in close to physiological conditions, because it minimizes lateral forces and enables a precise control of the force applied [28].

8.3.6 Tip-Sample Geometrical Dilation

All of the operational modes described above give rise to an effect which is inherent to the very nature of the AFM approach: the so-called tip-sample geometrical dilation. Because of this geometry-based effect, the topography of an object in an AFM image appears as laterally expanded (see Fig. 8.8). AFM users must understand and control this artefact as much as possible in order to preclude any misunderstanding of the data obtained.

When surface asperities are of dimensions comparable to the tip radius, which is very common with many samples imaged by AFM, including virus particles, the size of the tip plays an important role on image definition and the apparent image lateral expansion due to the dilation of certain features of the image by the finite tip size [29]. Figure 8.8a, b present AFM images of single-walled carbon nanotubes on silicon oxide; these images were obtained using tips of different radii, lower in Fig. 8.8a and higher in Fig. 8.8b. A carbon nanotube is rolled-up graphene in the form of a cylinder of a few nanometers in diameter. Fig. 8.8c shows the geometric dilation effects with the different tip radii: a single topography scan line across a carbon nanotube results wider for Fig. 8.8b (black) than for Fig. 8.8a (red). This effect can be seen to arise from the geometrical considerations shown in Fig. 8.8d, and the tip radius r_t can be calculated as $r_t = b^2/2h$, where b and h are

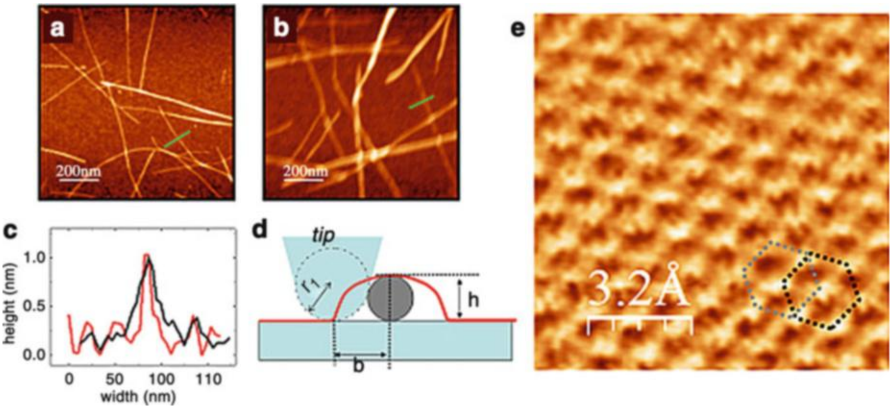


Fig. 8.8 Tip-sample dilation and atomic corrugation in HOPG. (**a**) and (**b**) show AFM images of the same samples of carbon nanotubes on silicon oxide. The topographic profiles obtained along the *green lines* in (**a**) and (**b**) show different widths due to dilation in (**c**). (**d**) illustrates the geometric parameters in the dilation process. The *red line* depicts the dilated section of the carbon nanotube section (*grey circle*). (**e**) presents the atomic corrugation of HOPG, where the two carbon layers results in two hexagonal networks

the width and the height of the sample, respectively. By using this formula, it can be calculated that the topographies of Fig. 8.8a, b have been obtained using tips with a radius of 15 nm or 70 nm, respectively. We can conclude that the sharper the tip, the better the image (which will become less laterally expanded and more defined). A popular and impressive experiment frequently used to teach AFM is to image highly oriented pyrolitic graphite (HOPG) in air [30], where the elastic deformation of the tip-sample contact plus the dilation effects result in atomic corrugation (Fig. 8.8e), although atomic defects are not visualized [31]. During imaging by AFM of virus particles and other biomolecular complexes, the height of the object at each defined point can be determined with subnanometer precision; however, the tip-sample dilation effects are important when considering lateral image dimensions and the object fine details.

8.4 Imaging Viruses and Other Biological Objects

8.4.1 Nucleic Acids, Proteins and Biomolecular Complexes

During the virus life cycle, many interactions between viral components and between viral and cellular components occur. We will review here some applications of AFM for imaging biological molecules, biomolecular complexes and their interactions.

AFM imaging in biology can be carried out in air or in liquid. Some relevant examples of AFM imaging of biomolecules in air are described first. For example, DM is frequently used for imaging DNA on a mica surface. As both DNA and mica are negatively charged, $MgCl_2$ is added to the DNA solution; the Mg^{2+} ions are sandwiched between DNA and mica, resulting in the adsorption of DNA molecules on the surface. Subsequently, the sample is dried out and DM is used to image the DNA molecules on the surface. Some structural and physical aspects of the DNA molecule have been the focus of much research using AFM [32]; in addition, an important biological application of AFM is to investigate the binding of proteins to DNA [33–35]. In this kind of single-molecule experiments the researcher is not probing the average result of the bulk reaction, but rather the action of single protein molecules on a single DNA molecule. Once the protein solution is pipetted into the DNA solution, the reaction starts and the DNA-protein complex is adsorbed onto the mica at the appropriate time and air-dried. Therefore, on the surface there is a snapshot of the process taking place between proteins and DNA, and the topography provides single molecule information of the protein-DNA complex.

However, in biology we are particularly interested in imaging in a liquid milieu in close to physiological conditions. This is an important point since biomolecules are normally fully functional only in aqueous solutions. In such a case, several particular aspects should be considered. First, a strong enough attachment of the biomolecule to the supporting surface is required to achieve good resolution during the AFM imaging process in liquid. Although biomolecules can be covalently linked to a chemically modified surface [36], covalent modifications could potentially damage the biomolecules, and therefore this treatment tends to be avoided. Fortunately, physisorption is usually sufficient and thus, the specimens can be directly adsorbed to a solid surface in a physiological buffer. The relevant forces that drive the physisorption process are the van der Waals force, the electrostatic double-layer force (EDL force) and the hydrophobic effect [37]. Unlike the van der Waals interaction, the EDL force depends strongly on the concentration and valence of charged solutes, as well as the surface charge density of both surface and specimen. The EDL force between two equally charged surfaces is repulsive and hence opposite to the van der Waals attraction [38]. One of the most extended AFM applications is to image 2-dimensional protein crystals in liquid, as well as lipid membranes, by using contact mode. Biological macromolecules become attached to the surface (mica, silicon, gold, glass, etc.) when there is a net attractive force between them and the surface pulling their surfaces into contact. This force per unit of surface (pressure) can be estimated to the light of the DLVO theory [15] like the sum of the electrostatic force between surface and molecule F_{el}, and the van der Waals interaction F_{vdW}.

$$F_{DLVO} = F_{el}(z) + F_{vdW}(z) = \frac{2\sigma_{surf}\sigma_{sample}}{\varepsilon_e\varepsilon_0} e^{-z/\lambda_D} - \frac{H_a}{6\pi z^3}$$

where z is the distance between the surface and specimen; σ_{surf} and σ_{sample} are the charge densities of surface and specimen, respectively; ε_e is the dielectric constant of the electrolyte; ε_0 is the permittivity of vacuum; λ_D is the Debye length, that depends on the electrolyte valence [37]; and H_a the Hamaker constant. The adsorption of a sample onto freshly cleaved mica (atomically flat) can be manipulated by adjusting both the ion content and the pH of the buffer solution.

AFM in liquid is being increasingly used in molecular and cell biology [39–44]. A major application is to produce topographic images of single biomolecules and, especially, larger biomolecular complexes, lipid membranes and their embedded proteins, and whole cells in close to physiological conditions. Furthermore, by introducing variations in the conditions (*i.e.*, pH changes, addition of other biological molecules, etc.) the resulting changes in the general surface structure of the biomolecule or biological object of interest can be identified. Interestingly, in this kind of set-up the cantilever can be used also to apply forces that trigger conformational changes in single protein~~s~~ molecules such as in the chaperonine GroEL; the changes thus induced may mimic those that occur under the action of physiological agents and can be revealed in real-time by AFM imaging [45].

AFM is being used also to visualize single proteins doing mechanochemical work. DM in liquid is generally used for this purpose since it is faster [46]. Maximum peak forces of a few nN are applied [23], in relation to the stiffness of the sample [47], and, although these forces could in principle be damaging to the sample, they are applied for very short periods of time such as 10 % of an oscillating period (*i.e.* for cantilever with a resonance frequency of 10 kHz in liquid, the forces are applied during 10 μs every 0.1 ms). Using this method, it has been possible to visualize, for example, the activity of RNA polymerase on DNA [48] or the conformational changes in a DNA-repair complex upon binding DNA [49]. Further high speed AFM developments have enabled to monitor the activity of single molecular machines in real-time [50].

In addition to imaging, AFM is also being applied in biology to measure the forces that participate in interactions between and within single biomolecules. AFM and optical tweezers (see Chap. 9) are major single-molecule techniques contributing to a better physics-based understanding of protein folding and unfolding, protein-ligand recognition, macromolecular assembly and disassembly, cell adhesion, etc. (*e.g.*, see references [32–35, 39–44, 46].

8.4.2 Virus Particles

Structural and chemico-physical characterization has been critical to understand the biology of viruses. X-ray crystallography (Chap. 4) and EM (cryo-EM) (Chap. 3) techniques have traditionally been used and provide direct three-dimensional structural information on whole virus particles, allowing both the interior and the surface of the virus to be visualized at high resolution. However, a limitation of these techniques is that they are both averaging ("bulk") techniques and thus they present an average time and space model of the entire population of particles in the crystal or of the many particles selected for cryo-EM image averaging. As such, these techniques generally provide limited information on possible structural differences between individual particles in the population that distinguish them from the average structure. For this reason, the beautifully symmetrical and apparently perfect models of larger viruses derived from these techniques may be somewhat deceptive, as they may not be entirely representative of the individual viruses. Therefore, the imaging of individual virus particles may provide unique insights on the structural biology of viruses and should be considered as an important part of virus characterization [51]. In addition, AFM allows imaging viruses in liquid state and following in real-time the structural changes of individual virus particles in response to different conditions and added agents, provided the changes are not too fast (see also Sect. 6). This may enable the analysis, for example, of the stages of virus assembly or disassembly [52]. Virus particles can be disrupted upon indentations above the rupture strength [3, 53, 54], and mechanical fatigue below the rupture can be used to monitor in real time the intermediate stages of adenovirus disassembly [1]. Moreover, AFM can be used to determine some physical properties of virus particles including mechanical features (see Chap. 18). Finally, AFM enables to image and quantitate the interactions of single virus particles with other molecules and even host cells [55].

Since individual virus particles adsorb weakly to typical surfaces, they are prone to undesired modifications by lateral forces in CM or DM operation, as they are not held by a surrounding neighbourhood, unlike a protein in a 2-dimensional crystal. The typical approach for surpassing this limitation is the use of structural fixation agents, such as glutaraldehyde. In such conditions AFM may provide images whose resolution is comparable to that of some EM images [56], although still far less than that of the best structural models of virus particles recently obtained by cryo-EM (see Chap. 3). Nevertheless, since glutaraldheyde structurally reinforces the specimens subjected to study [57, 58], it precludes any characterization of processes or properties of intact native viruses, such as disassembly or mechanical features [59].

Chemical fixation is, however, not required in many instances to obtain good images or probe mechanical properties of viruses. The adsorption of viruses on solid surfaces is based on the same principles than those involved in the adsorption of DNA and proteins (Sect. 8.4.1). Although the typical approach implies the convenient sylanization of glass surfaces [60], treated mica [61] and Highly Oriented Pyrolitic Graphite (HOPG) (21) have also proved as adequate substrates for virus attachment.

In general, JM has been the preferred mode to analyze unfixed viruses, since the loading forces can be accurately controlled and lateral forces are minimized. Firstly, let us exemplify the tip-virus geometrical dilation with a prolate icosahedral virus. Figure 8.9a shows a topographic image of a bacteriophage ϕ29 virion. Its comparison with the EM model [62] is shown in Fig. 8.9b: AFM enables a reasonable structural delineation of the viral head, the neck region and the tail. When the tip-sample geometrical algorithm [29] implemented in the WSxM software [63] is applied, one observes (Fig. 8.9b) how the EM model (light contour) is dilated to the black contour in the AFM image (see Sect. 3.6). Figure 8.9c demonstrates how the geometrical dilation of the image of a ϕ29 virion adsorbed

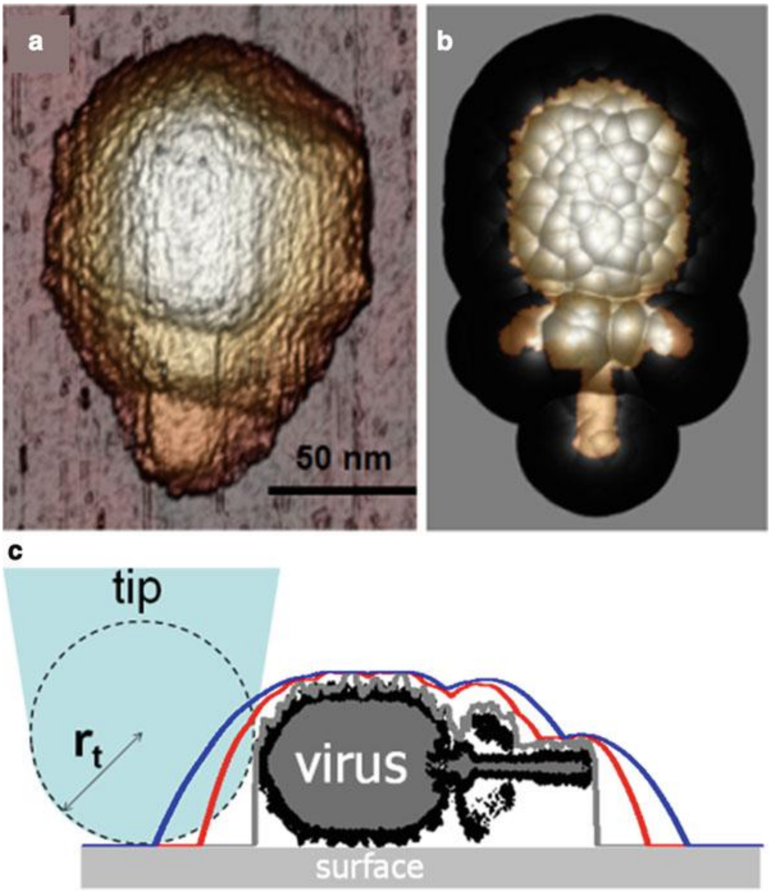


Fig. 8.9 Dilation effects in bacteriophage ϕ29 AFM images. (**a**) shows the topography of a single phage ϕ29 particle. The *bright color* image in (**b**) represents a EM structural model of ϕ29, and the *black area* indicates the dilation effects corresponding to a tip of 10 nm in diameter. The *cartoon* in (**c**) indicates the dilation effect as a function of the tip size

on a surface increases with the tip radius. A 0.5 nm-radius tip produces the grey color profile, which accurately follows the virus shape. In turn, tips of 5 and 10 nm in radius (red and blue lines, respectively) show an evident dilation and also blurring of the structural details in the virion image.

In the following, let us consider some geometrical aspects of the adsorption of icosahedral viruses (see Chap. 2) to a solid surface. For instance, the roughly spherical minute virus of mice (MVM) has a capsid of icosahedral symmetry with conspicuous protrusions at the threefold symmetry axes. Single MVM particles adsorb to the solid surface in different orientations, which frequently leaves a threefold, twofold or fivefold symmetry axis close to the top of the particle (see Figs. 8.10a–c, respectively). The possibility to identify the orientation of single

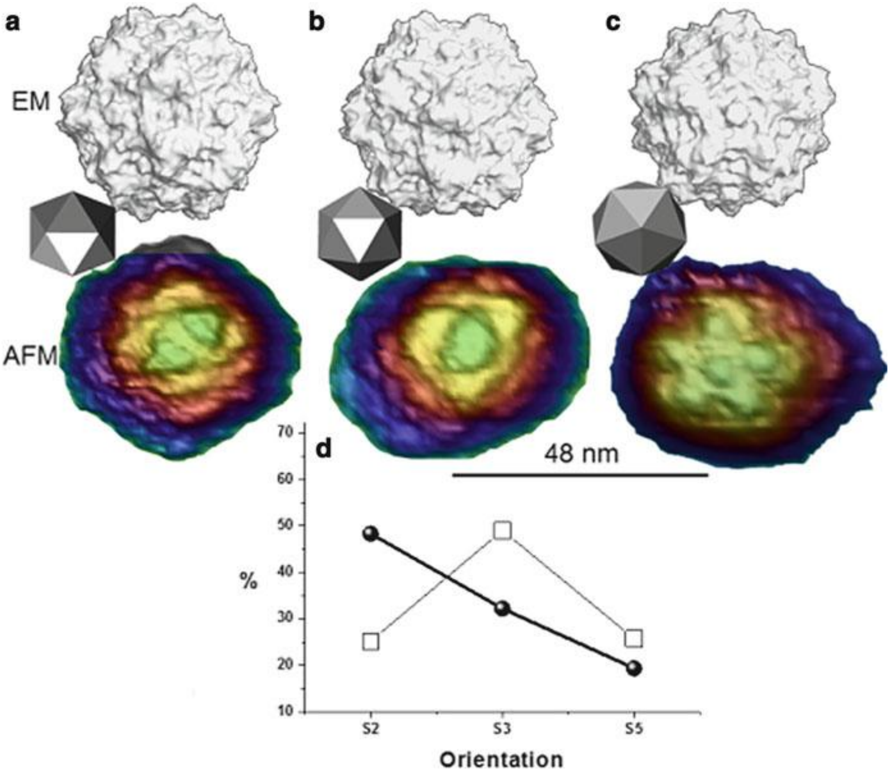


Fig. 8.10 AFM imaging of MVM. (**a**), (**b**) and (**c**) show the topographies corresponding to single MVM particles oriented with a twofold, threefold or fivefold symmetry axis on top, respectively. (**d**) represents the percentage of adsorption of each symmetry axis (*squares*) and the theoretical expectations if the adsorption through edges, faces and vertices of a perfect icosahedron were equiprobable

virus particles by resolving protrusions or other structural details allows not only a rather detailed characterization of their topography, but also the determination of local physical properties. For example, by making nanoindentations on single MVM particles adsorbed in different orientations and obtaining FZ curves, it has been shown that the viral single-stranded DNA genome is responsible for an anisotropic mechanical stiffening the virion relative to the DNA-free capsid [64], something that could be important for preserving viral infectivity. Moreover, it was possible to selectively disrupt these DNA-protein interactions or replace amino acid residues in selected regions of the MVM capsid and thereby engineer virus particles with altered mechanical properties [5] (see Chap. 18). Human adenovirus represents another example in which the orientation of single particles imaged by AFM could be determined on the adsorption of icosahedral viruses on a surface (Fig. 8.11). In this case, viruses attach to mica previously treated with $NiCl_2$ [28]. Comparison among

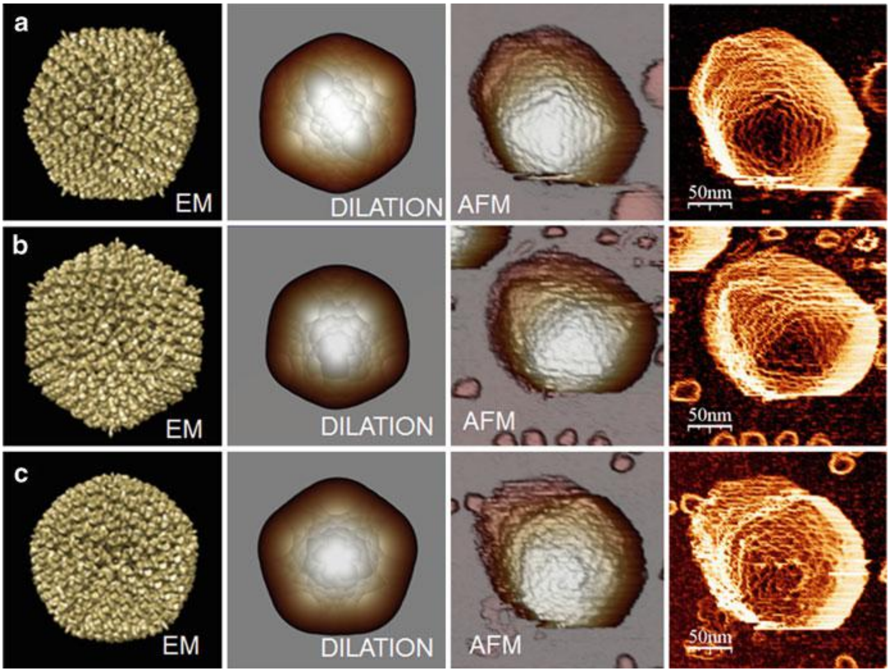


Fig. 8.11 AFM imaging of human adenovirus. (**a**), (**b**) and (**c**) show the topographies corresponding to single adenovirus particles oriented with a twofold, threefold or fivefold symmetry axis on top, respectively. AFM images are compared with EM and EM dilated structural models. The *right column* shows AFM topographic images that have been filtered to enhance the borders by obtaining the cosine of the angle between the normal vector of the surface and the normal direction of the paper sheet

EM data, EM geometrically dilated images and AFM images enabled a rapid identification of the particle orientation (see Figs. 8.11a–c for particles oriented along a twofold, threefold or fivefold symmetry axis, respectively).

It may be instructive to quantify the adsorption geometries found in MVM particles (Fig. 8.10d). Black circles represent the orientation likelihood of an icosahedron on a surface provided that the adsorption through vertices (fivefold symmetry), edges (twofold symmetry) and faces (threefold symmetry) is equally probable. If such were the case, a twofold symmetry orientation would exhibit the highest adsorption probability because in an icosahedron there are more edges (30) than faces (20) and vertices (12). On the other hand, empty squares depict the experimental results, which prime the threefold symmetry axis adsorption with almost a 50 % of the adsorbed particles. Since the threefold orientation maximizes the icosahedron area of contact, in contrast with edges or vertices, the particle-surface adsorption forces may depend strongly on the virus-surface contact area (Fig. 8.10d).

8.5 Understanding Viruses: Some Major Contributions of AFM

AFM enables new possibilities for studying virus particles, complementing classic structural approaches such as EM and X-ray crystallography. First, AFM may routinely work with individual viral particles. Second, they can be studied in liquid phase, in conditions which are closer to the *in vivo* environment. Although the resolution of AFM images is not comparable to that of structural models obtained by cryo-EM or X-ray crystallography, AFM enables the identification of virus elements which are not symmetrically ordered. Third, AFM may eventually allow the study in real-time of dynamic processes including the assembly, disassembly and conformational rearrangements of virus particles. Fourth, AFM has opened the possibility to investigate previously unexplored physical properties of virus particles, which may provide novel insights into the relationships between virus structure, properties, and functions. Along this line, AFM offers the possibility of probing the mechanical stiffness of virions and capsids and of selectively disrupting elements in individual virus particles (see Chap. 18). These experiments may contribute to understand virus assembly and disassembly, stability and dynamics.

8.6 Perspectives

Some of the current developments in AFM imaging aim to increase the data acquisition speed (a typical AFM image takes a few minutes to be acquired), in order to directly visualize the dynamics of relatively fast biological processes in real time, including conformational changes and the assembly or disassembly of virus particles. For example, AFM imaging at high video rates (80 ms/frame) has been already used to show conformational changes of single myosin V proteins on mica [65]. The methodology to achieve such a high speed AFM involves several technical developments, including the use of very soft cantilevers with a high frequency of resonance in liquids, which is achieved by decreasing the cantilever thickness and reducing the cantilever width and length proportionally.

Mechanical properties of biological objects are generally analyzed in indentation experiments [60]. However, there is also the possibility of using DM operating in liquids to extract information about mechanical properties of virus particles and other biological objects (*e.g*., their stiffness), by mapping the phase lag [66].

The development of non-invasive AFM imaging techniques where sample destruction is minimized is also important. For example, frequency modulation AFM [67, 68] is a dynamic technique where small forces of a few tens of pN only can be applied to the surface. This promising technique is based on the use of three simultaneous feedbacks. A phase lock loop ensures that the cantilever is always in

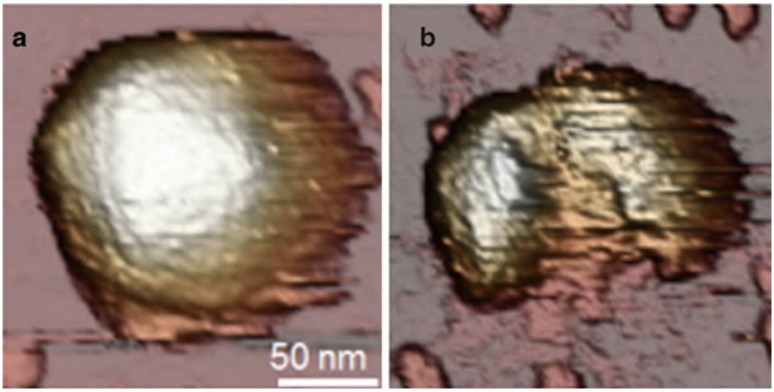


Fig. 8.12 Breaking a virus. (**a**) and (**b**) show a single human adenovirus particle before and after the breakage provoked by pushing the virus with the AFM tip beyond the rupture force in a force vs. z-piezo displacement experiment (Taken from Ref. [69]. Reproduced with permission)

resonance while a second feedback (working over the phase lock loop) changes the tip-sample gap to keep a set point frequency, such that its output gives the topography. Lastly, a third feedback is used to maintain the oscillation amplitude constant by changing the amplitude of the cantilever driving signal, which results in more stable operation.

Finally, it is important to stress the ability of AFM to mechanically manipulate and modify individual virus particles. The first experimental approaches, published very recently, are based upon the application of mechanical stress above the rupture force, thus provoking a rather uncontrolled dismantling of virus particles (Fig. 8.12) [53, 54, 69]. In the future the adequate control of the mechanical stress dosage may unveil further important insights into virus assembly, disassembly and dynamics [1].

In summary, the ongoing developments in AFM promise to improve its applicability in structural virology by: (i) further reducing unwanted sample damage; (ii) increasing the spatial resolution; (iii) increasing the temporal resolution so that relatively fast structural changes in virus particles can be followed in real time.

Acknowledgements I thank my students Aida Llauró-Portell, Alvaro Ortega-Esteban and Mercedes Hernando-Pérez, who are carrying out the hardest part of the work. I also want to thank my collaborators Nuria Verdaguer, Mauricio G. Mateu, José López Carrascosa, David Reguera, Julio Gómez-Herrero and Carmen San Martín. I acknowledge funding by grants from the Ministry of Science and Innovation of Spain, PIB2010US-00233, FIS2011-29493, Consolider CSD2010-00024, Comunidad de Madrid No. S2009/MAT-1467, and FIS2011-16090-E.

References and Further Reading

1. Ortega-Esteban A, Pérez-Berná AJ, Menéndez-Conejero R, Flint SJ, San Martín C, de Pablo PJ (2013) Monitoring dynamics of human adenovirus disassembly induced by mechanical fatigue. Sci Rep 3: 1434
2. Kuznetsov YG, Xiao CA, Sun SY, Raoult D, Rossmann M, McPherson A (2010) Atomic force microscopy investigation of the giant mimivirus. Virology 404:127–137
3. Roos WH, Radtke K, Kniesmeijer E, Geertsema H, Sodeik B, Wuite GJL (2009) Scaffold expulsion and genome packaging trigger stabilization of herpes simplex virus capsids. Proc Natl Acad Sci USA 106:9673–9678
4. Kuznetsov YG, Low A, Fan H, McPherson A (2004) Atomic force microscopy investigation of wild-type Moloney murine leukemia virus particles and virus particles lacking the envelope protein. Virology 323:189–196
5. Carrasco C, Castellanos M, de Pablo PJ, Mateu MG (2008) Manipulation of the mechanical properties of a virus by protein engineering. Proc Natl Acad Sci USA 105:4150–4155
6. Hernando-Pérez M, Pascual E, Carrasco C, Ionel A, Carrascosa JL, de Pablo PJ (2009) Study of the mechanical properties of bacteriophage T7. Biophys J 96(3):422a–423a
7. Binnig G, Rohrer H (1982) Scanning tunneling microscopy. Helv Physica Acta 55:726–735
8. Chen CJ (1993) Introduction to scanning tunneling microscopy. Oxford University Press, Oxford
9. Baro AM, Miranda R et al (1985) Determination of surface topography of biological specimens at high-resolution by scanning tunnelling microscopy. Nature 315:253–254
10. Binnig G, Quate CF, Gerber C (1986) Atomic force microscope. Phys Rev Lett 56:930–933
11. Schmalz G (1929) Uber Glatte und Ebenheit als physikalisches und physiologishes Problem. Verein Deutscher Ingenieure (Oct 12), pp. 1461–1467
12. Meyer G, Amer NM (1988) Novel optical approach to atomic force microscopy. Appl Phys Lett 53:1045–1047
13. Sader JE, Chon JWM, Mulvaney P (1999) Calibration of rectangular atomic force microscope cantilevers. Rev Sci Instrum 70:3967–3969
14. Butt HJ, Jaschke M (1995) Calculation of thermal noise in atomic-force microscopy. Nanotechnology 6:1–7
15. Israelachvili J (2002) Intermolecular and surface forces. Academic Press, London
16. Johnson KL (1985) Contact mechanics. Cambridge University Press, Cambridge
17. Ohnesorge F, Binnig G (1993) True atomic-resolution by atomic force microscopy through repulsive and attractive forces. Science 260:1451–1456
18. Giessibl FJ (1995) Atomic-resolution of the silicon (111)-(7x7) surface by atomic-force microscopy. Science 267:68–71
19. Sugimoto Y, Pou P, Abe M, Jelinek P, Perez R, Morita S, Custance O (2007) Chemical identification of individual surface atoms by atomic force microscopy. Nature 446:64–67
20. Carpick RW, Ogletree DF, Salmeron M (1997) Lateral stiffness: a new nanomechanical measurement for the determination of shear strengths with friction force microscopy. Appl Phys Lett 70:1548–1550
21. Martin Y, Williams CC, Wickramasinghe HK (1987) Atomic force microscope force mapping and profiling on a sub 100-a scale. J Appl Phys 61:4723–4729
22. Garcia R, Perez R (2002) Dynamic atomic force microscopy methods. Surf Sci Rep 47:197–301
23. Legleiter J, Park M, Cusick B, Kowalewski T (2006) Scanning probe acceleration microscopy (SPAM) in fluids: mapping mechanical properties of surfaces at the nanoscale. Proc Natl Acad Sci 103:4813–4818
24. Kuznetsov YG, Malkin AJ, Lucas RW, Plomp M, McPherson A (2001) Imaging of viruses by atomic force microscopy. J Gen Virol 82:2025–2034

25. Miyatani T, Horii M, Rosa A, Fujihira M, Marti O (1997) Mapping of electrical double-layer force between tip and sample surfaces in water with pulsed-force-mode atomic force microscopy. Appl Phys Lett 71:2632–2634
26. de Pablo PJ, Colchero J, Gomez-Herrero J, Baro AM (1998) Jumping mode scanning force microscopy. Appl Phys Lett 73:3300–3302
27. de Pablo PJ, Colchero J, Gomez-Herrero J, Baro AM, Schaefer DM, Howell S, Walsh B, Reifenberger R (1999) Adhesion maps using scanning force microscopy techniques. J Adhes 71:339–356
28. Ortega-Esteban A, Horcas I, Hernando-Pérez M, Ares P, Pérez-Berná AJ, San Martín C, Carrascosa JL, de Pablo PJ, Gómez-Herrero J (2012) Minimizing tip-sample forces in jumping mode atomic force microscopy in liquid. Ultramicroscopy 114:56–61
29. Villarrubia JS (1997) Algorithms for scanned probe microscope image simulation, surface reconstruction, and tip estimation. J Res Natl Inst Stan Technol 102:425–454
30. Marti O, Drake B, Gould S, Hansma PK (1988) Atomic resolution atomic force microscopy of graphite and the native oxide on silicon. J Vac Sci Technol a-Vac Surf Films 6:287–290
31. Soler JM, Baro AM, Garcia N, Rohrer H (1986) Interatomic forces in scanning tunneling microscopy—giant corrugations of the graphite surface. Phys Rev Lett 57:444–447
32. Hansma HG, Sinsheimer RL, Li MQ, Hansma PK (1992) Atomic force microscopy of single-stranded and double-stranded DNA. Nucleic Acids Res 20:3585–3590
33. Lyubchenko YL, Jacobs BL, Lindsay SM, Stasiak A (1995) Atomic-force microscopy of nucleoprotein complexes. Scanning Microscopy 9:705–727
34. Dame RT, Wyman C, Goosen N (2003) Insights into the regulation of transcription by scanning force microscopy. Journal of Microscopy-Oxford 212:244–253
35. Janicijevic A, Ristic D, Wyman C (2003) The molecular machines of DNA repair: scanning force microscopy analysis of their architecture. Journal of Microscopy-Oxford 212:264–272
36. Wagner P, Hegner M, Guntherodt HJ, Semenza G (1995) Formation and in-situ modification of monolayers chemisorbed on ultraflat template-stripped gold surfaces. Langmuir 11:3867–3875
37. Muller DJ, Amrein M, Engel A (1997) Adsorption of biological molecules to a solid support for scanning probe microscopy. J Struct Biol 119:172–188
38. Muller DJ, Janovjak H, Lehto T, Kuerschner L, Anderson K (2002) Observing structure, function and assembly of single proteins by AFM. Prog Biophys Mol Biol 79:1–43
39. Hansma HG, Pietrasanta L (1998) Atomic force microscopy and other scanning probe microscopies. Curr Opin Chem Biol 2:579–584
40. Horber JKH, Miles MJ (2003) Scanning probe evolution in biology. Science 302:1002–1005
41. Hinterdorfer P, Dufrene YF (2006) Detection and localization of single molecular recognition events using atomic force microscopy. Nat Meth 3:347–355
42. Neuman KC, Nagy A (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. Nat Meth 5:491–505
43. Kirmizis D, Logothetidis S (2010) Atomic force microscopy probing in the measurement of cell mechanics. Int J Nanomed 5:137–145
44. Kurland NE, Drira Z, Yadavalli VK (2012) Measurement of nanomechanical properties of biomolecules using atomic force microscopy. Micron 43:116–128
45. Viani MB, Pietrasanta LI, Thompson JB, Chand A, Gebeshuber IC, Kindt JH, Richter M, Hansma HG, Hansma PK (2000) Probing protein-protein interactions in real time. Nat Struct Biol 7:644–647
46. Moreno-Herrero F, Colchero J, Gomez-Herrero J, Baro AM (2004) Atomic force microscopy contact, tapping, and jumping modes for imaging biological samples in liquids. Phys Rev E 69:031915
47. Xu X, Carrasco C, de Pablo PJ, Gomez-Herrero J, Raman A (2008) Unmasking imaging forces on soft biological samples in liquids when using dynamic atomic force microscopy: a case study on viral capsids. Biophys J 95:2520–2528

48. Kasas S, Thomson NH, Smith BL, Hansma HG, Zhu X, Guthold M, Bustamante C, Kool ET, Kashlev M, Hansma PK (1997) Escherichia coli RNA polymerase activity observed using atomic force microscopy. Biochemistry 36:461–468
49. Moreno-Herrero F, de Jager M, Dekker NH, Kanaar R, Wyman C, Dekker C (2005) Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. Nature 437:440–443
50. Kodera N, Yamamoto D, Ishikawa R, Ando T (2010) Video imaging of walking myosin V by high-speed atomic force microscopy. Nature 468:72–76
51. Plomp M, Rice MK, Wagner EK, McPherson A, Malkin AJ (2002) Rapid visualization at high resolution of pathogens by atomic force microscopy – Structural studies of herpes simplex virus-1. Am J Pathol 160:1959–1966
52. Kuznetsov YG, McPherson A (2011) Atomic force microscopy in imaging of viruses and virus-infected cells. Microbiol Mol Biol Rev 75:268–285
53. Ivanovska IL, Miranda R, Carrascosa JL, Wuite GJL, Schmidt CF (2011) Discrete fracture patterns of virus shells reveal mechanical building blocks. Proc Natl Acad Sci USA 108:12611–12616
54. Castellanos M, Perez R, Carrillo PJP, Pablo PJ, Mateu MG (2012) Mechanical disassembly of single virus particles reveals kinetic intermediates predicted by theory. Biophys J 102:2615–2624
55. Sieben C, Kappel C, Zhu R, Wozniak A, Rankl C, Hinterdorfer P, Grubmüller H, Herrmann A (2012) Influenza virus binds its host cell using multiple dynamic interactions. Proc Natl Acad Sci USA 109:13626–13631
56. Xiao C, Kuznetsov YG, Sun S, Hafenstein SL, Kostyuchenko VA, Chipman PR, Suzan-Monti M, Raoult D, McPherson A, Rossmann MG (2009) Structural studies of the giant mimivirus. PLoS Biol 7:958–966
57. Vinckier A, Heyvaert I, Dhoore A, Mckittrick T, Vanhaesendonck C, Engelborghs Y, Hellemans L (1995) Immobilizing and imaging microtubules by atomic-force microscopy. Ultramicroscopy 57:337–343
58. Carrasco C, Luque A, Hernando-Pérez M, Miranda R, Carrascosa JL, Serena PA, de Ridder M, Raman A, Gómez-Herrero J, Schaap IA, Reguera D, de Pablo PJ (2011) Built-in mechanical stress in viral shells. Biophys J 100:1100–1108
59. Roos WH, Bruinsma R, Wuite GJL (2010) Physical virology. Nat Phys 6:733–743
60. Ivanovska IL, de Pablo PJ, Ibarra B, Sgalari G, MacKintosh FC, Carrascosa JL, Schmidt CF, Wuite GJ (2004) Bacteriophage capsids: Tough nanoshells with complex elastic properties. Proc Natl Acad Sci USA 101:7600–7605
61. Kienberger F, Zhu R, Moser R, Blaas D, Hinterdorfer P (2004) Monitoring RNA release from human rhinovirus by dynamic force microscopy. J Virol 78:3203–3209
62. Tang JH, Olson N, Jardine PJ, Grimes S, Anderson DL, Baker TS (2008) DNA poised for release in bacteriophage phi 29. Structure 16:935–943
63. Horcas I, Fernandez R, Gomez-Rodriguez JM, Colchero J, Gomez-Herrero J, Baro AM (2007) WSXM: A software for scanning probe microscopy and a tool for nanotechnology. Rev Sci Instrum 78:013705
64. Carrasco C, Carreira A, Schaap IAT, Serena PA, Gomez-Herrero J, Mateu MG, de Pablo PJ (2006) DNA-mediated anisotropic mechanical reinforcement of a virus. Proc Natl Acad Sci USA 103:13706–13711
65. Ando T, Kodera N, Takai E, Maruyama D, Saito K, Toda A (2001) A high-speed atomic force microscope for studying biological macromolecules. Proc Natl Acad Sci USA 98:12468–12472
66. Melcher J, Carrasco C, Xu X, Carrascosa JL, Gómez-Herrero J, de Pablo PJ, Raman A (2009) Origins of phase contrast in the atomic forces microscopy in liquids. Proc Nat Acad Sci 106:13655–13660
67. Hoogenboom BW, Hug HJ, Pellmont Y, Martin S, Frederix PLTM, Fotiadis D, Engel A (2006) Quantitative dynamic-mode scanning force microscopy in liquid. Appl Phys Lett 88:193109

68. Martinez-Martin D, Carrasco C, Hernando-Perez M, de Pablo PJ, Gomez-Herrero J, Perez R, Mateu MG, Carrascosa JL, Kiracofe D, Melcher J, Raman A (2012) Resolving structure and mechanical properties at the nanoscale of viruses with frequency modulation atomic force microscopy. PLoS One 7:e30204
69. Pérez-Berná AJ, Ortega-Esteban A, Menéndez-Conejero R, Winkler DC, Menéndez M, Steven AC, Flint SJ, de Pablo PJ, San Martín C (2012) The role of capsid maturation on adenovirus priming for sequential uncoating. J Biol Chem 287:31582–31595

Further Reading

Samorí P (ed) (2006) Scanning force microscopies beyond imaging. Wiley-VCH Weinheim, Germany

Baró A, Reifenberger R (2012) Atomic force microscopy in liquid. Wiley-VCH Weinheim, Germany

Also especially recommended for further reading are references [3, 4, 7, 21, 22, 45, 49] *listed above.*

Chapter 9
Optical Tweezers to Study Viruses

J. Ricardo Arias-Gonzalez

Abstract A virus is a complex molecular machine that propagates by channeling its genetic information from cell to cell. Unlike macroscopic engines, it operates in a nanoscopic world under continuous thermal agitation. Viruses have developed efficient passive and active strategies to pack and release nucleic acids. Some aspects of the dynamic behavior of viruses and their substrates can be studied using structural and biochemical techniques. Recently, physical techniques have been applied to dynamic studies of viruses in which their intrinsic mechanical activity can be measured directly. Optical tweezers are a technology that can be used to measure the force, torque and strain produced by molecular motors, as a function of time and at the single-molecule level. Thanks to this technique, some bacteriophages are now known to be powerful nanomachines; they exert force in the piconewton range and their motors work in a highly coordinated fashion for packaging the viral nucleic acid genome. Nucleic acids, whose elasticity and condensation behavior are inherently coupled to the viral packaging mechanisms, are also amenable to examination with optical tweezers. In this chapter, we provide a comprehensive analysis of this laser-based tool, its combination with imaging methods and its application to the study of viruses and viral molecules.

Keywords Biophysics • Virus • Bacteriophage • Capsid • DNA • RNA • Molecular motor • Machine • Single-molecule • Mechanochemistry • Optics • Optical tweezers • Magnetic tweezers • Dynamics • Manipulation • Force • Pressure • Elasticity • Condensation • DNA packaging

---

J.R. Arias-Gonzalez (✉)
Instituto Madrileño de Estudios Avanzados en Nanociencia (IMDEA Nanociencia),
c/Faraday 9, Campus de Cantoblanco, 28049 Madrid, Spain

Department of Macromolecular Structure, Centro Nacional de Biotecnología (CSIC),
c/Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain

CNB-CSIC-IMDEA Nanociencia associated unit "Unidad de Nanobiotecnología",
Campus de Cantoblanco, 28049 Madrid, Spain
e-mail: ricardo.arias@imdea.org

M.G. Mateu (ed.), *Structure and Physics of Viruses; An Integrated Textbook*,
Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_9,

© Springer Science+Business Media Dordrecht 2013

Abbreviations

3D	Three dimensional
AFM	Atomic force microscopy
bp	Base pair
Dig	Digoxigenin
ds	Double-stranded
GC	Guanine-cytosine
IR	Infrared
NA	Numerical aperture
α-Dig	Anti-digoxigenin

9.1 Introduction: Life Machinery at the Nanoscale

Optical tweezers are an instrument able to measure the forces and motion involved in the dynamics of biomolecular motors. From this viewpoint, they constitute a unique instrument to study the physics underlying some viral processes. In contrast to other structural techniques described in Chaps. 3, 4, 5, 7 and 8 that allow the visualization of virions and capsids, dynamic experiments using optical tweezers are unable to 'photograph' the main characters of this book, but can quantitatively track their mechanical activity in real time. This chapter will introduce a new method with which the virus is followed while it is at work, packaging its nucleic acid into the viral capsid. Characteristic results of these experiments are forces and packaging velocities as a function of time or as a function of the amount of packaged genome. A related problem is the mechanical behavior of nucleic acids; force-extension curves of DNA/RNA to determine their elasticity can be also obtained with this technique.

The machinery that makes up life is very small and highly condensed. The dimensions of a globular protein or the thickness of DNA are measured in nanometers. This machinery works in crowded environments (in a cell, the total concentration of macromolecules may reach 300–400 g/l), and is subject to continuous agitation. In these conditions, movement is strongly affected by thermal fluctuations (Brownian motion), and structures are markedly distorted by mechanical stress. The dynamic picture of the motion and activity of a protein or macromolecular complex is that of a boat sailing against furious winds and waves, with rapid changes in direction. Throughout evolution, proteins have developed strategies to rectify and bias thermal fluctuations using the energy from ATP or GTP in a way similar to the orientation of the boat's sail and rudder. These strategies are different from the macroscopic, human-made machinery, whose dynamic behavior more closely resembles a motorboat in a quiet sea, for which fuel energy is used to propel the boat in a chosen direction. Biological machinery is stochastic and soft, unlike macroscopic machines, whose movements are predictable and whose structures are typically rigid. Biological machinery is nonetheless highly efficient in terms of the ratio [developed work]/[spent energy], much more so than human technology to date. A further singularity of biomolecular machinery is its ability to interact with its environment. Biological machines are in essence biochemical entities; they change shape, size and mass as a result of their function. The action of most proteins and macromolecular assemblies, including viruses, is thus *mechanochemical*, and their complete study must include not only biochemical aspects, but also mechanical analyses (see also Chaps. 8 and 18).

Optical tweezers, described in this chapter, is a technique that allows such mechanochemical analysis. Viruses, like other large biological structures such as ribosomes, ATP synthase or the bacterial flagellum, are complex machines, and not passive structures. Many complex viruses in particular are coordinated assemblies of motors, structural scaffolds and information carriers. Many bacteriophages are fascinating from a technological point of view, because they are some of the most powerful machines at the nanoscale acting as packaging devices for their own nucleic acid. In turn, the nucleic acids represent soft ‘hard drives’ that store and spread information. Some nucleic acid physical properties, such as their elasticity and denaturation behavior, may have evolved in tune with the machinery that processes them, rendering these processes more efficient over the ages. Understanding DNA and RNA from a mechanochemical point of view may thus shed light on viral performance. This chapter will also be a platform to introduce these novel views of the viral machinery and its peripherals.

9.2 Basic Concepts and General Experimental Design

The forces developed by protein machines including viruses are in the piconewton (pN) range, which compares with one trillionth (10^{-12}) the weight of a lemon. Measuring these tiny forces was not technologically easy before the invention of the laser in the 1960s; use of this light technology to move and trap micro- and nanostructures began in the 1970s. By the mid-1980s, this application had been developed into a technique commonly known as optical or laser tweezers. Its use in molecular and cell biology became relevant by the 1990s, although experiments manipulating viruses, cells or organelles within living cells already saw the light in the previous decade [1].

The forces generated by electromagnetic radiation over a physical object are termed *optical forces*. The strength of these forces is generally in the piconewton range, although this depends on the material properties of the object, on the radiation power, and on beam shape. An instrument based on the use of optical forces will therefore be sensitive to the tiny forces exerted by molecular machinery. Since electromagnetic radiation can penetrate the medium without traceable effects, optical forces are a non-contact solution to measuring and exerting force

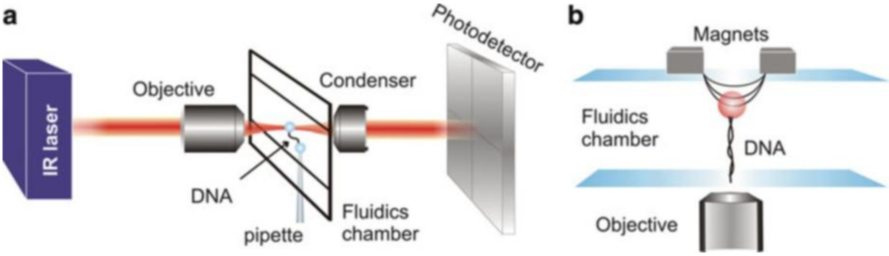


Fig. 9.1 Optical and magnetic tweezers. (**a**) Laser light is highly focused by a microscope objective and collected by a condenser lens. A bead or cell can be trapped in the focal region, which is produced inside a fluidics chamber. A micropipette inserted in the chamber can hold a second bead by suction. Micrometer-sized beads are typically used as mechanical handles to hold a biopolymer by both ends. The collected exiting rays are registered by a position-sensitive photodetector. The scheme shows a single-molecule experiment in which a DNA fiber is stretched between two beads. (**b**) The same experiment is performed using two small magnets that pull a magnetic microsphere inside a fluidics chamber. The DNA is held between the microsphere and the bottom coverslip. A microscope objective is used to monitor the experiment

on biological particles in their liquid environment and within living cells. Inertial and gravitational effects are insignificant in the timescales associated with the dynamics of most biological systems.

For reasons that will become clear later, a strongly focused laser beam generates optical forces in all directions towards the focal region, provided that these objects do not absorb light. The result is a three-dimensional (3D) *optical trap*. Most biological structures, natural or synthetic, do not substantially absorb visible or infrared (IR) radiation and can therefore be maneuvered by a strongly focused laser beam within these spectral bandwidths. If the light scattered from the optical trap is analyzed, the forces acting on the trapped object can be deduced. This is the basis of the optical tweezers, an instrument able to trap nanometer and micrometer-sized particles and to measure the external forces that act on them. The resulting displacements can be measured in the sub-nanometer range with millisecond time resolution.

A simplified version of an optical tweezers instrument is shown in Fig. 9.1 (in comparison to its magnetic counterpart, described in Sect. 9.3.7). The basic elements in the optical tweezers can be grouped into those common to an optical microscope and those related to the generation of the optical trap [2]. The *microscope* elements are an objective, a fluidics chamber and a condenser lens. The *trap* elements are a laser and a photodetector that registers the light at the exit of the condenser. To produce a strongly focused beam, a high-numerical aperture (NA) objective is required (water or oil immersion). Near-IR lasers are optimal for manipulating biological structures, because neither water nor biological structures strongly absorb radiation at this wavelength. Laser power in the focal region of several hundred milliwatts provides a good balance that prevents photodamage and produces forces up to a few hundreds of piconewtons. Magnetic tweezers use two small magnets to generate high field gradients in the vicinity of a paramagnetic

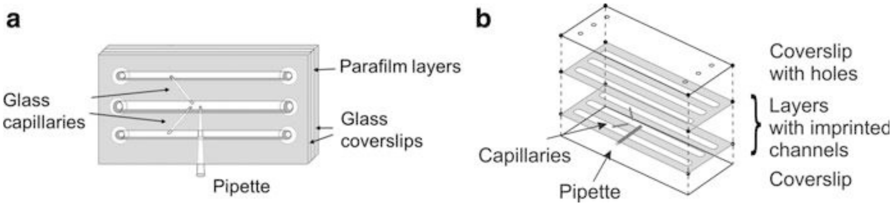


Fig. 9.2 Fluidics chamber. (**a**) The specimens (cells or beads) are passed sequentially through a system of channels. In this design, the experiment takes place in the central channel, which is connected to the *top* and the *bottom* channels by glass capillaries. The scheme also shows the disposition of a micropipette, useful in a variety of experiments. (**b**) Mounting a custom-made fluidics chamber for optical tweezers. After assembling the tubes and parafilm layers, the components are heat-sealed

bead, as shown in Fig. 9.1b. An objective behind the fluidics chamber is used to follow the experiment. This instrument is a good alternative for low forces and to generate torque, as we will see in Sect. 9.3.7.

The fluidics chamber is the test tube for mechanochemical assays. It is usually built using two microscope coverslips sandwiching parafilm layers with imprinted channels. The design shown in Fig. 9.2 allows specimens to be flowed through a top and a bottom channel connected with the central channel by glass capillaries. The optical trap is attained in the central channel. Buffers can be easily exchanged and flowed at a controlled rate, thus adding the ability to exert hydrodynamic pressure. A micropipette can be placed inside the chamber to suck beads and additional tubes can be used to dispense chemicals (for example, ATP) directly to the vicinity of the optical trap. These alternatives are very convenient in a variety of experiments that will be described later in this chapter. In contrast to most structural techniques, optical tweezers allow experiments in physiological conditions that can be changed easily and rapidly (for example, chemical concentration and pH) without affecting the performance of the instrument, which allow the detection in real time of processes triggered by such changes in conditions.

9.2.1 Single-Molecule Approach

As indicated by the above description, this method allows one-by-one trapping and manipulation of specimens (individual molecules and macromolecular assemblies, including virus particles). This ability is the basis of the *single-molecule* approach [3]. *Bulk* experiments, in which a large number of identical (but unavoidably unsynchronized) reactions take place, hinder access to signals from intermediate states of the specimens. Ensemble methods such as electron microscopy, X-ray crystallography or nuclear magnetic resonance (NMR) allow ‘taking snapshots’ of specific instants in the temporal flow of events for a certain type of molecule, but only the average state is eventually reconstructed from the data.

In addition, data and reconstructed states might be influenced by the cooperative disposition (periodic or non-periodic) of the ensemble before the experiment, yielding misleading structural information or masking rare but potentially important conformations. In fact, bulk experiments only allow observation of average dynamics. Single-molecule experiments permit the unraveling of real-time dynamics and conformations, because the signals are not confused by multiple, asynchronous events.

Another drawback of bulk experiments is that they tend to smear out the information from *thermal fluctuations*. The idealized dynamic picture envisaged from these experiments is a 'movie' at long timescales, in which molecular trajectories are smooth and well defined. At short timescales, this movie is very different: dynamics are fast and random (stochastic). As explained in the introduction, proteins and viruses dwell in a background of thermal agitation that affects their dynamic performance. This inherent dynamics can only be captured by real-time monitoring of individual processes. Fluctuations represent an integral part of protein function, and single-molecule information must therefore be included in the overall description of biomolecular processes. The average energy of a thermal fluctuation is k_BT, where k_B is the Boltzmann constant and T, the temperature. At 25 °C, this energy is $\sim 4.1 \cdot 10^{-21}$ J, a notable number if we consider the free energy released in the hydrolysis of one ATP molecule at physiological conditions, ~20–24 k_BT.

9.2.2 Molecular Machinery

Living in a thermal bath of violent agitations might seem to render any machinery inefficient. As pointed out in the introduction, however, biological machines have adapted to this apparently uncomfortable situation by developing a mechanism to extract energy from the thermal bath. This mechanism, the *Brownian ratchet*, consists roughly of rectifying random pushes (like a sailboat) using energy from ATP/GTP. In cases in which thermal pushes are not sufficient (for example, in large structures like the bacterial flagellum), protein machinery might bias diffusional steps by burning ATP as a *power stroke* (like a motorboat), a mechanism similar to that of macroscopic engines. In general, both mechanisms are combined by proteins in an intricate fashion to generate mechanical work [4].

9.2.3 Limitations of Optical Tweezers

Optical tweezers do not perform real-time microscopy, *i.e.*, structures cannot be resolved during the assay. Optical tweezers provide measurements of force, distance and time, but cannot ascertain the disposition of protein domains at a given moment or the structure of nucleic acids, for example, under a stretching force. The specimens are grossly considered as point-like or fiber-like structures. Combining fluorescence with optical tweezers (see perspectives in Sect. 9.6) can help determine the position and movement of a certain molecule, but in most cases the experimenter is ‘blind’: only the position of micrometer-sized structures, such as cells or the beads shown in Fig. 9.1, can be observed during the assay. In these conditions, structural and biochemical information must be known in advance if experiments are to be designed precisely and correctly interpreted.

9.3 Optical Tweezers

9.3.1 Optical Forces

Light can be viewed for many purposes as a flux of small particles called photons, which can interact with other objects and exchange energy as well as linear and angular momentum, thus generating forces and torques. This simple reasoning is the physical basis underlying optical forces: due to light-matter interaction, light incident on an object can push or modify the object’s trajectory in complex ways. These forces are very small, however, and we therefore do not feel the so-called radiation pressure from the Sun, for example, or from the light of a bulb. The fact that light interacts with matter nonetheless explains photodamage and mutations by ultraviolet or X-ray radiation at long exposure times and powers. These radiations are made up of very high energy photons which, rather than pushing biological matter, alter its structure at the atomic level. At visible and IR wavelengths, radiation is made up of photons with much less energy; in this case, the stream of coherent photons pushes small objects in a similar fashion as water exiting a hose at high pressure. This radiation pressure explains, for instance, why a comet’s dust tail points away from the Sun. Radiation pressure was long ago suggested as a propulsion method for spacecraft, by using thin mirrors as ‘solar sails’, although such sails would require areas of square kilometers.

Optical forces are therefore too tiny to be practical for the manipulation of macroscopic objects, but their effect is patent on atoms and on micrometer- and nanometer-sized objects. By adequately shaping the photon bundle from a laser source, it is possible not only to push tiny structures but also to trap them in 3D. The first optical trap was designed to manipulate atoms, which paved the way for the study of many interesting physical phenomena [5]. Trapping of small particles that scale to a few nanometers came soon afterwards, an interest that remains active for nanotechnological applications. Trapping of cells was then tested. Whereas light manipulation of small objects and atoms is useful for understanding light-matter interactions, however, the interest of optical traps in biochemical research resides in their use as force sensors, and as a method to mediate mechanochemical reactions.

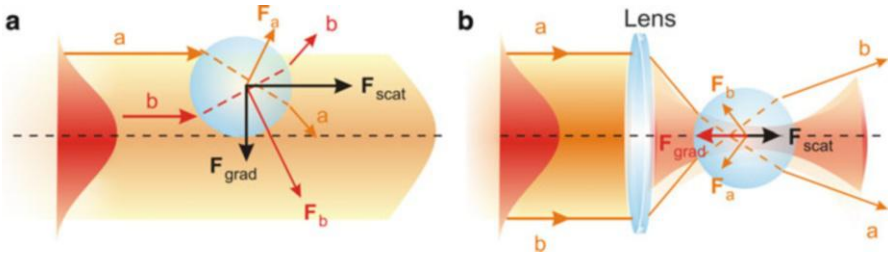


Fig. 9.3 Optical forces acting on a non-absorbing sphere. (**a**) Qualitative view of the forces exerted by a Gaussian laser beam on a sphere. The diagram illustrates the origin of the scattering (F_{scat}) and gradient (F_{grad}) components of the optical force. (**b**) Qualitative view of the forces exerted by a focused Gaussian beam on a sphere on-axis. The diagram shows the axial stability established by the gradient force along the centerline of the beam, towards the focal region. See text for details

9.3.2 An Optical Trap

Here we will analyze how to engineer the light from a laser source to confine an object in 3D. The internal structure of a complex biological particle cannot be resolved by visible light wavelengths; consequently, structure is not essential to understand the trapping analysis. We will assume that the particle is dielectric (*i.e.*, non-absorbing) and has a low refractive index (~1.5). These assumptions are common to biological structures including viruses and many synthetic particles. The main effect of light on this kind of object is refraction. The physics of trapping absorbing objects (metallic or semiconductor particles) or non-absorbing particles under resonant conditions are more complex, but to date they have not been exploited in biological experiments. In addition, absorbing and resonant light effects in biological structures are not desired if one wants to observe biological function without perturbation. These effects will therefore not be discussed here.

Figure 9.3a shows a spherical particle under the action of a laser beam whose wavelength is smaller than the size of the sphere. A ray of light can be conceived as a stream of photons, each carrying an associated momentum. The photons following the trajectories labeled "*a*" and "*b*" will be refracted (neglecting relatively minor surface reflections), as depicted, due to their interaction with the particle; as a result, their momentum will change. The momentum change experienced by each photon is transferred to the particle, thus generating the forces F_a and F_b. As more photons follow path "*b*" than "*a*" – because beam intensity is higher along the former path – force F_b is stronger than F_a. After decomposition of vectors F_a and F_b in two axes, one in the propagation direction of the laser and the other perpendicular to it, and extension of this ray analysis to all symmetrical pairs of rays, the net effect of laser light can be understood in terms of two forces: the *scattering force*, which pushes the particle along the incident beam, and the *gradient force*, which points towards the centerline of the beam. From the diagram, it is clear that there is transverse stability for a particle on-axis, since here $F_a = F_b$. To establish a 3D

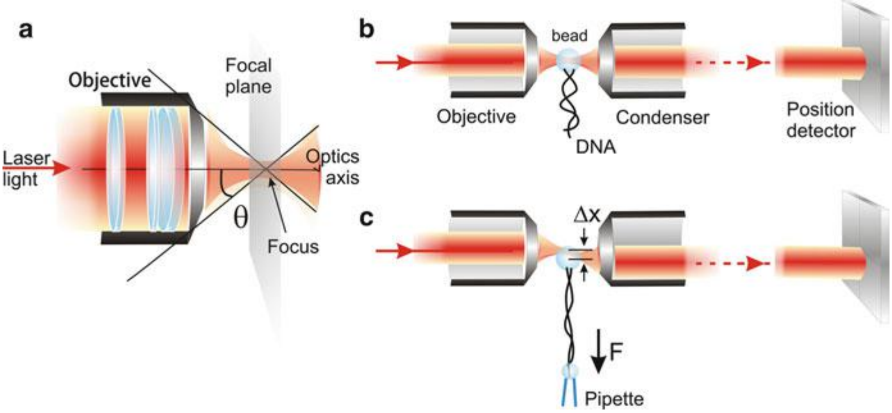


Fig. 9.4 Trapping and detecting force and displacements. (**a**) Representation of a microscope objective and the aperture angle, θ, of the cone of rays which are made to converge to a focus. In this configuration, laser light overfills the objective. (**b**) Diagram illustrating the configuration of the objective in front of a condenser lens, used to collect the light at the exit of the trap, and a photodetector to analyze the scattered light. Between the condenser lens and the photodetector, there are other optical elements not shown here for simplicity (see Fig. 9.5). (**c**) Diagram showing how an external force can be measured by registering light flux offsets in the photodetector. In (**b**) and (**c**), laser light underfills the objectives. Diagrams not to scale

trap, axial stability must be achieved, that is, the scattering force must be compensated as well. Arthur Ashkin first made this possible using two equal laser beams facing each other; a trap generated in this fashion is known as a *dual-beam trap*. Some years later, by focusing the laser beam with an objective lens, he showed that a single-beam trap is also possible (find advanced literature in [1]). Figure 9.3b shows the mechanical stability of a particle in such a trap: the focal region is now a global maximum of optical intensity and in these conditions, an axial gradient force pointing towards the focus is generated before and after the light passes through that region. If the focus is sufficiently strong, the gradient effect also dominates the axial stability thus resulting in the all-optical capture of the particle.

9.3.3 Optical Trapping and Manipulation of Micro-and Nanoparticles

The photon flux from a laser source describes approximately parallel, linear trajectories at the output. Using simple optical elements like lenses and mirrors, the photon flux can be guided to a microscope objective (Fig. 9.4a). The objective can be viewed as a lens that causes the parallel trajectories to converge on an ideal point called the focus. In this approximate description, we do not consider that photons traveling far from the center of the bundle interact with the edges of the objective. These photons will deviate from the linear, perfectly converging trajectories, hence will also slightly affect the trajectories of the remainder of the photons in the bundle. These effects are known more rigorously as *diffraction*; the result is that the photon bundle does not converge to a point, but to a focal volume with a size on the order of the light wavelength.

Laser light in fact behaves not only as a set of traveling particles, but also like a wave (*wave-particle duality*). It is actually an electromagnetic wave, whose electric and magnetic fields oscillate in both space and time. Laser light is characterized by its wavelength, which is a measure of the distance along which the electromagnetic wave describes a complete oscillation, and by its power in units of energy per transversal area. The intensity of the electromagnetic flux at the output of the laser is highest for the central photon trajectory and decreases rapidly for trajectories away from the center. This type of "light bundle" is called a Gaussian beam (Fig. 9.3a) and is common to most commercial lasers. Other names for these light bundles are TEM_{00} *beams*, and another term for the laser sources that produce them are *single mode lasers*. For optical trapping, it is common to use continuous wave (*CW*) lasers rather than pulsed lasers; in simple terms, they are switched on/off at high frequencies. The laser powers required for optical tweezers are of some hundreds of mW.

The microscope objective is actually a set of many lenses arranged to produce a high focusing effect (Fig. 9.4a). The focusing power is characterized by the numerical aperture (NA); for reasons that will be explained later, the higher the NA, the higher the trapping strength. Objectives are fabricated to work in air, water or oil, and the higher the refractive index of the immersion medium, the higher the NA of the objective. For optical tweezers applications, it is thus common to use water and oil immersion objectives. Let n be the refractive index of the immersion medium and θ, the angle of convergence of the rays after the objective (Fig. 9.4a). The NA is then defined by the equation $NA = nsin\theta$. The maximum numerical aperture, $NA = n$, is achieved when the rays are ideally made to converge on the focus at 90° angles (in practice, objectives cannot curve the rays that way). Water has a refractive index of 1.333 and usual oils, 1.51. Water immersion objectives normally have $NA = 1.2$ and oil immersion objectives can achieve $NA = 1.4$. As explained briefly in Sect. 9.2, the optical trap is attained in a fluidics chamber (Fig. 9.2), a test tube that will be described later to explain how to perform single-molecule experiments.

9.3.4 Force Measurement

If we only need to use the optical trap to manipulate objects, the key instrumentation described so far is sufficient. If we want to measure forces, however, we need to analyze the light scattered from the optical trap. The light exiting the trap is a diverging light ray bundle that can be collected by another objective or by a cost-effective condenser lens. After the condenser lens, the light is redirected to a position-sensitive photodetector. This type of photodetector measures power and registers light flux offsets, as depicted in Fig. 9.4b and c. An external force acting on the optically trapped particle deflects the exiting light cone and subsequently changes its projection after the condenser lens on the photodetector plane. Another method widely used in optical tweezers to measure bead displacement is back-focal plane interferometry, in which displacements are deduced from the interference pattern between the trapping laser and the scattered light (find advanced literature in [7]).

The use of modified position-sensitive photodetectors to register laser light intensity distribution is the basis for determining the externally applied force, through the momentum conservation principle [6]. This force calibration method has several advantages (it is independent of particle size, shape or refractive index, of buffer viscosity or refractive index, and of variations in laser power), although a simpler method is used in most optical tweezers design. This method is based on the fact that the external force, F, acting on the optically trapped objects is approximately proportional to the generated displacement, Δx (see Fig. 9.4c). In other words, $F = \kappa \Delta x$, where κ is the trap stiffness, a parameter that characterizes the 'invisible' spring that binds the particle to the optical trap. This parameter depends on the laser power and on the properties of both the particle and the immersion medium. Methods like the use of the Stokes law, or the Brownian motion of the bead in the trap can be used to accurately determine the trap stiffness, as described in-depth in other technical reviews and articles [2].

9.3.5 Optical Tweezers Instrument

The fact that the main elements of an optical microscope are involved in the design of a laser trap made the use of conventional microscopes the first and easiest strategy to set up an optical tweezers instrument. In Sect. 9.2, we divided the main components of the optical tweezers into those corresponding to the 'microscope' and to the 'trap' parts. The different elements are nonetheless optimally tuned by setting all components on an optical table with vibration isolation. Most optical tweezers instruments are homemade designed, with little use of purpose-designed commercial apparatus to date.

Experiments in the optical tweezers use microsphere position as the visual reference. Videomicroscopy can easily be integrated in optical trap design, because the objective that focuses the laser beam can also be used to generate real-time images of the events in the focal plane (Fig. 9.4a). A white-light lamp positioned on the condenser lens side is often used to illuminate the fluidics chamber interior in Köhler conditions, as in an optical microscope. A CCD camera at the back focal plane of the objective is used to visualize the experiment. Images from the camera and signals from the photodetectors are processed by a computer.

Figure 9.5 represents the component layout in an optical tweezers instrument. This scheme is not standardized, but integrates the general concepts and components of the instrument. In-depth design considerations are discussed in advanced reviews and technical articles [2, 6, 7]. In this scheme, two counter-propagating lasers

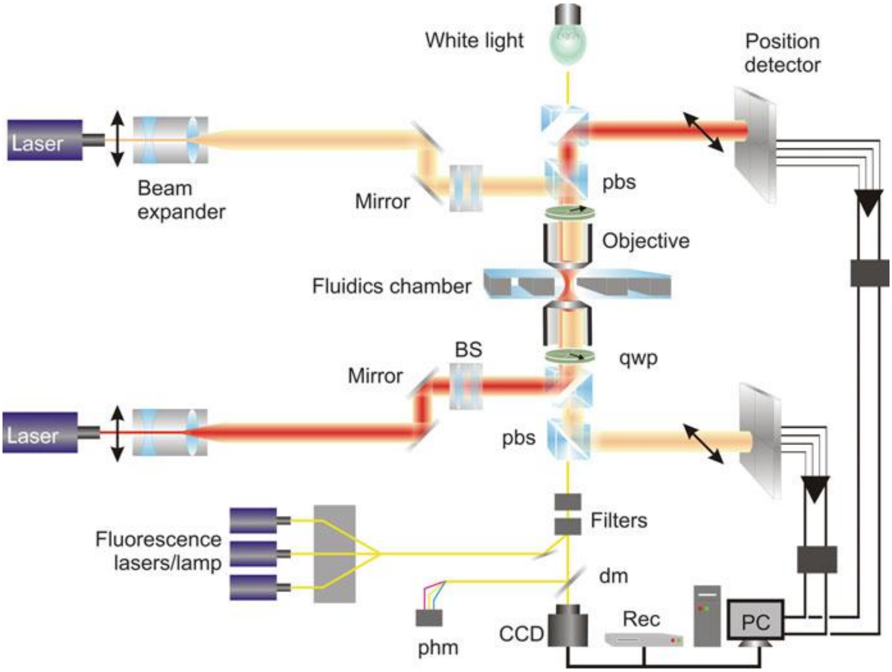


Fig. 9.5 Optical tweezers setup. Schematic summary of the most common designs (not to scale). The *arrows* show polarization of the beams at the exit of the laser sources and before the photodetectors. In this design, the polarizations of the beams in the trap are made circular and mutually orthogonal [6]. Additional light sources can be introduced in the setup to implement confocal microscopy and fluorescence. The laser wavelengths of the confocal module must be different from the infrared wavelengths used to generate the optical trap(s). *BS* beam steering, *pbs* polarizing beam splitter, *qwp* quarter wave plate, *dm* dichroic mirror, *Rec* video recorder, *phm* photomultiplier

converge on the fluidics chamber, for which two objectives are used instead of a pair objective-condenser lenses. If the two lasers converge on the same focus, the instrument is a *dual-beam optical tweezers*. In this design, a micropipette is used to grab a second bead, as in the single-beam optical tweezers described above. If the two lasers are brought to separate focuses, each could be used to grab an independent bead without the need for a micropipette; the instrument so designed is a *dual-trap optical tweezers*. The equipment can also be engineered to split the light from a single laser source into two optical paths, a strategy that increases resolution [8].

In contrast to single-beam optical tweezers, the lasers in dual-beam instruments do not need to overfill the objectives to generate an efficient trap (see Fig. 9.4). Underfilling reduces energy concentration in the focal region, which in turn decreases radiative effects on the trapped specimen. This situation is also exploited to calibrate the optical tweezers by direct measurement of light momentum, as indicated earlier.

When two beams share the same trajectory, it is of paramount importance to avoid interferences. Electromagnetic waves can interfere, generating spurious effects in the optical trap that could lead to undesired consequences in manipulation and measurement. To prevent this, beams from distinct sources should encounter each other with orthogonal polarizations. Polarization is a property of light waves that describes the direction of oscillation of the electric and magnetic fields. Light from a laser source is often linearly polarized, as shown in Fig. 9.5. Quarter-wave plates can rotate the polarization of each laser beam immediately before it enters the microscope objectives, so that they do not interfere with each other in the focal region. Polarizing beam splitters are optical elements that transmit or reflect the light depending on its polarization. These elements can be used to guide the two beams before and after the objectives (Fig. 9.5).

The two beams can be steered using motorized elements, such as mirrors and lenses, which can be computer-controlled. These elements allow moving one trap relative to the other in dual-trap optical tweezers. Beam expanders are also used to adjust the area of the laser spot at the exit of the laser sources to the entrance aperture of the objectives. They are used to over- or underfill the objectives, depending on the optical tweezers design. Beam expanders are normally made up of two lenses of different focal length. A pinhole can be positioned between them to filter the beam spatially. Astigmatism lenses can be introduced in the laser paths to correct this aberration (not shown in Fig. 9.5).

Finally, it is important to say a word about imaging possibilities. Unless fluorescence imaging is used, only the position of microscopic objects will allow monitoring experiments in the optical tweezers. Cells can be tracked easily, as can large organelles like centrosomes [9], but molecules cannot. Experiments with proteins and nucleic acids thus involve their attachment to microspheres in the so-called molecular construct (see Sect. 9.4.2). These microspheres, generally made of polystyrene or silica, act as mechanical handles to tether the specimen at specific sites. Correct design of the molecular construct and as much structural information as possible is necessary to perform and interpret the assays correctly. Single-molecule fluorescence, including the use of quantum dots,[1] and confocal microscopy[2] can be integrated in the optical tweezers; these are of great help to provide dynamic information on the mechanochemical experiments (see Sect. 9.6.3). Figure 9.5 illustrates one way of combining fluorescence in the optical tweezers setup.

---

[1] *Quantum dots* are luminescent semiconductor nanocrystals which are used as a new class of fluorescent probes. In comparison with organic dyes and fluorescent proteins, they are brighter and exhibit both greater photostability and longer lifetimes. They can also be tuned to specific wavelengths with narrower bandwidths than fluorescent molecules.

[2] *Confocal microscopy* is a high-resolution imaging technique based on a point-by-point illumination of the sample. The light source is often a focused laser beam. Lasers of several wavelengths and filters can be used to combine this technique with fluorescent imaging.

9.3.6 Electromagnetic Theory

The forces that generate a laser source are electromagnetic. In this section, we present a rigorous theory to explain the origin of these interactions on non-absorbing nanoparticles. This section will help to better understand the concepts presented on optical forces in Sect. 9.3.2. The following is not essential for performing experiments or understanding the instrumental description above, and can therefore be skipped.

A particle is small when its characteristic diameter is much smaller than the illumination wavelength. Laser wavelengths for optical tweezers purposes are in the near-IR range, usually from 800 to 1,100 nm. Longer wavelengths are strongly absorbed by water and generate undesired heating and convection effects. A particle is then small (also known as a Rayleigh or dipolar particle) when its diameter is below ~100 nm. The Lorentz force in the presence of an electromagnetic field is [10]:

$$\vec{f} = (\vec{\wp}\cdot\nabla)\vec{\varepsilon} + \frac{\partial\vec{\wp}}{\partial t}\times\vec{\beta} \tag{9.1}$$

where $\vec{\wp}$ is the induced electric dipole moment and $\vec{\varepsilon}, \vec{\beta}$ are the electric and magnetic vectors, respectively (SI units). Experiments in optical tweezers are performed in liquid media and particle motion is therefore greatly slowed by viscous forces. In these conditions, the particles cannot follow the rapid oscillations of the electromagnetic fields; consequently, the time-average force is the observed value. The electromagnetic field from a single mode laser is harmonic and can thus be expressed in complex representation as $\vec{\varepsilon}(\vec{r},t) = \Re e\{\vec{E}(\vec{r})exp(-i\omega t)\}$ and $\vec{\beta}(\vec{r},t) = \Re\mathrm{e}\{\vec{B}(\vec{r})\exp(-i\omega t)\}$, and the dipolar moment as $\vec{\wp}(\vec{r},t) = \Re e\{\vec{p}(\vec{r})\exp(-i\omega t)\}$; $\vec{E}, \vec{B}$ y $\vec{p}$ are, in this formalism, complex functions of position. $\Re e$ denotes the real part and ω is the angular frequency of the field. The i-th component of the time-average force is thus (see [11] and references therein):

$$\langle f_i(\vec{r},t)\rangle = F_i(\vec{r}) = \frac{1}{2}\Re e\left\{\alpha\sum_{j=x,y,z}E_j(\vec{r})\frac{\partial E_j^*(\vec{r})}{\partial x_i}\right\}, i = x, y, z \tag{9.2}$$

where * denotes the complex conjugate. Equation (9.2) is the expression of the observed optical force, $\vec{F} = \left(F_x, F_y, F_z\right)$, over a small particle with complex electric polarizability α in the liquid medium.

The electric field can be expressed in paraxial representation, that is, a beam or evanescent field with a main propagating direction along $\vec{k}$, as $\vec{E}(\vec{r}) = \vec{E}_0(\vec{r})$ $\exp\left(i\vec{k}\,.\,\vec{r}\right)$. The modulus of vector $\vec{k}$ is given by $k = n(\omega/c)$, being n the refractive index of the liquid medium and c the speed of light in vacuum. Substituting this expression into Eq. (9.2), one obtains [11]:

$$\vec{F} = \frac{1}{2}\vec{k}\Im m\{\alpha\}\left|\vec{E}_0\right|^2 + \frac{1}{4}\Re e\{\alpha\}\nabla\left|\vec{E}_0\right|^2 - \frac{1}{2}\Im m\{\alpha\}\Im m\left\{\left(\vec{E}_0\cdot\nabla\right)\vec{E}_0^*\right\}, \qquad (9.3)$$

where $\Im m$ denotes the imaginary part. The first term in Eq. (9.3) is proportional to the field intensity, $\left|\vec{E}_0\right|^2$, and points along the propagating direction, $\vec{k}$. This term therefore represents the contribution of the radiation pressure to the scattering force. Since $\Im m\{\alpha\} > 0$, this force pushes the particle in the laser-propagating direction. The second term is proportional to the gradient of the field intensity, $\nabla\left|\vec{E}_0\right|^2$, and it is thus known as the gradient force. Since $\Re e\{\alpha\} > 0$ for non-absorbing particles, this force attracts the particle towards the maximum of the field intensity. This term is responsible for optical trapping, as was also explained qualitatively in Sect. 9.3.2. The last term is not significant in typical experiments but may induce rotation of the trapped particle due to spin angular momentum exchange between the field and the particle (see Sect. 9.6.2).

Optical forces on many viral particles can be calculated by using Eq. (9.3). Beads used as mechanical handles are usually in the micrometer range and therefore beyond this approximation. The above description nonetheless allows understanding the physics of the problem, which was previously treated with a ray optics description in Sect. 9.3.2. The force on these larger particles can be calculated using the following expression [10, 11]:

$$\begin{aligned}\vec{F} =& \frac{1}{2}\Re e\left\{\int_{\Sigma} d^2r\left[\left(\in \vec{E}(\vec{r},\omega)\cdot\vec{n}\right)\vec{E}^*(\vec{r},\omega) + \left(\mu\vec{H}(\vec{r},\omega)\cdot\vec{n}\right)\vec{H}^*(\vec{r},\omega)\right]\right.\\ &\left. -\frac{1}{2}\left(\in\left|\vec{E}(\vec{r},\omega)\right|^2\right) + \left(\mu|H(\vec{r},\omega)|^2\right)\vec{n}\right\} \qquad (9.4)\end{aligned}$$

where Σ is a surface enclosing the particle and $\vec{n}$ is its local outward normal unit vector. $\in$ and μ are the electric and magnetic relative permittivities of the medium, respectively. The complex representation of the H-field, $\vec{\aleph}(\vec{r},t) = \Re e\left\{\vec{H}(\vec{r})\exp(-i\omega t)\right\}$, is related to the complex B-field by $\vec{B}(\vec{r}) = \mu\vec{H}(\vec{r})$. Equation (9.4) uses the time-harmonic expression of Maxwell's stress tensor. Computation of forces from this expression involves numerical calculations in most cases. This might be useful to predict the behavior of special particles in the trap but for most biological experiments, it is often only required to characterize empirically the field-particle interaction by a spring constant (see Sect. 9.3.4).

9.3.7 Magnetic Tweezers

Magnetic tweezers are a cost-effective instrument to perform single-molecule experiments with applications similar to those of optical tweezers [12]. This instrument can be used to measure and exert forces in mechanochemical reactions, although to date it has been little used to study viruses and viral components, with the exception of their nucleic acids. To provide a broad palette of possibilities for virus study, we present this technique briefly.

Typical magnetic tweezers setups (Fig. 9.1b) use permanent magnets to generate a field gradient inside a fluidics chamber. By moving these magnets along the field gradient, forces on magnetic particles in a fluidics chamber can be increased/decreased; by rotating them, torque can be induced. Electromagnets can also be used to generate stretching forces. In this case, adjusting the current running through the magnetic coils allows change in the force strength. Superparamagnetic beads are used as mechanical handles for biological molecules in these experiments. Due to space constraints, magnets must be placed outside the fluidics chamber, which prevents the highest field gradients from approaching the magnetic beads, thus establishing an upper boundary to the forces generated. Maximum forces on micrometer-sized superparamagnetic particles in these conditions are usually below ~10 pN. Recent developments using microscopic magnetic poles inside the fluidics chamber can greatly increase the forces, to a range similar to that of optical tweezers [13]. Minimum generated forces are in the range of 0.01 pN.

The force generated by two magnets on a superparamagnetic bead is

$$\vec{F} = \frac{1}{2}\nabla\left(\vec{m}\cdot\vec{\beta}\right), \tag{9.5}$$

where $\vec{m}$ is the induced magnetic moment on the microsphere in the presence of the magnetic field, $\vec{\beta}$, arising from the magnets. Force calibration can be performed against the viscous drag or using Brownian fluctuations. A microscope objective and a camera below the liquid chamber are used to follow the experiments, specifically, to deduce the forces and displacements of the bead. Illumination comes from above the fluidics chamber in this design (Fig. 9.1b). This apparatus can also be built on an inverted microscope.

Although this instrument constitutes the magnetic analogue of optical tweezers, it is not based on the generation of a trap. In other words, particles are not grabbed but dragged. To be manipulated, single particles must be tethered. For example, the DNA molecule in the experiment in Fig. 9.1b is always subject to a pulling force. In addition, the capacity of magnetic tweezers to manipulate a particle in 3D is limited. The use of more than two magnetic poles increases maneuverability, as shown in recent reports [13], but 3D control of the particle can remain a problem for many experiments.

Particles that do not show a magnetic response cannot be manipulated directly with magnetic fields. Cells and organelles, for example, cannot be dragged with magnetic tweezers. While this is a limitation for *in vitro* experiments, this fact is being exploited in the manipulation of magnetic nanoparticles inside a cell with high specificity [13].

Magnetic and optical tweezers, together with atomic force microscopy (AFM) (Chap. 8), cover a range of forces from 10^{-2} to 10^{5} pN for single molecule manipulation. These techniques thus provide modern means to look into biochemical reactions in new ways. In addition to the research scope and experimental conditions, spatial and temporal resolution issues should be considered to aid in making the most appropriate choices [14].

9.4 Operation

Forces on biological specimens of a few nanometers are very small, and it is thus difficult to generate a stable trap in liquid media. Thermal fluctuations and instrument noise make trapping efficiency very low and residence times very short. It is also difficult to control specimen orientation and to assure that a single specimen is in the trap. The methodology for manipulating nucleic acids, viruses and proteins therefore involves the use of beads as mechanical handles. We will now explain general strategies for performing experiments in the optical tweezers.

9.4.1 Methodology

Most experiments in the optical tweezers require that the molecular system under study be tethered to at least two specific sites. For this purpose, a second bead must be introduced in the experiments, either by suction by a micropipette (shown in Fig. 9.4c, bottom), or using a second optical trap. As stated above, these two beads are mechanical handles to biochemically link, for example, DNA molecules by two ends, or more complex systems such as a virus particle and the viral DNA to be packaged (see below). To exert force, one of the beads must move relative to the other (Fig. 9.4c). In optical designs in which the second bead is held by suction, the micropipette is moved by a computer-controlled stage. In designs in which this micropipette is replaced by another optical trap, the second trap can be displaced by steering one laser beam relative to the other with external mirrors and lenses (Fig. 9.5).

In DNA stretching experiments, the molecule is often modified at one end to insert biotin labels and at the other to insert digoxigenin (Dig) labels. The molecule is then tethered between two beads, one coated with streptavidin to provide biotin-streptavidin linkage and the other with anti-digoxigenin (α-Dig) to provide a Dig-α-Dig linkage. The experiment is followed by videomicroscopy, observing the position of the two beads. DNA extension is measured by recording the relative positions of the two beads in real time. In setups in which a micropipette is used, a low-power laser can be introduced to track the position of the second bead; light scattered from the bead can be detected in another photodetector. Alternatively, if the micropipette rigidly attached to the microchamber is moved relative to the

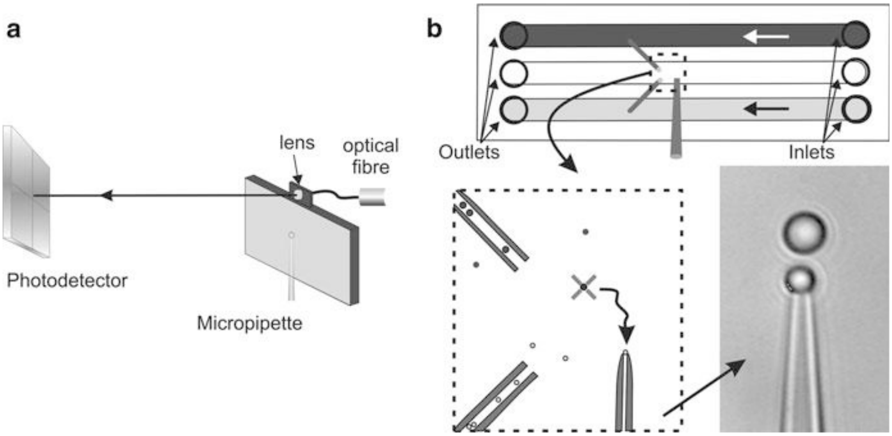


Fig. 9.6 Optical tweezers methodology. (**a**) Use of a light-lever device to track the position of the chamber with the rigidly attached micropipette. (**b**) *Top*: a three-channel microfluidics chamber is filled with buffer. Two suspensions of different types of microspheres are then flowed into the top and bottom channels through inlet holes; the fluid exits the channels through outlet holes. *Bottom*, *left*: microspheres from the top or bottom channel are optically trapped at the exit of the dispenser tubes and deposited on the micropipette in the middle of the central channel. *Bottom*, *right*: micrograph as seen by videomicroscopy of a ~3-μm diameter bead in the optical trap and a ~2-μm diameter bead held by the micropipette

optical trap (Fig. 9.6a), a so-called light-lever device can be used [6]. This device comprises a single-mode laser and a lens attached to the pipette/chamber frame as in Fig. 9.6a. The lens collimates light from the laser and directs it onto a photodetector. A similar light-lever configuration is used in AFM to detect the position of the cantilever in reflection (see Chap. 8).

The procedure for a DNA stretching experiment in the fluidics chamber is shown in Fig. 9.6b. The optical trap is first used to grab a single streptavidin-coated microsphere from one of the channels and to place it on the micropipette. An α-Dig-coated bead with attached DNA fibers is then trapped from the other channel and brought toward the bead on the micropipette until a stable biotin-streptavidin bond is formed.

9.4.2 Molecular Construct

The design of the molecular construct can be a bottleneck in single-molecule optical tweezers experiments. Good design consists of biochemically engineering the specimen for attachment to mechanical handles. Force is a vector magnitude and therefore marks a precise coordinate in the experiments, unlike other parameters such as temperature or pH. It is important to tether the specimen correctly from the domains that will produce mechanical work or will extend.

To study the elasticity of a polymer, for example DNA, the molecular construct is very simple and consists of tethering the molecule from opposite ends, as explained above. If one wants to unzip double-stranded (ds) DNA to study mechanically induced melting, the molecule must be tethered on opposite strands of the same end. When proteins are involved, structural data are needed to determine the sites expected to unfold. In experiments with molecular motors, structural data and functional models are desirable, to establish the domains involved in motion and the direction the molecule is expected to translate or rotate.

The design of the molecular construct can be seen as the analogue of the crystallization procedure in X-ray diffraction (Chap. 4): proteins need different approaches for their structural analysis with this technique and established strategies help, but do not constitute a standard for new experiments. Much of the success in single-molecule experiments relies on correct construct design. Figure 9.7 shows a molecular construct devised to study bacteriophage packaging by *Bacillus subtilis* phage Φ29. Phage capsids with prepackaged 6.6-μm-long dsDNA were bound directly to a bead on a micropipette, tethering the free DNA end to an optically trapped bead [17]. We will analyze later in this chapter how this construct is used to study the Φ29 packaging mechanism.

9.4.3 Operation Modes

There are four basic operation methods: *force-extension*, *constant-force*, *passive* and *force-jump* modes. The first is used to stretch polymers for analysis of their elastic behavior; Fig. 9.7 shows the force-extension curve of a DNA molecule in this mode. Polymer extension is observed as the force increases continuously; for this reason, this mode is also known as force-ramp.

The constant-force or force-feedback mode is used to analyze the activity of a protein machine as a function of a constant load. In this mode, the position of the optical trap relative to the second attachment point is adjusted to maintain constant tension on the bead. The protein kinesin has been tested using this strategy, as has packaging activity of Φ29 (Fig. 9.7) (find particular examples in [1]). Application of distinct constant forces allows observation of varying responses of the specimen.

The passive mode consists of fixing the position of the micropipette and/or traps and letting the specimen work. In this mode, the distance between the beads is not necessarily constant, as the position of the trapped bead undergoes small displacements relative to its resting position, resulting from the activity of the tethered specimen. This mode can be used to the study of DNA polymerization, kinesin transport and viral packaging, and allows measurement of maximum forces developed by protein motors, also known as stall forces.

In force-jump mode, the applied force is changed rapidly in discrete force steps rather than continuously (characteristic of the ramp-mode). This mode is useful for studying kinetics between several molecular states that are stable at different forces.

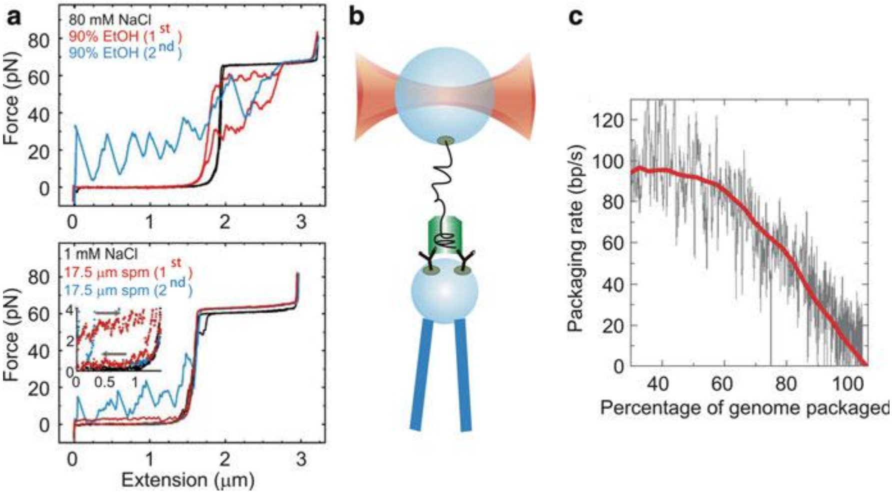


Fig. 9.7 Some applications of optical tweezers. (**a**) DNA stretching experiments. In each graph, a single molecule of dsDNA is first stretched and relaxed in physiological buffer (*black curve*), followed by two successive stretch/relax cycles (*red* and *blue curves*) in a buffer-90 % ethanol mixture (*top*) or in buffer with a low concentration of spermine (*bottom*) (Reproduced from [15]. With permission). (**b**) Molecular construct to study Φ29 packaging activity. The bead in the optical trap is streptavidin-coated for attachment to the biotin-labeled free end of the DNA. The virus is attached to a protein G-coated bead held by a micropipette. The DNA-capsid complex is captured in the tweezers in a stalled, partially prepackaged preparation before the start of the assay. (Current strategies are not restricted to using prepackaged molecular constructs [16]) (**c**) Packaging rate *versus* the amount of DNA packaged, relative to the original 19.3-kbp Φ29 genome. *Grey line*, trace for a single complex; *red line*, average of eight such measurements (Reproduced from [17] with permission)

For instance, protein folding or RNA-hairpin open/close conformations can be monitored through changes in the molecular extension with this operation mode [18].

9.5 Understanding Viruses: Some Major Contributions of Optical Tweezers

In contrast to the techniques studied in previous chapters (Chaps. 3, 4, 5, 6, 7 and 8), viruses have been little studied with optical tweezers so far. The reason is the novelty of the single-molecule approaches that involve mechanical manipulation of biological specimens. There is currently much interest not only in characterizing viral mechanochemical functions in real time, but also in using this machinery for nanotechnological applications. In this section, we present some major and illustrative contributions provided by optical tweezers in virology.

9.5.1 *Elasticity of Nucleic Acids and Relevance for Packaging in Viruses*

With the advent of optical tweezers in molecular biology, DNA and RNA in their double- or single-stranded configurations were promptly studied [19]. The first assays in the 1990s elucidated the elasticity of these information carriers. These studies are important for understanding the energetic barriers and the function of proteins and molecular complexes that act on nucleic acids in processes such as replication, transcription, unwinding or packaging. From the point of view of structural and physical virology, conformational dynamics of nucleic acids is important, for example, for understanding nucleic acid packaging and storage inside viral capsids, from passive encapsulation into preassembled viral capsids to the active packaging process in bacteriophages (see also Chap. 12).

A molecular construct for mechanical analysis of DNA is shown in Fig. 9.4c. Figure 9.7a shows the results of a standard experiment in which individual dsDNA molecules are stretched and then relaxed in different conditions. For these experiments, DNA was obtained from a linear fragment of plasmid piJ702 from *Streptomyces*, whose makeup consists of 70 % guanine-cytosine (GC) base pairs (bp). Each molecule was first stretched and relaxed in physiological conditions (aqueous buffer with 80 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, pH 7.5) and subsequently stretched and relaxed twice in a buffer-90 % ethanol mixture (Fig. 9.7a, top) or in the same buffer with 1 mM NaCl and 17.5 μM spermine (Fig. 9.7a, bottom). The ability of optical tweezers to deal with different conditions in the same assay is demonstrated here. Ethanol was used to mimic the behavior of DNA at low humidity, an environment similar to that inside viral capsids; spermine, a polycation involved in genome condensation, was used to study DNA organization in free form, not constrained by, for example, encapsulation in viral proheads.

The force-extension curves in Fig. 9.7a in physiological buffer (black trace) show the characteristic elasticity regimes of DNA: the entropic regime, in which the molecule behaves as an entropic spring, and the intrinsic regime, in which it is stretched beyond its contour length. At 65 pN, the molecule overstretches as a result of the almost total unwinding it experiences at high stress [19]. Force-extension curves of DNA in the ethanol mixture or with spermine show either a plateau at entropic forces of ~3 pN or a stick-release pattern. These effects indicate the unraveling of collapsed structures, such as coiled DNA or toroidal supercoiling, in the form of a smooth decondensation process or a violent unfolding of the polymer (see [15] and references therein).

The mechanical properties of DNA, as characterized by its persistence length and stretch modulus of the molecules, have been also studied as a function of the GC composition or of external conditions such as pH, ionic strength, temperature and humidity level (find advanced material in [1, 20–22]). Optical tweezers have also been used in more varied configurations to study DNA unzipping (the mechanical separation of the two strands), DNA twisting (with optical and magnetic tweezers), and the opening of RNA hairpins [18, 19]. The influence of several

DNA-binding proteins on the mechanical properties of the complex has also been analyzed [23].

In conclusion, from the point of view of viral nucleic acid storage and packaging in viral capsids, these and other single-molecule manipulation studies have contributed to the understanding of the elastic, electrostatic and entropic penalties of confining and organizing genetic material at high pressure in a small, closed space. These studies have also shown that nucleic acid structure is dependent on mechanical tension; as indicated, a few piconewtons are sufficient to straighten a dsDNA molecule; some tens of piconewtons are sufficient to promote base-pair unstacking and twist changes in physiological conditions, both of which are mechanical denaturation effects that lead to geometric changes in the DNA double helix. Forces in these ranges are developed by the bacteriophage portal machinery, which will be reviewed in the next section. A dynamic description of viral packaging in the presence of mechanical stress must therefore include the tension-induced conformational alterations of the viral nucleic acid that the portal machinery must process.

9.5.2 Packaging Activity of Bacteriophages

Bacteriophages are the most abundant organisms on earth (see Chap. 17). Nonetheless, very few virologists have analyzed their machinery at the single particle level [24–26]. The pioneering work of Smith and Bustamante's team in 2001 reported the direct activity of this impressive, powerful machinery [17]. These results have complemented structural and biochemical studies (see Chap. 12) and also opened entirely new perspectives on the working of viral nanomachines. An artistic view of a typical bacteriophage manipulation assay in the optical tweezers is represented in Fig. 9.8.

Measurement of the portal motor of bacteriophage Φ29 in the passive mode showed that its packaging machinery is highly processive (*i.e.* that can undergo many catalytic cycles without dissociation) and one of the strongest known in nature. Forces against an external load (like that exerted by a micropipette, Fig. 9.7b) or by another optical trap (Fig. 9.9a) can be above 60 pN. These forces are near those that distort the dsDNA in the overstretching transition (see force-extension curves in Fig. 9.7a). Constant-force mode studies indicated that packaging rates, which can reach 100 base pairs per second (100 bp/s), are punctuated by pauses and slips, and that these rates are dependent on the percentage of total genome packaged (see Fig. 9.7c).

Since that time, the Smith and Bustamante labs have contributed to understanding the mechanochemical behavior of this machinery. The former extended the use of optical tweezers to other phages in an assay similar to that in Fig. 9.7b. Studies of T4 [29] and λ [30] phages showed that high force generation is a common property of viral packaging motors. Among other results, this group found that T4 works at a strikingly high translocation rate, on average 700 bp/s, with maximum rates of

Fig. 9.8 A bacteriophage packaging assay in the optical tweezers. This artist's conception depicts a bacteriophage Φ29 packaging its genome against a force applied from a laser-trapped bead. The viral head is attached to the upper bead by chemical bonds represented by chains. The *in vitro* experiment is performed in aqueous buffer. A fibreless Φ29 virion map (cross section) at 7.8-Å resolution, obtained from EMdatabank.org (ID: emd-1,420 [27]), has been modified for illustrative purposes

2,000 bp/s [29]. In λ phage, procapsid expansion (see Chap. 13) was detected at 30 % of total packaging [30]. The Smith group also analyzed ionic effect on packaging rates and on portal motor function [31], as well as pressure within the capsids at different stages of the packaging process. They developed a new strategy to analyze the packaging process at near-zero force, which also involves initiation of the process without using a prepackaged configuration [16].

The dependence of the internal force on capsid filling is also a relevant analysis for its implication in the capsid pressure. The Smith team reported internal forces exceeding 100 pN, which are also indicative of the resistance of the capsid shells; these forces can be characterized in indentation experiments using AFM (see Chaps. 8 and 18). One of these studies showed that the capsid can withstand indentation forces up to ~1,000 pN [32], a result compatible with internal forces of 100 pN after complete capsid filling [16]. In addition, Smith and co-workers also analyzed the effect of mutations in the phage λ motor on its packaging dynamics [33, 34].

The Bustamante group determined the kinetic parameters of the Φ29 packaging motor and their dependence on external load. These researchers also showed that

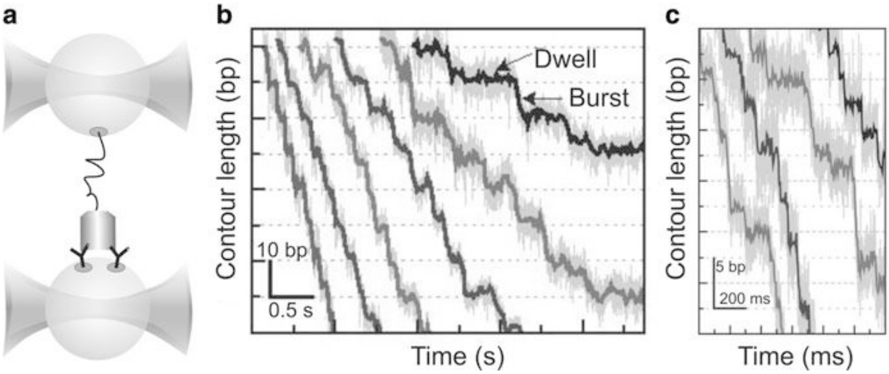


Fig. 9.9 High resolution measurements of packaging activity. Detecting a structure in motion on increasingly smaller scales involves a simultaneous improvement on spatial and temporal resolutions. (**a**) A dual-trap optical tweezers in which two separate traps are built from the same laser beam allow better sensitivity, as detailed elsewhere [7, 8]. (**b**) Step size in Φ29. The graph shows representative packaging traces collected under low external load ~8 pN, and different [ATP] (from *left* to *right*): 250, 100, 50, 25, 10 and 5 μM, all boxcar-filtered and decimated to 50 Hz. Data at 1.25 kHz are plotted in *light grey*. Packaging is observed to occur in large 10-bp "bursts" separated by flat "dwells". (**c**) Fine structure of the 10-bp bursts: each step consists of four 2.5-bp substeps. The graph shows representative packaging traces collected with external loads of ~40 pN and 250 μM [ATP], all boxcar-filtered and decimated to 100 Hz. Data in *light grey* are plotted at 1.25 kHz (Reproduced from [28] with permission)

the motor subunits act in a coordinated, successive fashion with high processivity [35]. Their study suggests that the packaging rate decrease shown in Fig. 9.7c at high capsid fillings results from build-up of internal pressure due to DNA confinement, which exerts an opposing force and slows the motor. Recent studies from this lab characterized the intersubunit coordination of the homomeric ring ATPase, which powers the DNA packaging process in Φ29. ATPase characterizations of this kind are important, because this machinery is also involved in other varied tasks in the cell, from chromosome segregation to protein degradation [28]. For these studies, the Bustamante lab used high-resolution optical tweezers (see Sect. 9.6). They found that packaging occurs in increments of 10 bp, in turn composed of four 2.5-bp steps, as shown in Fig. 9.9. An important conclusion is that a non-integer number of base-pair steps can take place in protein mechanics, in contrast to earlier models that assumed an integer number. In another assay using several modified DNAs, these researchers dissected phosphate interactions in the DNA-motor engagement and showed that promiscuous, nonspecific contacts with the substrates are common in translocase-substrate interactions [36].

Promiscuity by viral DNA packaging machinery was also recently studied by Rao and colleagues [37], who reported similar packaging activity rates for proheads and previously filled heads of T4 bacteriophage. Full heads and those emptied of most of their packaged DNA can reassemble the packaging machinery and reuse it. This group also found that several exogenous DNA molecules can be used as a substrate to fill the T4 capsid in several initiation steps [37].

The use of magnetic tweezers to provide a tethering force over a single DNA-phage complex was used in conjunction with single molecule fluorescence. In this experiment, Bustamante and colleagues tested the possibility of connector rotation in Φ29 [38]. Using an impressive variety of instrumentation and techniques, they showed that absolute rotations do not take place; only transient rotation of the connector relative to the capsid shell, followed by return to the original position, is likely to be involved in the packaging process.

The mechanical properties and structural transitions of nucleic acids in the presence of force are useful information to devise and test DNA confinement models, as prompted in the previous subsection. Several theoretical models have been proposed to understand DNA organization inside the capsids (see [25] and references therein). These models together with the above experimental analysis on single viruses are needed to understand several aspects of the relationship between a virus and its substrate; for example, the connection between internal pressure and DNA ejection forces.

Finally, single-molecule manipulation studies have provided new insights into the theoretical modeling of bacteriophage packaging mechanisms, which would be limited if based only on information from structural studies, solution biochemistry and genetics [39, 40]. For example, the possibility that DNA rotates during packaging, a process compatible with that in which the connector does not rotate can be directly tested. (A translocation mechanism driven by a (non-rotating) compression/extension mechanism is not supported by experimental evidence). The connection between DNA structure and the viral packaging mechanism at high force can also be enlightened. Although the portion of DNA engaged in the ATPases is closely in B-form due to structural constraints in the connector, the DNA fragment outside the connector and capsid should strongly deviate from a double helix at forces above 65 pN. It is thus tempting to speculate that ATPases might have trouble in doing good contacts with the DNA polymer above 65 pN thus decreasing the packaging efficiency. This fact might be one of the reasons why average maximum packaging forces are around 50–60 pN in Φ29. In order to test this scenario, constant forces above 65 pN should be directly applied against the motor.

9.5.3 Manipulation of Individual Viruses and Other Configurations

Direct trapping of individual virions, without using mechanical handles, is possible and has been shown using the tobacco mosaic virus in one of the pioneering manipulation experiments by Ashkin. The insights into the field of virology provided by these studies were limited, however, and this method has not found broad applications so far [41]. However, large viruses of hundreds of nanometers are amenable to individual examination in an optical trap, and this setup can be used to test direct responses to external physical stimuli (electric fields, local heat, or a combination of these with chemical changes such as those of pH and ionic strength), as shown for other biological structures [9]. Motion against fluid force can be used to characterize capsid friction and hydrodynamic size. The physical characterization of virus particles may be important not only to understand virus function, but also for nanotechnological applications of virus particles (see Chap. 22).

Virus-cell attachments have also been characterized by force-measuring optical tweezers. In these experiments, a laser-trapped bead coated by influenza virions was approached to a host cell to measure the interactions between an individual virus and the cell receptors. Rupture forces in the range of 10–25 pN and sequential unbinding events were detected and confirmed by similar AFM assays. Molecular dynamics simulations (see Chap. 19) were complementarily used to explain unbinding pathways and interaction dynamics [42]. In another configuration, cell-cell transmission of HIV-1 virus through a synapse was demonstrated by manipulation of single cells with optical tweezers [43].

9.5.4 A Note on Viral Nucleic Acid Replication and Transcription

The replicative protein machinery of viruses, which involve a variety of force-developing motors such as polymerases and helicases, is the object of a growing number of mechanistic studies at the single-molecule level using optical tweezers. These studies exceed the scope of this book; however, interested readers will find illustrative contributions in the literature [44–46].

9.6 Perspectives

There is little debate today as to the importance of viewing processes that take place inside cells as a product of nanoscale machinery. Proteins, nucleic acids and biomolecular complexes are not only chemical entities, they are known to perform mechanical work, an activity inherently related to their function. Manipulation techniques have become key instruments for identifying the mechanochemical activity of this biological machinery. These techniques are thus used not only as confirmatory tests for chemical and structural models, but also address novel mechanisms that cannot otherwise be determined [3].

Optical tweezers are one of the most important tools in this endeavor, and have furnished continuous advances since Ashkin's pioneering work in 1970 [47]. In the previous section Sect. (9.5) and in Chemla and Smith's chapter [25], some perspectives for the study of viruses themselves are discussed. Here we focus on optical advances and hybrid instruments.

9.6.1 Improvements in Resolution

One of the most important challenges in the field of biomolecular manipulation is to increase resolution in measurements. In bacteriophages, motors that translocate or move nucleic acids are inextricably coupled to their geometric structure. They use these motors to encapsulate nucleic acids to near crystalline densities inside a nanoscopic container. DNA is a double-helical sequence with a distance between base pairs of 0.34 nm in B-form. Resolving discrete steps in DNA packaging can involve spatial resolution of single base pairs, which is limited by drift and noise. Noise can be that intrinsic to the experiments, caused by fundamental Brownian fluctuations, and extrinsic, caused by instrumental and environmental disturbances (power fluctuations of the lasers, electronic noise, acoustic and mechanical room vibrations, etc.). A drift in micropipette position or changes in optical components (refractive index variations or thermal expansion/contraction of lenses and mirrors) due to temperature gradients, for example, can alter the calibration. Noise and drift can thus lead to false signals or reduce the accuracy of measurements.

Temporal resolution is another issue that is continually under development. Unlike structural studies in which the specimen is static, dynamic studies involve detection of conformational changes in structures in motion on their most fundamental temporal scale. Resolution is actually a combined function of space and time in optical tweezers and thus, improvements in resolution entail a simultaneous advance on spatial and temporal detection. Improved sensitivity can uncover the discrete stepping structure of a coordinated machine, as shown in Fig. 9.9 for the Φ29 packaging mechanism [28]; in this example, the 10-bp steps of each translocation burst (Fig. 9.9b) was shown to be composed of four 2.5-bp substeps (Fig. 9.9c) by a high-resolution instrument [8].

Recent advances have pushed spatial and temporal resolution to the Angstrom and millisecond scales, respectively, with measurement stability within the second range. The use of two traps from the same laser beam [8] has specifically allowed direct observation of intersubunit coordination and of step size in the packaging motor of Φ29 [28]. Future progress is open to new ideas for specific treatment of each experimental noise source; for example, the use of smaller beads as mechanical handles decreases the effects of Brownian noise, although they are technically more difficult to manipulate [7].

9.6.2 Adding Translation Coordinates and Rotation

Measurement in more degrees of freedom constitutes another challenge. The result of several coordinated molecular motors is not only linear displacement, but also a general three-dimensional motion that might involve several translations and rotations. Multiple optical traps permit measurement of motion in more than one linear coordinate, as well as control over more than one molecule [48]. To date, molecular specimens have been biochemically linked by two specific sites to beads in one or two traps, but the use of other configurations with more than two traps was shown, for example, to allow simultaneous manipulation of several molecules in multiple directions [49].

Another development in manipulation that could help to understand viral packaging processes is the combination of torque measurements in the optical tweezers. Torque is easily produced in magnetic tweezers (see Sect. 9.3.7) and has been generated in optical tweezers using a rotating micropipette [19]; however, torque measurements are technically more difficult. Recent advances point toward use of the intrinsic properties of light [50]. Light can exchange angular momentum -spin or orbital angular momentum, or both- and can thus exert rotation on particles. Conversely, angular momentum changes in the light scattered from a particle can be analyzed as a method to measure torque [51].

The helical structure of DNA suggests that its binding machinery follows its helical pitch to generate rotation. As discussed above, DNA may move by ratcheting and rotation in the Φ29 phage portal motor in its translocation into the viral capsid, a mechanism that could be examined by a combination of fluorescence microscopy and optical tweezers (see below). Torque-measuring optical tweezers could be then applied to characterize torsion-induced mechanochemical packaging pathways by over- and underwinding the DNA double helix.

9.6.3 Combination with Fluorescence Imaging

Imaging of molecules while they are being manipulated is a long-standing demand. There is currently no method with which to perform structural studies with atomic resolution of a molecule in motion or under mechanical stress. The conditions involved in methods such as X-ray diffraction or NMR are incompatible with those used in liquid chambers. It is nonetheless possible to introduce methods developed for optical microscopes in the tweezers setup.

Combining fluorescence with optical tweezers not only provides control over the experiment, but also allows observation of new conformations induced by force or torque. Single-molecule fluorescence and confocal imaging have already been incorporated successfully to optical tweezers, and their applicability is shown by a number of studies (see [7, 52] and references therein). A general strategy to integrate fluorescence detection in optical tweezers was outlined briefly in Fig. 9.5 [53–55]. Fluorescence can be obtained using fluorophores or with quantum dots, which add stability in signal emission. The use of a pair of dyes allows detection of nanoscale distance changes of a protein or of a complex of coordinated proteins using Förster resonant energy transfer (FRET) photophysical interactions [56].

Fluorescence has been exploited to understand DNA structure under tension as well as the portal motor dynamics of bacteriophages [38, 52]. These studies suggest a fruitful combination of methods to visualize bacteriophage motor dynamics under the control of force and torque. DNA could be fluorescently labeled to follow its twisting dynamics relative to the connector in the packaging experiment.

9.6.4 Conclusions

We have introduced a technique with which the mechanical action of viruses as well as the elastic properties and condensation dynamics of nucleic acids can be analyzed. Optical tweezers are a laser-based tool with which to trap and manipulate individual particles with piconewton forces; they are used in molecular biology to measure pairs of associated magnitudes such as force-extension and torque-rotation as a function of time. The resulting dynamic information from these experiments is yielding not only demonstrative experiments of previously conjectured mechanistic models, but also unparalleled characterization of the viral machinery. Since the experiments are performed in a liquid chamber, optical tweezers can be adapted to a large variety of physiological conditions. They can also be merged with optical microscopies, allowing information from manipulation experiments and real-time imaging to be integrated in the same experiment.

The dynamic information provided by optical tweezers and its combination with structural and biochemical analysis provide rigorous ways to model viral function and stability. These models in turn permit more precise thermodynamic characterization of the viral machinery, and form a basis for the development of nanotechnological applications.

Acknowledgements It is a pleasure to acknowledge J.R. Moffitt and J.L. Carrascosa for technical insights into different aspects of the chapter, R. Bocanegra, L. Quintana for careful reading of the manuscript, C. Mark and S. Hormeño for editorial and illustration assistance, respectively, and M. de la Guía for graphic design of Fig. 9.8. This work was funded by the Spanish Ministry of Science and Innovation under the "Ramon y Cajal" program (Grant No. RYC-2007-01765).

References and Further Reading

1. Hormeno S, Arias-Gonzalez JR (2006) Exploring mechanochemical processes in the cell with optical tweezers. Biol Cell 98:679–695
2. Svoboda K, Block SM (1994) Biological applications of optical forces. Annu Rev Biophys Biomol Struct 23:247–285
3. Bustamante C (2008) In singulo biochemistry: when less is more. Annu Rev Biochem 77:45–50
4. Oster G, Wang H (2003) How protein motors convert chemical energy into mechanical work. In: Schliwa M (ed) Molecular motors. Wiley-VCH, Weinheim, pp 207–227
5. Ketterle W (1999) Experimental studies of Bose-Einstein condensation. Phys Today 52:30–35
6. Smith SB, Cui Y, Bustamante C (2003) Optical-trap force transducer that operates by direct measurement of light momentum. Methods Enzymol 361:134–162
7. Moffitt JR, Chemla YR, Smith SB, Bustamante C (2008) Recent advances in optical tweezers. Annu Rev Biochem 77:205–228

8. Moffitt JR, Chemla YR, Izhaky D, Bustamante C (2006) Differential detection of dual traps improves the spatial resolution of optical tweezers. Proc Natl Acad Sci U S A 103:9006–9011
9. Hormeno S, Ibarra B, Chichon FJ, Habermann K, Lange BM, Valpuesta JM, Carrascosa JL, Arias-Gonzalez JR (2009) Single centrosome manipulation reveals its electric charge and associated dynamic structure. Biophys J 97:1022–1030
10. Jackson JD (1999) Classical electrodynamics. Wiley Hoboken, USA
11. Arias-Gonzalez JR, Nieto-Vesperinas M (2003) Optical forces on small particles: attractive and repulsive nature and Plasmon-resonance conditions. J Opt Soc Am A 20:1201–1209
12. Tanase M, Biais N, Sheetz M (2007) Magnetic tweezers in cell biology. In: Yu-Li W, Dennis ED (eds) Methods in cell biology, vol 83. Academic Press, pp 473–493, vol 83
13. Neuman KC, Nagy A (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. Nat Methods 5:491–505
14. Bustamante C, Macosko JC, Wuite GJ (2000) Grabbing the cat by the tail: manipulating molecules one by one. Nat Rev Mol Cell Biol 1:130–136
15. Hormeno S, Moreno-Herrero F, Ibarra B, Carrascosa JL, Valpuesta JM, Arias-Gonzalez JR (2011) Condensation prevails over B-a transition in the structure of DNA at low humidity. Biophys J 100:2006–2015
16. Rickgauer JP, Fuller DN, Grimes S, Jardine PJ, Anderson DL, Smith DE (2008) Portal motor velocity and internal force resisting viral DNA packaging in bacteriophage phi29. Biophys J 94:159–167
17. Smith DE, Tans SJ, Smith SB, Grimes S, Anderson DL, Bustamante C (2001) The bacteriophage straight phi29 portal motor can package DNA against a large internal force. Nature 413:748–752
18. Li PT, Collin D, Smith SB, Bustamante C, Tinoco I Jr (2006) Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. Biophys J 90:250–260
19. Bustamante C, Bryant Z, Smith SB (2003) Ten years of tension: single-molecule DNA mechanics. Nature 421:423–427
20. Williams MC, Rouzina I, Bloomfield VA (2002) Thermodynamics of DNA interactions from single molecule stretching experiments. Acc Chem Res 35:159–166
21. Mao H, Arias-Gonzalez JR, Smith SB, Tinoco I Jr, Bustamante C (2005) Temperature control methods in a laser tweezers system. Biophys J 89:1308–1316
22. Hormeno S, Ibarra B, Valpuesta JM, Carrascosa JL, Arias-Gonzalez JR (2012) Mechanical stability of low-humidity single DNA molecules. Biopolymers 97:199–208
23. Chaurasiya KR, Paramanathan T, McCauley MJ, Williams MC (2010) Biophysical characterization of DNA binding from single molecule force measurements. Phys Life Rev 7:299–341
24. Casjens SR (2011) The DNA-packaging nanomotor of tailed bacteriophages. Nat Rev Microbiol 9:647–657
25. Chemla YR, Smith DE (2012) Single-molecule studies of viral DNA packaging. Adv Exp Med Biol 726:549–584
26. Feiss M, Rao VB (2012) The bacteriophage DNA packaging machine. Adv Exp Med Biol 726:489–509
27. Tang J, Olson N, Jardine PJ, Grimes S, Anderson DL, Baker TS (2008) DNA poised for release in bacteriophage phi29. Structure 16:935–943
28. Moffitt JR, Chemla YR, Aathavan K, Grimes S, Jardine PJ, Anderson DL, Bustamante C (2009) Intersubunit coordination in a homomeric ring ATPase. Nature 457:446–450
29. Fuller DN, Raymer DM, Kottadiel VI, Rao VB, Smith DE (2007) Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. Proc Natl Acad Sci U S A 104:16868–16873
30. Fuller DN, Raymer DM, Rickgauer JP, Robertson RM, Catalano CE, Anderson DL, Grimes S, Smith DE (2007) Measurements of single DNA molecule packaging dynamics in bacteriophage lambda reveal high forces, high motor processivity, and capsid transformations. J Mol Biol 373:1113–1122

31. Fuller DN, Rickgauer JP, Jardine PJ, Grimes S, Anderson DL, Smith DE (2007) Ionic effects on viral DNA packaging and portal motor function in bacteriophage phi 29. Proc Natl Acad Sci U S A 104:11245–11250
32. Ivanovska IL, de Pablo PJ, Ibarra B, Sgalari G, MacKintosh FC, Carrascosa JL, Schmidt CF, Wuite GJ (2004) Bacteriophage capsids: tough nanoshells with complex elastic properties. Proc Natl Acad Sci U S A 101:7600–7605
33. Tsay JM, Sippy J, DelToro D, Andrews BT, Draper B, Rao V, Catalano CE, Feiss M, Smith DE (2010) Mutations altering a structurally conserved loop-helix-loop region of a viral packaging motor change DNA translocation velocity and processivity. J Biol Chem 285:24282–24289
34. Tsay JM, Sippy J, Feiss M, Smith DE (2009) The Q motif of a viral packaging motor governs its force generation and communicates ATP recognition to DNA interaction. Proc Natl Acad Sci U S A 106:14355–14360
35. Chemla YR, Aathavan K, Michaelis J, Grimes S, Jardine PJ, Anderson DL, Bustamante C (2005) Mechanism of force generation of a viral DNA packaging motor. Cell 122:683–692
36. Aathavan K, Politzer AT, Kaplan A, Moffitt JR, Chemla YR, Grimes S, Jardine PJ, Anderson DL, Bustamante C (2009) Substrate interactions and promiscuity in a viral DNA packaging motor. Nature 461:669–673
37. Zhang Z, Kottadiel VI, Vafabakhsh R, Dai L, Chemla YR, Ha T, Rao VB (2011) A promiscuous DNA packaging machine from bacteriophage T4. PLoS Biol 9:e1000592
38. Hugel T, Michaelis J, Hetherington CL, Jardine PJ, Grimes S, Walter JM, Falk W, Anderson DL, Bustamante C (2007) Experimental test of connector rotation during DNA packaging into bacteriophage phi29 capsids. PLoS Biol 5:e59
39. Purohit PK, Kondev J, Phillips R (2003) Mechanics of DNA packaging in viruses. Proc Natl Acad Sci U S A 100:3173–3178
40. Yu J, Moffitt J, Hetherington CL, Bustamante C, Oster G (2010) Mechanochemistry of a viral DNA packaging motor. J Mol Biol 400:186–203
41. Ashkin A, Dziedzic JM (1987) Optical trapping and manipulation of viruses and bacteria. Science 235:1517–1520
42. Sieben C, Kappel C, Zhu R, Wozniak A, Rankl C, Hinterdorfer P, Grubmuller H, Herrmann A (2012) Influenza virus binds its host cell using multiple dynamic interactions. Proc Natl Acad Sci U S A 109:13626–13631
43. McNerney GP, Hubner W, Chen BK, Huser T (2010) Manipulating CD4+ T cells by optical tweezers for the initiation of cell-cell transfer of HIV-1. J Biophoton 3:216–223
44. Herbert KM, Greenleaf WJ, Block SM (2008) Single-molecule studies of RNA polymerase: motoring along. Annu Rev Biochem 77:149–176
45. Pyle AM (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. Annu Rev Biophys 37:317–336
46. van Oijen AM, Loparo JJ (2010) Single-molecule studies of the replisome. Annu Rev Biophys 39:429–448
47. Ashkin A (1970) Acceleration and trapping of particles by radiation pressure. Phys Rev Lett 24:156–159
48. Grier DG (2003) A revolution in optical manipulation. Nature 424:810–816
49. Dame RT, Noom MC, Wuite GJ (2006) Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. Nature 444:387–390
50. Gutierrez-Medina B, Andreasson JO, Greenleaf WJ, Laporta A, Block SM (2010) An optical apparatus for rotation and trapping. Methods Enzymol 475:377–404
51. Parkin S, Knoner G, Singer W, Nieminen TA, Heckenberg NR, Rubinsztein-Dunlop H (2007) Optical torque on microscopic objects. Methods Cell Biol 82:525–561
52. van Mameren J, Peterman EJ, Wuite GJ (2008) See me, feel me: methods to concurrently visualize and manipulate single DNA molecules and associated proteins. Nucleic Acids Res 36:4381–4389

53. Lang MJ, Fordyce PM, Engh AM, Neuman KC, Block SM (2004) Simultaneous, coincident optical trapping and single-molecule fluorescence. Nat Methods 1:133–139
54. Richardson AC, Reihani N, Oddershede LB (2006) In: Dholakia K, Spalding GC (eds) Combining confocal microscopy with precise force-scope optical tweezers, vol 6326. SPIE, San Diego, pp 632628–632637
55. Vossen DLJ, van der Horst A, Dogterom M, van Blaaderen A (2004) Optical tweezers and confocal microscopy for simultaneous three-dimensional manipulation and imaging in concentrated colloidal dispersions. Rev Sci Instrum 75:2960–2970
56. Toprak E, Selvin PR (2007) New fluorescent tools for watching nanometer-scale conformational changes of single molecules. Annu Rev Biophys Biomol Struct 36:349–369

Further Reading

Berns MW, Greulich KO (2007) Laser manipulation of cells and tissues. Methods Cell Biol 82
Schliwa M (ed) (2003) Molecular motors. Wiley-VCH Weinheim (Germany)
Sheetz MP (1998) Laser tweezers in cell biology. Methods Cell Biol 55

Also especially recommended for further reading are references [1, 7, 13, 24, 25] *listed above.*