Computational and Structural Biotechnology Journal 21 (2023) 2858–2872

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Mini-review

Structural studies of protein–nucleic acid complexes: A brief overview of the selected techniques

Kamil Szpotkowski, Klaudia Wójcik, Anna Kurzyńska-Kokorniak *

Department of Ribonucleoprotein Biochemistry, Institute of Bioorganic Chemistry Polish Academy of Sciences, 61-704 Poznan, Poland

ARTICLE INFO

Article history:
Received 30 October 2022
Received in revised form 28 April 2023
Accepted 28 April 2023
Available online 29 April 2023

Keywords:
Nucleic acid binding proteins
Protein-nucleic acid complex
Nucleoproteins
RNA-protein interactions
DNA-protein interactions
Crystallography
Nuclear Magnetic Resonance (NMR)
Cryogenic Electron Microscopy (cryo-EM)
Atomic force microscopy (AFM)
Small Angle X-ray Scattering (SAXS)
Small Angle Neutron Scattering (SANS)
Circular Dichroism (DC)
Infrared Scpectroscopy (IR)

ABSTRACT

Protein–nucleic acid complexes are involved in all vital processes, including replication, transcription, translation, regulation of gene expression and cell metabolism. Knowledge of the biological functions and molecular mechanisms beyond the activity of the macromolecular complexes can be determined from their tertiary structures. Undoubtably, performing structural studies of protein-nucleic acid complexes is challenging, mainly because these types of complexes are often unstable. In addition, their individual components may display extremely different surface charges, causing the complexes to precipitate at higher concentrations used in many structural studies. Due to the variety of protein-nucleic acid complexes and their different biophysical properties, no simple and universal guideline exists that helps scientists chose a method to successfully determine the structure of a specific protein-nucleic acid complex. In this review, we provide a summary of the following experimental methods, which can be applied to study the structures of protein-nucleic acid complexes: X-ray and neutron crystallography, nuclear magnetic resonance (NMR) spectroscopy, cryogenic electron microscopy (cryo-EM), atomic force microscopy (AFM), small angle scattering (SAS) methods, circular dichroism (CD) and infrared (IR) spectroscopy. Each method is discussed regarding its historical context, advancements over the past decades and recent years, and weaknesses and strengths. When a single method does not provide satisfactory data on the selected protein–nucleic acid complex, a combination of several methods should be considered as a hybrid approach; thus, specific structural problems can be solved when studying protein-nucleic acid complexes.

1. Introduction

Protein–nucleic acid complexes are defined as conjugates of proteins and nucleic acids, including DNA or RNA or both. Sometimes protein–nucleic acid complexes are also called nucleoproteins; however, it should be emphasized that the term “nucleoproteins” indicates a group of proteins that are associated with nucleic acids. Nucleoproteins can consist of only one protein and one nucleic acid molecule, e.g., simple signal recognition particles (SRPs) [1], but they can also form more complex structures, such as ribosomes, nucleosomes or viral nucleocapsids [2,3]. Nucleoproteins play a very important role in all living organisms. They are involved in all vital processes, including replication, transcription, translation, regulation of gene expression and cell metabolism [4]. Over the years, as scientific fields have developed, especially structural and functional biochemistry, biology and biotechnology, it was revealed that information regarding the function of biomolecules is encrypted in their structures [4].

Historically, the oldest method of studying the structure of biomolecules is crystallography. Its development since the 1960 s has laid the foundations for today’s structural biology. Crystallographic methods can provide information about the structure of protein–nucleic acid complexes with atomic resolution; nevertheless, many protein–nucleic acid complexes fail to crystallize readily or at all. In this case, other methods should be considered to determine the structure of protein-nucleic acid complexes. These methods include nuclear magnetic resonance (NMR) spectroscopy, cryogenic electron microscopy (cryo-EM), atomic force microscopy (AFM), small angle scattering (SAS) methods, circular dichroism (CD) and infrared (IR) spectroscopy, and their modifications. In contrast to the crystallographic methods, NMR spectroscopy, cryo-EM and AFM, approaches such as SAS, CD and IR spectroscopy, do not provide information about the structure of biomolecules with atomic resolution. X-ray crystallography and cryo-EM are used to study large

* Correspondence to: Noskowskiego 12/14, 61-704 Poznan, Poland.
E-mail address: akurzyns@man.poznan.pl (A. Kurzyńska-Kokorniak).

https://doi.org/10.1016/j.csbj.2023.04.028

and rigid nucleoprotein complexes, for which they provide structural snapshots reflecting their functional activity [4]. AFM allows for visualization, with atomic resolution, of even single molecules. Through in-solution structural techniques, such as NMR spectroscopy, SAS, CD and IR spectroscopy, the dynamics of protein-nucleic acid complexes can be analyzed, as well as the mechanisms of interactions between their components. Moreover, techniques to analyze biomolecule structures in-solution are important because all known biological processes in the human body occur in water environments. Water and its unique properties are key factors in the folding processes of biomolecules, including proteins and their complexes with nucleic acids [5,6]. Through analyses of the biomolecule structure in near-physiological conditions, combined with in vitro and *in cellulo* studies, the biological functions and cellular activities of biomolecules can be reasoned.

Undoubtably, substantial progress in structural biology of protein–nucleic acid complexes was achieved when genetic engineering methods were introduced, especially various methods of cloning, protein expression, and in vitro transcription, as well as the nonenzymatic chemical synthesis of nucleic acids. Since then, proteins and nucleic acids could be produced in large quantities with good quality, e.g., for use in structural studies.

The purpose of this mini-review is to provide a brief overview of the techniques that can be applied to study the structures of protein-nucleic acid complexes: X-ray and neutron crystallography, NMR spectroscopy, cryo-EM, AFM, SAS methods, and CD and IR spectroscopy. Each of the methods is discussed regarding its historical context, advancements over the past decades and recent years, and weaknesses and strengths. The manuscript is dedicated to a broad audience of scientists and researchers interested in structural studies of protein-nucleic acid complexes, including those who have only little or no experience in the field.

2. Crystallography

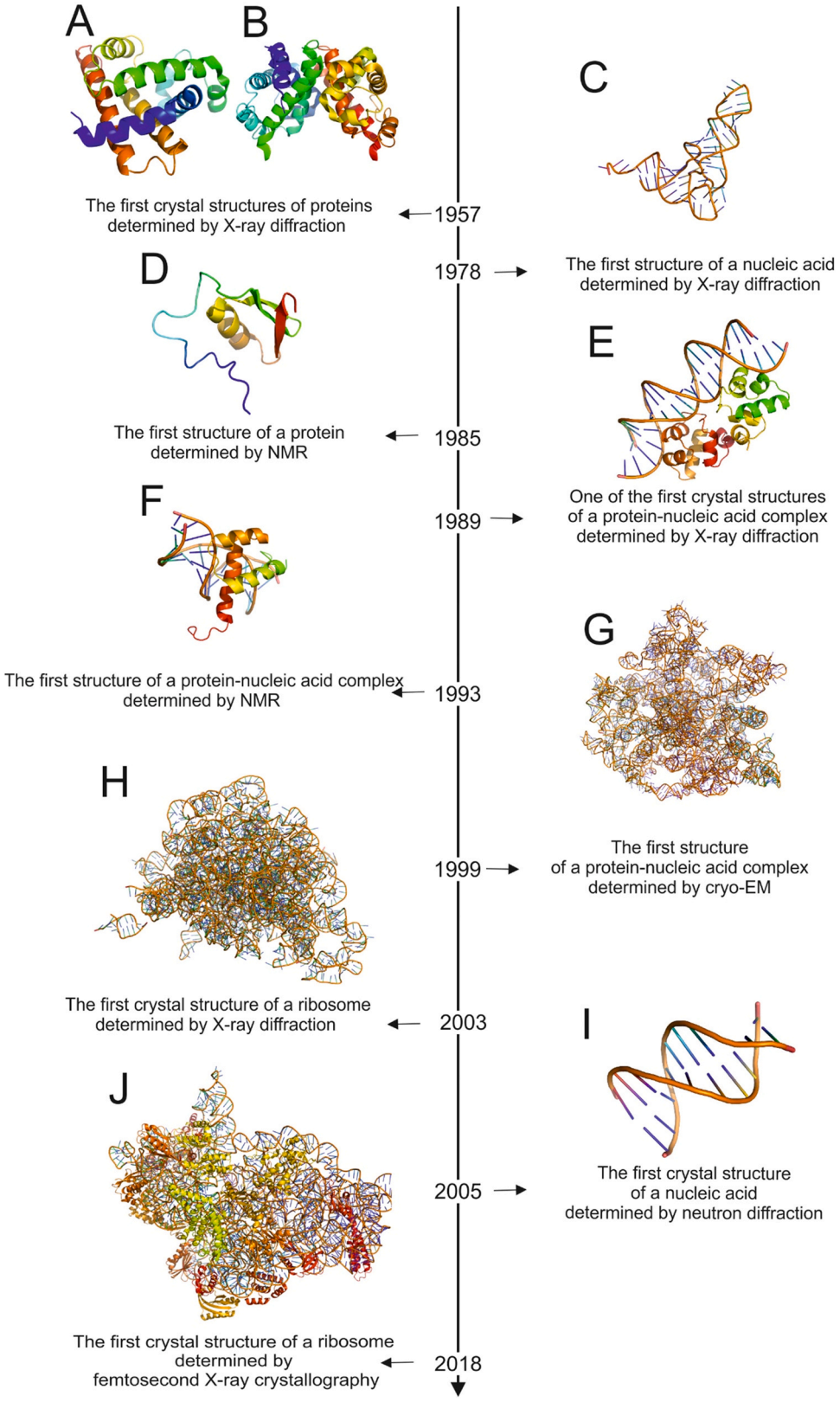
When W. Roentgen discovered X-rays in 1895, it was certain that science and medical diagnostics would be revolutionized. X-rays have been adapted for determining the structures of molecules since the 1920 s. The first crystal structure of a protein was published in 1957 by J. Kenrew et al., and it was the structure of an oxygen-binding protein, myoglobin, at 2 Å resolution [7]. In 1959, M. Perutz determined the structure of horse deoxyhaemoglobin at 2.8 Å resolution in the same laboratory [8]. One of the first crystal structures of a protein-nucleic acid complexes was the structure of a repressor-operator complex of bacteriophage 434, which was published in 1988 by A.K. Aggarwal et al. [9]. However, the first, most spectacular nucleoprotein structure determined by this method was for the ribosome of archaeon *Haloarcula marismortui*; it was published in 1991 by A. Yonath [10] and her research team (summarized in Fig. 1).

Through crystallography, the molecular structure can be determined with atomic resolution, and the physical basis of this technique involves the diffraction of electrons or neutrons by crystals formed by molecules. The scheme of diffraction is quite simple. The monochromatic beam passes through the crystal and is diffracted by the sample's crystal lattice, producing a characteristic diffraction pattern on the detector. This pattern depends on the spatial orientation of atoms in the crystal molecule. In general, crystallography is the method of choice for determining the structure of large and stable protein-nucleic acid complexes. When determining the crystal structure of protein-nucleic acid complexes by X-ray or neutron crystallography, the following steps are performed: (i) a protein and nucleic acid are prepared and then these two components are mixed to obtain a protein-nucleic acid complex, (ii) initial crystallization conditions are determined, (iii) crystal quality is optimized, (iv) diffraction data are collected, (v) the structure of the three-dimensional model is determined and refined, and (vi) the refined model is analyzed [11]. Crystallization, despite being automatized, remains the bottleneck of successful structure determination. Crystal formation is a multiparametric process that depends on many physical, chemical and biochemical factors [12], and this process can be supported by precipitating agents and nucleation additives. First, precipitating agents are chemical compounds that reduce protein solubility. They reinforce the attraction among molecules, e.g., by altering the activity coefficient of water (i.e., the mole fraction of water in the aqueous fraction) or by increasing molecular crowding. For most proteins, the degree of their solubility depends strongly on the kind of anion but weakly on the kind of cation present in the solution. The "Hofmeister series" (lyotropic series) are ways to classify ions based on their ability to salt out or salt in proteins [13]. Second, nucleation additives can bind to a protein and thus modify and/or stabilize the protein conformation or perturb protein−protein and protein-solvent interactions.

Crystallizing protein-nucleic acid complexes is an inherently difficult process because these complexes are stable only under some experimental conditions [11]. These conditions mostly depend on the individual components of the complexes and the sample buffer compositions. Some rules can be followed to increase the probability that protein-nucleic acid complexes are successfully crystallized. Notably, many nucleic acid-binding proteins contain some flexible regions and can thus fit into the structure of the interacting RNA or DNA [14]. These flexible regions may contain disordered domains and disordered or flexible loops that may negatively affect the formation of the well-ordered crystal lattice. Therefore, usually only the ordered protein domains are chosen for crystallization. Additionally, longer nucleic acids are also flexible and polymorphic. Consequently, only shorter, rigid fragments of nucleic acids are chosen for crystallization [15]. Another problem is the stability of the protein-nucleic acid complexes. Usually, the conditions under which an individual protein-nucleic acid complex is stable and can form a crystal must be determined experimentally. This process involves searching for the optimal buffer conditions, the type and concentration of the ions, precipitating agents and nucleation additives, etc. [16]. Typically, a protein and nucleic acid are mixed at a 1:1.2–1.5 molar ratio. These conditions occur because (i) the entire nucleic acids is usually not perfectly folded, (ii) the substrate concentration is often an approximate estimation and (iii) an excess of nucleic acid is not an issue for crystallization [14]. When crystals are obtained, before measurements, they are mounted in a small loop and frozen in a stream of liquid nitrogen. Through data collection from frozen crystals, we can obtain a complete dataset from a single crystal.

It must be emphasized that X-rays interact with electrons, which is why hydrogens are almost invisible during X-ray diffraction; only some hydrogens can be seen in high-resolution structures [17]. For this reason, neutron crystallography is used to directly locate the hydrogen atom positions in a protein or nucleic acid structure. Neutron crystallography is a very useful technique for structural studies on the mechanism of interactions between proteins and nucleic acids because these interactions involve hydrogen atoms and are largely based on hydrogen bonds [18]. An undoubted advantage of using neutrons in crystallography is that neutrons do not destroy samples and measurements can be obtained at room temperature. Unfortunately, there are far fewer neutron beamlines than X-ray beamlines worldwide. Moreover, to perform neutron crystallography, the sample must be deuterated, which is a complex and expensive procedure [19].

In 1996, a new method based on crystallization in living cells was reported by G.Y. Fan et al. [20]. The authors presented the first protein crystalized in living cells, and it was calcineurin, a calcium- and calmodulin-dependent serine/threonine protein phosphatase [20]. The development of *in cellulo* crystallization was possible due to new synchrotrons and free electron lasers. The X-ray free electron



(caption on next page)

Fig. 1. Milestones in the development of crystallography, nuclear magnetic resonance and cryogenic electron microscopy in structural studies of protein-nucleic acid complexes. Each of the presented structures provided new opportunities for researching protein-nucleic acid complexes. The first crystal structures of proteins were (A) myoglobin from *Physeter macrocephalus* (PDB entry 1MBN) and (B) horse hemoglobin (PDB entry 2DHB). Both structures were determined in 1957, but they were deposited in the Protein Data Bank (PDB) in 1973. (C) The first crystal structure of a nucleic acid was $tRNA^{PHE}$, which was deposited in PDB in 1978 (PDB entry 6TNA). (D) The first structure of protein determined by NMR was proteinase inhibitor IIA from bull seminal plasma, which was deposited in PDB in 1985 (PDB entry 2BUS). (E) One of the first crystal structures of a protein-nucleic acid complex was the repressor-operator complex of bacteriophage 434, which was deposited in PDB in 1989 (PDB entry 2OR1). (F) The first structure of a protein-nucleic acid complex determined by NMR was the structure of the Antennapedia homeodomain-DNA complex, which was deposited in PDB in 1993 (PDB entry 1AHD). (G) The first structure obtained by cryo-EM was 23 S rRNA, which was deposited in PDB in 1999 (PDB entry 1C2W). (H) The first crystal structure of ribosome was the ribosome of archaeon *Haloarcula marismortui*, which was published in 1991 but deposited in PDB in 2003 (PDB entry 1P9X). (I) The first and one of the few structures of nucleic acids determined by neutron diffraction crystallography was B-DNA decamer, which was deposited in PDB in 2005 (PDB entry 1WQZ). (J) One of the first crystal structures determined by serial femtosecond X-ray crystallography (SFC) was 30 S ribosomal subunit from *Thermus thermophiles*, which was determined in 2018 and deposited in PDB in 2018 (PDB entry 6CAS). The 3D structures were visualized by PyMOL.

laser (XFEL) is a light source that produces coherent X-ray pulses, with a peak brightness 10 orders of magnitude greater than that of synchrotrons [21]. The XFEL provided new possibilities for crystallography and resulted in the development of new crystallographic techniques, such as serial femtosecond crystallography (SFC) or single particle imaging (SPI). The SFC technique is dedicated to X-ray diffraction from nano- and microcrystals that are not suitable for single-crystal crystallography. Through this technique, the biomolecule structure and its dynamics can be studied during time-resolved measurements. The first structure obtained by SFC was the structure of Photosystem I [22]. In the SPI experiment, a large number of two-dimensional (2D) diffraction patterns (snapshots) of single molecules are recorded by a detector. Because these snapshots are extremely noisy and taken in random and unknown orientations, special algorithmic methods need to be applied to assign the orientation of the molecules to the single-particle X-ray diffraction patterns. Currently, SPI remains in the test phase, but future development of this technique can provide an excellent tool for studying the structure of biomolecules with atomic resolution. Notably, SPI does not require sample freezing, and the resolution of the images depends on the number of observed photons and hence the number of recorded images. The first nucleoprotein structures obtained by this technique were the structures of mimivirus [23] in 2016, coliphage virus [24] in 2020, and the structure of a small, hydrophobic, disulfide-rich, protein called crambin [25] in 2018.

3. Nuclear magnetic resonance spectroscopy (NMR)

The phenomenon of NMR was described for the first time in 1938 by I. Rabi, who won the Nobel Prize in Physics in 1944 for his discovery. Then, F. Bloch and E. Purcell expanded NMR for use on liquids and solids, and for their achievements, they shared the Nobel Prize in Physics in 1952.

NMR is a physical phenomenon in which the nuclei in a strong constant magnetic field are perturbed by a weak oscillating magnetic field and respond by producing an electromagnetic signal; this signal exhibits a frequency characteristic of the magnetic field at the nucleus [26]. This effect can be, however, observed only for atoms with an odd number of nucleons (i.e., the total of protons and neutrons), e.g., hydrogen ^{1}H, carbon ^{13}C, nitrogen ^{15}N, oxygen ^{17}O, fluorine ^{19}F, sodium ^{23}Na or phosphorus ^{31}P (i.e., the atoms with a nonzero spin). The most common types of NMR are ^{1}H and ^{13}C NMR spectroscopy. After crystallography, NMR is the second technique used to determine the structure of biomolecules with atomic resolution (Fig. 1).

Considering the protein-nucleic acid complexes, which comprise, inter alia, carbon atoms, it must be stressed that the most common isotope of carbon (^{12}C) is not active in the NMR effect [27]. Due to this limitation, protein or nucleic acid samples must be specially prepared to increase in the number of nuclei that produce signals during NMR experiments. For example, proteins can be produced in bacteria that are grown in media supplemented with nutrients comprising isotopically enriched ^{13}C and ^{15}N atoms. However, currently, many other techniques can increase the number of active nuclei in biomolecules for NMR [28].

In contrast to X-ray crystallography, with NMR spectroscopy, measurements can be obtained at room temperature and the structure, the dynamics, reaction state, and chemical environment of protein-nucleic acid complexes can be analyzed [29]. Moreover, while crystallographic methods cannot reveal disordered parts of a protein or flexible fragments of proteins and nucleic acids, these fragments of biomolecules can be visible in the NMR spectra. The advancements of NMR spectroscopy were possible due to the development of selective isotopic labeling techniques [30]. Importantly, the strategy of analyzing the structure of protein-nucleic acid complexes by NMR spectroscopy is based on selective labeling procedures. In general, one partner is labeled while the other remains unlabeled. In this way, only the selected NMR signals can be observed. In addition, when analyzing protein-nucleic acid complexes by NMR techniques, the charge of the protein and nucleic acid must be considered. The main mechanism of interactions between proteins and nucleic acids involves electrostatic interactions. Nucleic acids contain negatively charged phosphate groups, but fragments of the protein that bind the nucleic acid are positively charged. Strong electrostatic interactions between the protein and nucleic acid, which both occur at high concentrations in the NMR experiments, may result in protein-nucleic acid complex precipitation. To reduce the risk of sample precipitation, the following strategies can be applied: (i) increase the salt concentration of a sample buffer, (ii) reduce the number of nucleic acid charges by shortening its length, and (iii) use site-directed mutagenesis to replace basic residues on the protein surface outside the nucleic acid recognition interface [28].

The wide use of NMR spectroscopy is possible due to a chemical shift effect. Generally, a chemical shift is the value that determines the interactions among nuclei, electrons and the outer magnetic field (i.e., the value that a peak corresponds to within the NMR spectrum; this value is measured in parts per million (ppm)). The chemical shift value depends on the spatial orientation of the molecule and the neighboring chemical groups [31]. A major advantage of NMR spectroscopy is the possibility of measuring biomolecules in various sample states, including the solution state [27], solid-state [32], and membranous environments [33]. Importantly, through the development of NMR techniques, biomolecules can be studied even under natural conditions, i.e., in the cell [34].

A significant limitation of solution-state NMR spectroscopy is the molecular weight of biomolecules [35]. Regarding proteins, this technique is dedicated to molecules less than 50 kDa because (i) line-broadening of the spectrum is directly dependent on the molecular size (a spectrum line width is proportional to the number of nuclei in a given environment; when more nuclei are present, the spectral lines are broader), (ii) faster transverse relaxation rates (relaxation refers to the phenomenon of nuclei returning to their thermodynamically stable states after being excited to higher energy levels, and the transverse relaxation rate is proportional to the concentration of NMR active ions in the sample; the more NMR active ions there are, the slower the transverse relaxation rates), and (iii) overall spectral crowding due to the increase in the number of peaks [36]. This limit has been significantly extended, even up to 1 MDa, by the methyl-TROSY methodology (transverse relaxation-

optimized spectroscopy), in combination with the selective $^{13}CH_3$ methyl group labeling of highly deuterated proteins [37]. Fluorine ^{19}F is a very important isotope for methyl-selective labeling [28,38]. The advantages of fluorine ^{19}F labeling in NMR experiments are (i) elimination of background signals, (ii) 100% natural abundance, (iii) high sensitivity, and (iv) a large chemical shift range. For large challenging systems (e.g., complex nucleoproteins), single fluorinated phenylalanine and tryptophan residues can be used to analyze the key sites for interaction with nucleic acids [28].

It must be emphasized that in recent years, an increasing number of high molecular weight protein–RNA complexes (> 50 kDa) have been solved by coupling NMR spectroscopy with other techniques, e.g., SAS. One example is the structure of 390 kDa protein–RNA complexes of the archaeal Box C/D ribonucleoprotein bound to RNA [35,39]. The complementary techniques are used to yield additional restraints for structure calculation and validation from sparse NMR data.

Solid-state NMR (ssNMR) can be applied to insoluble and non-crystalline particles, such as membrane proteins [40], viral assemblies [41] or amyloid fibrils [42]. ssNMR can also be used to determine the structure of RNA, RNA bound to short peptides, and RNA as a part of complex nucleoprotein complexes [43,44]. Importantly, in contrast to solution-state NMR, the ssNMR line widths do not depend on the molecular weight; thus, complexes larger than 50 kDa can be studied with ssNMR. Notably, however, this technique requires selective isotope labeling because ssNMR lines are intrinsically broader than the solution NMR lines. Moreover, the sample for the ssNMR experiment must undergo special preparation, including micro(nano) crystallization, ethanol precipitation, lyophilization, freezing in the presence of a cryo-protectant or sedimentation of soluble macromolecules in the ssNMR rotor using ultracentrifugation [45]. In contrast, in solution-state NMR, samples can be studied under native conditions. The undoubted disadvantage of the ssNMR technique is the requirement of additional structural information from crystallography or solution-state NMR [46]. ssNMR was applied to determine the structure of some protein-nucleic complexes, e.g., bacterial DnaB helicase from *Helicobacter pylori* complexed with the transition-state ATP-analog and single-stranded DNA [47], archaeal pRN1 primase complexed with DNA in the presence and absence of bound ATP molecules [48] or DNA and the H2B histone protein component of the 200-kDa nucleosome core particle [49].

NMR spectroscopy also offers the possibility to measure translational diffusion, i.e., the tendency of a molecule to move under the influence of either a concentration gradient or Brownian motion, which can be very informative, especially when the experiment is performed under conditions near physiological situations. NMR diffusiometry spectroscopy is a very useful tool to study the dynamics of oligomerization and complex formation by biomolecules, e.g., complex formation between proteins and nucleic acids [50].

4. Cryogenic electron microscopy (cryo-EM)

In the last decade, the importance of cryo-EM has increased rapidly as a method for determining biomacromolecular structure at near-atomic resolution without the need for crystallization [51]. This progress mainly results from advances in microscopy and detector technology.

Cryo-EM involves using transmission electron microscopy to observe samples by applying very low temperature (cryogenic) conditions [52]. The principle of cryo-EM is to image biological macromolecules, such as proteins and protein complexes, which are frozen and fixed in glassy ice; thus, a projection of protein molecules is obtained in all directions. Then, a computer is applied to process and calculate a large number of 2D images and, based on these 2D images, the three-dimensional (3D) structure of the molecule is reconstructed [53]. The first electron microscope was constructed by E. Ruska and M. Knoll in the 1930 s. The general concept and a method of the 3D reconstruction of molecules were proposed by D. De Rosier and A. Klug in 1968, who reported the first 3D structure of a very large macromolecular complex, the T4 phage tail [53]. However, the biggest breakthrough was the development of cryogenics and rapid freezing technology by J. Dubochet and A. McDowal. J. Dubochet, J. Frank, and R. Henderson were awarded the 2017 Nobel Prize in chemistry for the development of cryo-EM [54]. The first structure determined by cryo-EM was published in 1990, and it was bacterial rhodopsin (resolution 7.5 Å) [55]. As of February 2023, the record of the cryo-EM structure resolution was 1.5 Å for apoferritin [56].

One of the advantages of cryo-EM is the possibility of obtaining the structure of protein at near-atomic resolution using a very small amount of sample (mostly 4 µl) with a concentration of 1 µg/µl for soluble proteins and 5 µg/µl for membrane proteins. However, cryo-EM is used for large proteins (> 100 kDa). For smaller complexes, the signal-to-noise ratio of cryo-EM images is problematic (the micrographs are noisy). To overcome this problem, the target protein can be linked to a larger protein scaffold (e.g., antibody) to increase its apparent size and therefore the image contrast.

Undoubtably, the cryo-EM method contributes significantly to knowledge on the structure of large macromolecular assemblies (Fig. 1). Research using cryo-EM has intensified over the past few years, resulting in the appearance of many macromolecular structures that were previously unsolvable by crystallographic methods. Examples of such macromolecular structures are the structures of the Dicer-type ribonucleases' complexes with RNA substrates.

Accordingly, in 2018, Z. Liu et al. published the cryo-EM structure of human Dicer in complex with its cofactor protein transactivation response element RNA-binding protein (TRBP) [57]. The authors revealed the precise spatial arrangement of human Dicer's multiple domains. They also confirmed the structures of a Dicer-TRBP complex bound with microRNA (miRNA) precursor, pre-miRNA, in two distinct conformations. Along with biochemical analysis, the results from this study provide insights into the mechanism of RNA processing by Dicer ribonucleases.

In 2021, X. Wei et al. reported single-particle cryo-EM structures of the Arabidopsis Dicer-like protein DCL1 complexed with a pri-miRNA and a pre-miRNA in cleavage-competent states. These structures uncovered the plasticity of the PAZ domain, which is critical for the recognition of both pri-miRNA and pre-miRNA in Dicer-like proteins (DCLs) [58]. Additionally, in 2021, Q. Wang et al. published the structure of the Arabidopsis DCL3-double-stranded RNA (dsRNA) complex in an active, dicing-competent state. These structural studies, complemented by functional data, provided insight into the mechanism of RNA cleavage by Dicer-type proteins [57].

The year 2022 was also fruitful for cryo-EM Dicer studies; (i) K. Jouravleva et al. developed a model of how the Drosophila Dicer-1-Loqs-PB complex influences the full cycle of pre-miRNA recognition, stepwise endonuclease cleavage and product release [59], and (ii) S. Su et al. revealed the molecular mechanism for the full cycle of ATP-dependent dsRNA processing by the Drosophila Dicer-2-Loqs-PD complex [60].

Cryo-EM was particularly useful when the COVID-19 pandemic broke out. Through this technique, the structure of replicating SARS-CoV-2 polymerase with the RNA template–product duplex was rapidly determined [61–63] (Fig. 2). Detailed knowledge of the virus structure and its replication mechanism is crucial for the development of drugs targeting the SARS-CoV-2 polymerase complex and thereby inhibiting virus replication. Examples of antiviral drugs targeting viral polymerases are remdesivir and favipiravir, which are nucleoside analogs. The cryo-EM structure of the SARS-CoV-2 polymerase in complex with the RNA template-primer duplex and

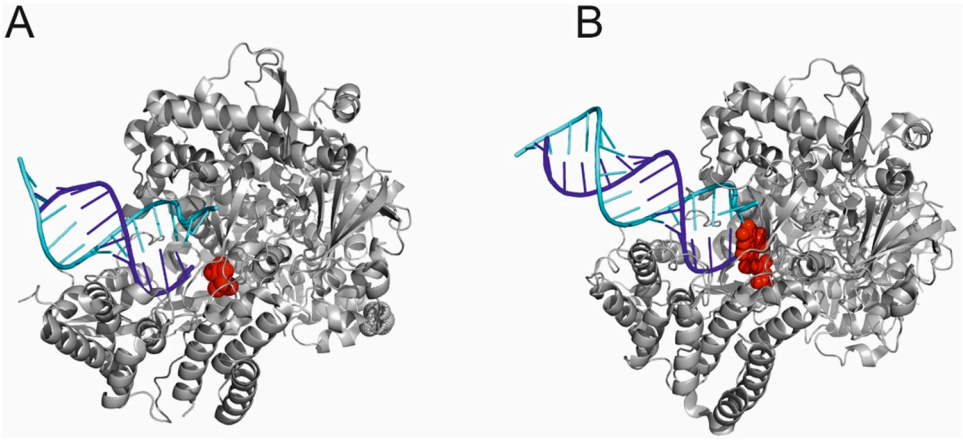


Fig. 2. Cryo-EM structures of the replicating SARS-CoV-2 polymerase complex. (A) Cryo-EM structure of SARS-CoV-2 polymerase in complex with the RNA template-primer duplex and remdesivir (PDB entry 7BV2). (B) The cryo-EM structure of favipiravir bound to the replicating SARS-CoV-2 polymerase complex (PDB entry 7CTT). SARS-CoV-2 polymerase is indicated in gray, remdesivir or favipiravir in red, the template strand in cyan, and the primer strand in blue. The 3D structures were visualized by PyMOL.

remdesivir revealed that remdesivir is covalently incorporated into the primer strand at the first replicated base pair, thus terminating chain elongation (Fig. 2A) [64,65]. In 2021, the cryo-EM structure of favipiravir bound to the replicating SARS-CoV-2 polymerase complex was reported [66]. These data revealed an unexpected base-pairing pattern between favipiravir and pyrimidine residues that may explain the capacity of favipiravir to mimic both adenine and guanine nucleotides (Fig. 2B).

The progress achieved in cryo-EM techniques is nicely reflected by the increase in the resolution of the molecular structures determined by this technique, which occurred with advances in microscope and detector technology. This can be exemplified by the cryo-EM structure of the DNA-dependent protein kinase (DNA-PK) complex. DNA-PK is a large protein complex central to the non-homologous end joining DNA repair pathway. It comprises the DNA-PK catalytic subunit (DNA-PKcs) and the heterodimer of DNA-binding proteins Ku70 and Ku80 [67]. First, in 1998, Y. Chiu et al. presented the structure of the DNA-PKcs complex at ~21 Å resolution [66]. Ten years later, D. Williams et al. revealed a 7 Å resolution structure of the same complex [68]. Then, in 2017, H. Sharif et al. reported the structures of human DNA-PKcs at 4.4 Å resolution and the DNA-PK holoenzyme at 5.8 Å resolution [67]. Through these data, a structural model for the DNA-PK interaction with DNA was proposed.

5. Atomic Force Microscopy (AFM)

The structure of the protein-nucleic acid complexes can also be investigated by using the atomic force microscopy (AFM). This technique was invented by G. Binning, C.F. Quate, and Ch. Gerber in 1986 [69]. AFM is a type of a scanning probe microscopy technique that enables visualization of the topography of the tested sample at a very high resolution (~1 Å). The images are obtained by using a cantilever with a sharp nanometer-sized tip (probe) at its end that is used to scan a surface of the sample [70]. The great advantage of this method is that (i) only very small amounts of the sample are needed for testing, (ii) samples do not require special preparation (e.g., labeling) [71], (iii) samples can be analyzed under various conditions: in liquid, air or ultrahigh vacuum [72], (iv) it is a nondestructive technique [73]. AFM can be used for force measurements (e.g., the bond strength between proteins and nucleic acids), topographic imaging (spatial mapping of the sample surface), and manipulation on single molecules or even single living cells [74]. Significant limitations of AFM are (i) the restriction of this technique to the sample surfaces, (ii) the small area of the imaging (a typical scanning area is ~ 150 µm x 150 µm) and (iii) the depth of the field (a typical depth of field is 10–20 µm, but it significantly decreases in high-resolution scanning mode) [75]. Moreover, the construction of the scanning tip limits thorough measurements of steep walls and overhangs of the sample [76]. A summary of the basic principles, advantages and limitations of the most common AFM bioimaging modes were provided, e.g., by Y. Dufrêne et al. [77], Y. Suzuki et al. [78], Y. Pan et al. [79], I. Volokhina et al. [80].

The first usage of AFM for imaging of protein-nucleic acid complexes was reported by J. Yang et al. in 1992; specifically, the authors visualized M13 phage DNA and the complex of M13-DNA with DNA polymerase [81]. AFM has also been used for determination of protein-nucleic acid binding constants and specificities: in 2005, Y. Yang et al. provided a detailed analysis of the DNA Mismatch Protein (MutS)–DNA interactions [82]. Later, in 2012, J. Yeh et al. demonstrated the use of AFM for assessing the stoichiometry of the complexes between the damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) and DNA duplexes. UV-DDB is part of a complex that initiates the nucleotide-excision repair (NER) pathway by recognizing damaged chromatin. The authors revealed that the dimeric UV-DDB can simultaneously bind to two DNA duplexes [83]. Moreover, AFM-based techniques can also be applied to visualize how proteins and nucleic acids assemble complex biological architectures. For example, in 2017, M. Shibata et al. demonstrated with high-speed AFM (HS-AFM) the real-space and real-time dynamics of CRISPR-Cas9 in action, including Cas9-RNA interactions and the Cas9-mediated DNA cleavage process [84]. Some more examples of AFM significant contributions to the study of protein-nucleic acid interactions, including information on protein binding specificity and affinity, protein binding stoichiometry, conformational changes in the nucleic acid structure induced by protein binding, complex conformation, and cooperativity, were extensively reviewed by, e.g., E. Beckwitt et al. [85], K. Main et al. [70].

6. Small angle scattering (SAS)

The SAS method was introduced in the 1930 s by A. Guinier [86], and the basics of the theory were developed in the 1940 s by Guinier, Krakty, Porod and Debye [86]. SAS is a powerful method for analyzing the structure of numerous biomolecules, including proteins, nucleic acids, and their complexes. Both X-rays and neutrons can be used for small angle scattering experiments. Accordingly, depending on the source of the scattering particles, small angle X-ray scattering (SAXS) and small angle neutron scattering (SANS) can be distinguished. The physical basis of both methods is similar, and the

only difference is the mechanism of interactions of photons and neutrons with matter.

The setup of an SAS experiment is very simple; a tested sample (an aqueous sample, i.e., a solution of molecules) and a control comprising a pure solvent are usually placed in a quartz capillary and are illuminated by a collimated monochromatic X-ray beam or neutrons (a collimated beam means a radiation beam with parallel rays, and a monochromatic beam means a beam with one wavelength; a typical wavelength used in SAXS is ∼1Å) [87]. Then, the scattered beams collected from a sample and a pure solvent are recorded. In the next step, the scattering pattern of the pure solvent is subtracted from the scattering pattern of the sample solution, leaving only the signal collected for the sample. The obtained scattering pattern provides information about the overall shape and size of the tested molecules. Due to the random orientations of molecules in a solution, the scattering pattern is isotropic (i.e., uniformly distributed in all orientations); thus, the scattering pattern, which is usually recorded by a two-dimensional detector, can be radially averaged [87,88]. A schematic of the method used to analyze the SAS data, on the example of the amino (N)-terminal domain of the restriction endonuclease McrA from *Escherichia coli* in complex with the 12 base-pair DNA duplex [89], is presented in Fig. 3.

6.1. Small angle X-ray scattering (SAXS)

The first experiment in which a protein structure was characterized by SAXS was performed in 1950. However, the widespread use of this method only occurred after the computerization and development of strong X-ray sources. In particular, a few high-performance synchrotrons have been developed in the last two decades, and this technology has resulted in the intensive development of the SAXS technique [88]. Due to greater X-ray availability, the SAXS method is more commonly used than SANS.

Through SAXS, structural parameters of molecules can be obtained, such as (i) the radius of gyration (a characteristic dimension of the molecule in solution that determines its shape and the distribution of mass), (ii) maximum diameter, (iii) volume, and (iv) molecular weight. Through this method, the homogeneity of the sample, a tendency toward oligomerization of the molecules, and a qualitative analysis of the degree of ordering of the molecule can also be evaluated (the Kratky plot gives quantitative information about flexibility and/or degree of unfolding of molecules).

For example, SAXS studies of protein-nucleic acid complexes showed that poly(C)-binding protein 2 (PCBP2) can regulate the function of p53 mRNA; in this case, the analysis of the exact PCBP2 binding site on p53 mRNA was supported by electrophoretic mobility shift assays [90]. SAXS was also successfully applied to study the structure of intrinsically disordered proteins (IDPs) [91,92]. An example of such a protein is the human RNA helicase DDX21 [93]. In this case, M. Marcaida et al. showed different orientations of DDX21 domains, depending on RNA binding. Moreover, analysis of DDX21 complexes with long RNA oligonucleotides revealed that the oligomerization of DDX2 is responsible for its activity [93].

Moreover, through SAXS, measurements can be obtained in different buffers under near-native conditions and under different temperature and pressure conditions. Importantly, the development of high flux sources provided the opportunity to analyze the dynamics of complex formation, e.g., analysis of complex formation between proteins and nucleic acids by time-resolved SAXS [94]. A combination of size-exclusion chromatography and SAXS, called SEC-SAXS, is a good method to study biomolecules that tend to aggregate. In a typical SEC-SAXS experiment, data are collected while the eluate of a size exclusion column flows through the SAXS sample cell [95].

SAXS is also used as a supporting method for crystallographic methods, NMR spectroscopy, CD spectroscopy and electron microscopy. For example, M. Meier et al. [96] and E. Ariyo et al. [97] combined NMR spectroscopy, SAXS and CD spectroscopy for the analysis of human RNA helicase DHX36 in complex with RNA G-quadruplexes.

The advent of the XFELs introduced a new modification of SAXS, i.e., fluctuation X-ray scattering (FXC) [98]. Generally, the basics of FXC are similar to those of SAXS experiments. The most significant difference is the use of femtosecond and extremely intense pulses of X-rays. On the femtosecond time scale, molecules are virtually frozen in space and time. This time scale enables structural details to be reconstructed with greater confidence than that with SAXS data alone [99].

6.2. Small angle neutron scattering (SANS)

Generally, the basics of the SANS experiment are similar to those of the SAXS experiment, but in SANS, neutrons are used for small angle scattering, while in SAXS, X-rays are applied. X-rays interact with electrons, which is only atoms heavier than the hydrogen atom are visible in SAXS experiments. In practice, this means that hydrogen atoms are not visible in SAXS experiments. Hydrogens are visible in SANS experiments; this can be explained by the strong interaction between the neutrons and hydrogen nuclei [100]. When neutrons interact with nuclei, the neutron scattering lengths depend, in an irregular way, on the atomic number. This feature is the basis for the application of phase contrast.

Neutrons have been used for biological research since the 1970 s, with the advent of high-flux neutron sources; neutrons can be produced either by nuclear fission or by spallation (high-energy nuclear reaction in which a target nucleus, struck by a bombarding particle, ejects numerous lighter particles) [101]. The development of SANS experiments was correlated with increasingly advanced computerization and progress in algorithms for SANS data analysis [102]. In contrast to SAXS, the SANS technique exhibits certain advantages, including the following: (i) neutrons do not cause radiation damage, while X-rays do, (ii) the choice of the buffer is more flexible (high salt concentration buffers do not significantly diminish the signal-to-noise ratio), (iii) the contribution of a hydration shell of a different density, as the bulk solvent, is less pronounced, and (iv) contrast variations can be used to focus on structural features of specific subunits in macromolecular complexes formed by several partners.

As mentioned above, the applicability of phase contrast during SANS experiments is possible due to the difference in the neutron interaction with different nuclei. For example, hydrogen and deuterium scatter neutrons very differently; thus, samples can be deuterated to distinguish specific areas or species of interest. Moreover, by replacing hydrogen with deuterium, the contrast of the sample relative to the solvent can be tuned. Several important observations can be implied directly from the dependence between neutron scattering length densities (SLD) and the percentage of the deuterated water in the solution. Hydrogenated RNAs and proteins exhibit different natural contrasts with respect to aqueous solvents. The contrast at a given H_2O:D_2O ratio depends on the exact amino acid or nucleotide sequence of a given protein or RNA. Deuteration homogenizes the respective SLD of proteins and RNAs, both on average and for individual amino acids and nucleotides [103]. The rules described in this paragraph also apply for neutron crystallography experiments.

Finally, notably, there is a data bank dedicated to SAXS and SANS data called the Small Angle Scattering Biological Data Bank (SASBDB). SASBDB was released to the general public in August 2014,

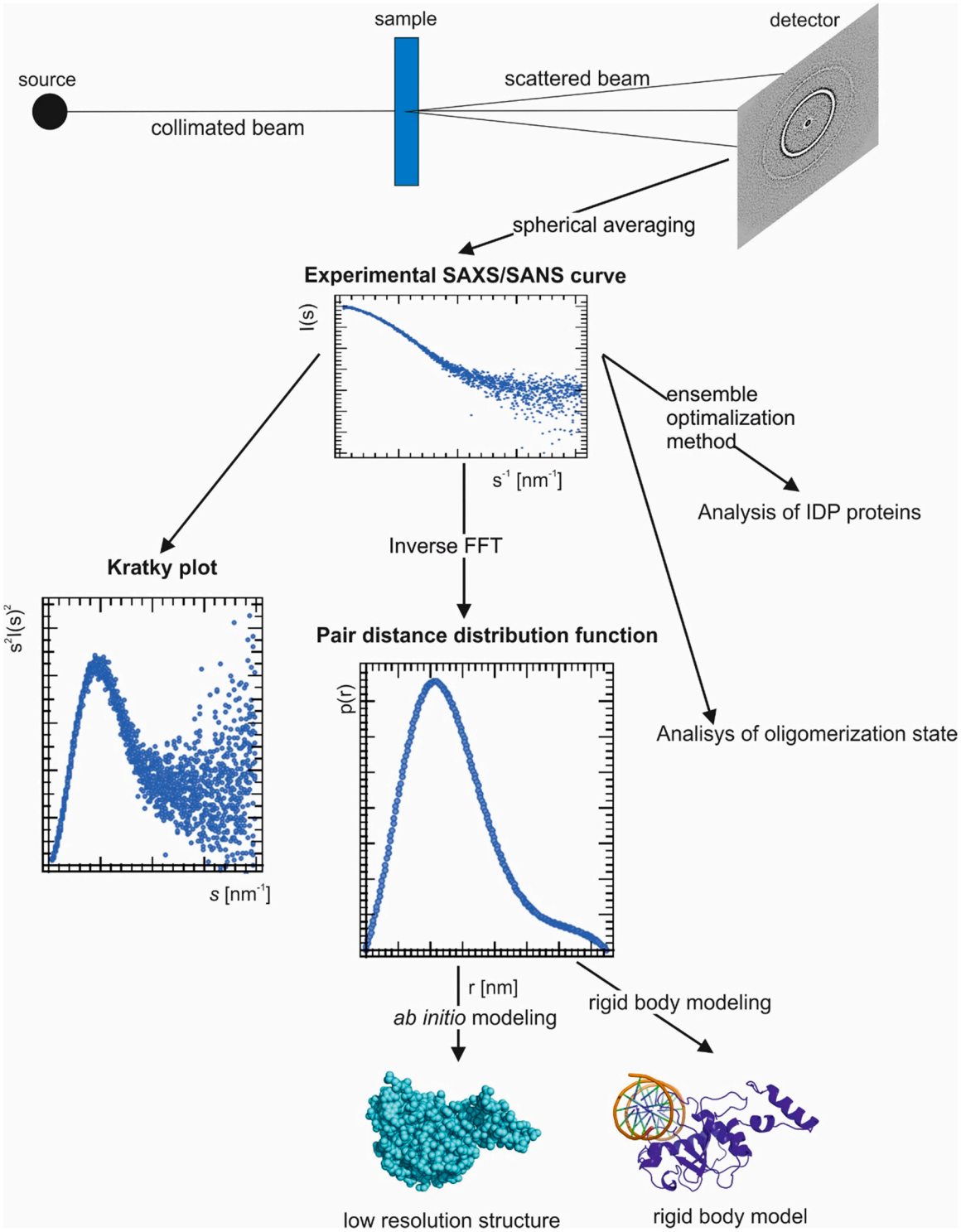


Fig. 3. Algorithm for analyzing SAXS/SANS data. Scheme of the small angle scattering experiment; the 2D scattering pattern is isotropic and can be radially averaged in the 1D scattering curve (shown in the figure as the experimental SAXS/SANS curve). Information on the radius of gyration is found at the smallest angles in the "Guinier region" (not shown in the figure). Guinier analysis is also used as quality control. The Kratky plot can quantitatively assess the flexibility or the degree of unfolding of the tested proteins and nucleic acids. Fully structured, compact particles are characterized by a bell-shaped Kratky plot with a well-defined maximum (as presented in the figure). Unfolded proteins and nucleic acids have a plateau in the Kratky plot at a high value of wave vector *s*. Partially unfolded (flexible) particles may show a combination of the bell shape and plateau plot or a plateau plot that slowly decays to zero. The Fourier transform (FFT) of the scattering curve provides the pair distance distribution function, which is a representation of the shape of the molecule in real space. The *p(r)* function is a histogram of all pairwise distances *r* between two scattering elements in the macromolecules weighted by their electron density contrast. The maximum distance inside the particle can be calculated from the *p(r)* function. The example presented in the figure refers to the N-terminal domain of the restriction endonuclease McrA from *Escherichia coli* in complex with the 12 base-pair DNA duplex [89] (SASBDB entry: SASDJZ2). For analysis of protein-nucleic acid complexes with known crystal structures it is possible to use the rigid body modeling approach.

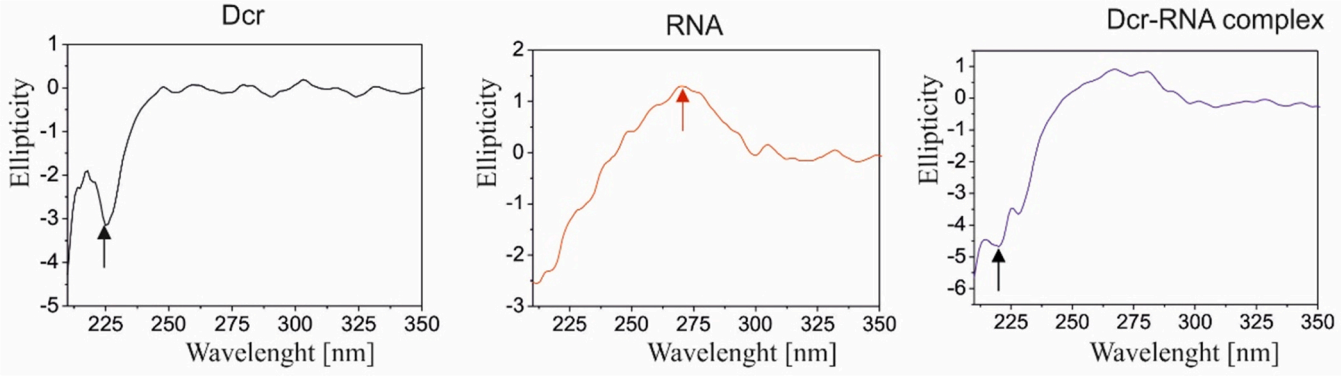


Fig. 4. Exemplary CD analysis for the protein-nucleic acid complex. CD analysis of the fragment of human Dicer (Dcr) in the complex with 42-nt RNA adopting a partial dsRNA structure. Exemplary CD spectra for Dcr (black), RNA (red), and the Dcr-RNA complex (blue). Analyzing the shape of a CD spectrum provides information on the structure of protein-nucleic acid complexes. CD data are commonly reported as ellipticity (θ), i.e., deviation (flattening) of an ellipse from the form of a circle or a sphere, usually reported in millidegrees (mdeg). A comparison of the shapes of the CD spectra, in a spectral range of 210–350 nm, was generated for the tested samples. Notably, the fragment of Dcr, that was used in the experiment, is mostly composed of α-helices. The ellipticity at 222 nm is routinely used to determine the α-helical content of a protein, and the double-stranded helical regions of RNA give a positive peak at ~270 nm. The CD spectrum of the Dcr protein shows a negative minimum value at 222 nm (indicated by a black arrow), which is associated with the presence of α-helices. In contrast, the CD spectrum of RNA has a maximum value at ~270 nm (indicated by a red arrow), which is characteristic of dsRNA structures. The CD spectrum of the Dcr-RNA complex exhibits a negative minimum value at 222 nm, typical for a protein with a dominant α-helical structure (indicated by a black arrow); however, the maximum value for dsRNA, at ~270 nm, is flattened. This result could be explained by the interactions between Dcr and RNA.

and as of February 2023, this data bank contained 3235 experimental datasets and 4460 models.

7. Circular dichroism (CD) spectroscopy

CD spectroscopy is a widely used technique in biochemistry, structural biology and pharmaceutical chemistry [104]. The CD phenomenon was discovered in the first half of the 19th century by J. Babptiste Biot, A. Fresnel and A. Cotton. CD spectroscopy is based on the differential absorption of left (L)- and right (R)-handed (i.e., circularly polarized) light by optically active molecules, such as chiral molecules. Chirality is a property of asymmetry, and a biomolecule is chiral if it is distinguishable from its mirror image (i.e., it cannot be superimposed onto it) [105].

Proteins and nucleic acids, which are composed of amino acids and nucleotides, respectively, are examples of naturally occurring chiral substances. All amino acids except glycine have at least one chiral center at their alpha-carbon in the peptide chain. According to the D-/L- naming convention, naturally occurring amino acids are found in the L-configuration. DNA has chiral centers in the atoms C1', C3', and C4' (in D-deoxyribose), while RNA has an additional chiral center at C2' (in D-ribose). Moreover, there are two other sources of chirality in nucleic acids, including (i) the helicity of the secondary structure and (ii) in some conditions, the long-range tertiary ordering of nucleic acids.

The first analysis of proteins by CD spectroscopy was performed in 1960. CD spectroscopy can be used to investigate the secondary structure of proteins based on electronic transitions in the far ultraviolet wavelength region (UV CD). Depending on the radiation source, the wavelength ranges from 260 to 190 nm (Xenon lamp) [106] or from 240 to 160 nm (synchrotron radiation) [107]. However, it must be stressed that the region below 200 nm causes difficulties related to the light source and optical device [106]. This problem can be overcome by synchrotron radiation, through which the CD spectra can be measured in the 160–260 nm region [108].

CD spectroscopy is very sensitive to conformational changes in the structure of macromolecules. It can provide information about the dynamics of folding of biomolecules such as proteins, nucleic acids, and their complexes. Importantly, proteins and nucleic acids generate specific CD spectra. The secondary structures of proteins can be determined in the far UV region (250–190 nm). Specifically, specific secondary structure motifs can be assigned to specific peaks in the CD spectrum, as follows: (i) random coil produce a positive peak at 212 nm and a negative peak at 195 nm, (ii) β-strand generates a positive peak at 196 nm and a negative peak at 218 nm, and (iii) α-helix produces a positive peak at 192 nm and two negative peaks at 222 and 208 nm [109]. For nucleic acids, CD spectroscopy can be used to determine the tertiary structures of DNA and RNA, such as helices, bulges, loops, and other mismatches, which contribute to complex tertiary structures. Each tertiary structure of RNA generates specific maxima in CD spectra, e.g., A-RNA has a positive peak near 260 nm and a negative peak near 210 nm. An exemplary CD spectra for the protein-nucleic acid complex is presented in Fig. 4. Other interesting tertiary structures are G-quadruplexes. G-quadruplexes are noncanonical structures formed by guanine-rich DNA and RNA molecules. They are organized in stacks of two or more G-quartets, in which four guanines are associated through Hoogsteen base pairing [110]. The following types of quadruplexes occur, depending on the strand arrangements: parallel, antiparallel, and hybrid: parallel–antiparallel. Parallel quadruplexes generate a dominating positive peak at 260 nm, and antiparallel quadruplexes produce a positive peak at 295 nm and a negative peak at 260 nm [111].

It must also be emphasized that NMR spectroscopy and SAS can be successfully used to determine the structure of G-quadruplexes. An interesting example is the 18-nucleotide oligomer BrG3:G6:BrG3:G6. The in-solution studies, carried out by SAXS and NMR spectroscopy, revealed the G-quadruplex fold for this oligomer, while crystallographic studies showed that this oligomer could adopt the duplex structure [112].

The CD technique has been used for over 50 years, and with the development of new radiation sources, various modifications to this method have been performed. One example is time-resolved CD, a method dedicated to studying the dynamics of conformational changes in proteins and structural changes in proteins and nucleic acids during their interactions [113]. This technique was first used in the 1970 s to analyze, with millisecond temporal resolution, photolysis processes in proteins [114] and biomolecular interactions [113]. Through the development of synchrotron radiation, biomolecules can be measured and analyzed on the picosecond time scale [113].

In 2009, the Protein Circular Dichroism Data Bank (PCDDB) was started (https://pcddb.cryst.bbk.ac.uk/). PCDDB is an online public repository that archives and distributes CD and synchrotron radiation CD (SRCD) spectral data and their associated experimental metadata. Moreover, the last two decades have enabled the development of computational approaches to predict protein secondary

Table 1
Maximum of the peak position, which characterizes the protein secondary structure in IR spectroscopy experiments.

Secondary structure	Band position in H_2O [cm^{-1}]		Band position in D_2O [cm^{-1}]	
	Average	Extremes	Average	Extremes
α-helix	1654	1648–1657	1652	1642–1660
β-strand	1633/1684	1623–1641, 1674–1695	1630/ 1679	1615–1638 1672–1694
Turns	1672	1662–1686	1671	1653–1691
Disordered	1654	1642–1657	1645	1639–1654

Based on A. Barth [117].

structure based on CD spectra. Examples of such approaches are K2D3, PDB2CD, SSNN, and BeStSel, which are web-based interfaces [115,116].

8. Infrared (IR) spectroscopy

IR spectroscopy is among the earliest methods used for studying the structure of biomolecules [117]. This method is a very useful tool for investigating protein and nucleic acid structures, the molecular mechanism of protein folding, protein enzymatic reactions, and the dynamics of the complexes [118].

The physical basis of this method involves the absorption of electromagnetic radiation from the IR range (780–1000 nm), which has a frequency similar to the frequency of vibration of molecules [119]. When passing through the sample of the tested substance, this radiation is absorbed by chemical groups of the molecule, increasing the amplitude of vibrations in the molecules of this substance. The absorption of IR radiation is accompanied by changes in the vibrational energy of the molecules. Since this energy is quantized, only radiation with certain specific energies, which is characteristic of the functional groups performing the vibrations (i.e., stretching, twisting or scissoring), is absorbed. The values of the frequency of characteristic vibrations can be presented in the form of tables (e.g., Table 1), and the absorption IR spectrum allows us to determine which functional groups are present in the analyzed sample.

The advantage of IR spectroscopy is the simplicity of obtaining spectra and the possibility of using phase contrast; this is possible because the frequency of the molecule vibration is strongly dependent on the weight of chemical groups. Isotope substitution can change the weight of the chemical group. This causes a change in the frequency, which is visible in a new peak on the spectrum. In practice, in protein-nucleic acid complex analysis, a common approach is the deuteration of nucleic acids. This causes the separation of the peaks of groups that bind with deuterated groups [120]. In the case of proteins, each secondary structure type can be assigned to a specific wavelength (summarized in Table 1). Because the position of the maximum of the peak is similar for α-helix and disordered fragments, buffers based on heavy water can be used to shift the maximum of the peak toward lower values (Table 1). Regarding nucleic acids, a typical region for characterizing their structures is between 1800 and 700 cm^{-1}.

IR spectroscopy was first used in the analysis of biomolecules in the mid-twentieth century; however, the widespread use of this method began after the Fourier transformed IR method was introduced, which accelerated the ability to obtain very good spectra. Large amounts of the sample and low resolution data are the main disadvantages of IR spectroscopy [117]. Fourier transform infrared (FTIR) nanospectroscopy (nano-FTIR) is a modification of IR spectroscopy that overcomes these disadvantages [121]. This technique is based on scattering-type scanning near-field optical microscopy (s-SNOM) [122], in which infrared images with nanoscale spatial resolution are obtained by recording the infrared light scattered at a scanning probe tip [123]. Through nano-FTIR, the protein and nucleic acid secondary structures can be analyzed on the nanometer scale and with sensitivity to the individual complexes. Moreover, through this technique, we can obtain the shape of molecules, as well as other parameters that can be obtained from AFM, e.g., the size and kinetics of folding. nano-FTIR was also adapted to study biomolecular complexes in living cells [124].

9. Summary and perspectives

Protein-nucleic acid complexes are central for the functioning of all living organisms. Determining the structure of individual protein-nucleic acid complexes is important not only for learning their specific functions but also, in the case of malfunctioning or harmful effects on the cell, for their putative application as therapeutic targets. Performing structural studies with protein-nucleic acid complexes is challenging, mainly because such types of complexes are often unstable, and their individual components may display extremely different surface properties (i.e., surface charges), which make the complex precipitate at higher concentrations used in many structural studies. Due to the variety of protein-nucleic acid complexes and their different biophysical properties, no simple and universal guideline is available to help scientists choose a method to successfully determine the structure of a specific protein-nucleic acid complex. In this manuscript, we provide a summary of the experimental methods that can be applied to structurally study protein-nucleic acid complexes (as summarized in Table 2).

Notably, however, each of the mentioned methods exhibits weaknesses and strengths, which is why the combination of different methods, as a hybrid approach, should be considered; in this way, specific problems encountered during studies of protein-nucleic acid complexes can be solved. This has been well illustrated by the example of the SARS-CoV-2 nucleocapsid phosphoprotein. The SARS-CoV-2 nucleocapsid phosphoprotein is an abundant RNA-binding protein that is a particularly attractive antiviral target because of its critical role in viral genome packaging [125]. Within 2020–2022, multiple papers were published revealing the atomic and molecular structure of the SARS-CoV-2 nucleocapsid phosphoprotein determined by X-ray crystallography. Nevertheless, the flexibility of the RNA-binding domains (the N-terminal domains) of this nucleoprotein was not revealed by X-ray crystallography alone. To support these structural studies, D.C. Dinesh et al. applied in-solution-state NMR spectroscopy, and they obtained a hybrid atomic model of the N-terminal domain of the SARS-CoV-2 nucleocapsid protein in complex with single-stranded and dsRNAs [126]. Additionally, L. Zinzula et al. utilized several methods involving X-ray crystallography, cryo-EM, size-exclusion chromatography coupled to SAXS and analytical ultracentrifugation, and differential scanning fluorimetry to obtain a high-resolution structure and perform biophysical characterization of the SARS-CoV-2 nucleocapsid phosphoprotein complex [125]. These studies proved that the N-terminal domains of the SARS-CoV-2 nucleocapsid protein are flexibly tethered to the carboxy (C)-terminal domain dimers and revealed that this nucleoprotein is largely disordered at physiological temperature due to the dynamic extension of its intrinsically disordered regions [127]. Moreover, combining cryo-EM and SAXS with CD spectroscopy allowed to determine the structure of the nucleoprotein of influenza D [120]. Cryo-EM was used to determine the 3D structure of the nucleoprotein, but a flexible N-terminal tail was not observed in that structure. To show the flexible parts of the analyzed nucleoprotein, SAXS and CD methods were applied.

Modern structural biology poses new challenges that can be achieved by new techniques. Notable, new techniques to observe the complexes formed between macromolecules in living cells have recently been developed. One of these methods is fluorescent confocal nanoscope in MINFLUX technology, which provides an unbeatable resolution of 1 nm in live cells. Through MINFLUX technology, the interaction and distance between the two molecules can be

Table 2
Characteristics of the experimental methods used for analysis of the biomolecules' structure.

Method	Source of radiation/ measurement tool	Radiation damage	Special requirements for the sample or buffer	Volume of the sample	Concentration of the sample	Contrast variation	Difficulty of sample preparation*	Type of information	Size limitation
X-ray crystallography	X-ray tube, synchrotrons	Yes	Good quality single crystal	**	**	No	+++	Structure with atomic resolution	No
Neutron crystallography	Nuclear reactor, spallation source available only in a large scale facility	No	Good quality single crystal	**	**	Yes	+++	Structure with atomic resolution	No
Serial femtosecond crystallography ***	X-ray free-electron laser	Yes	Large amounts of micro or nanocrystals	10–100 µl	1–30 µg/µl	No	+++	Structure with atomic resolution	No
Single particle imaging ***	X-ray free- electron laser	Yes	No	10–100 µl	2–30 µg/µl	No	+	Structure with atomic resolution	No
Nuclear magnetic resonance spectroscopy	Spectrometers	No	Sample should be labeled with $^{15}N/^{13}C$, measurement in D_2O	50–500 µl ****	5–20 µg/µl	No	+++	Structure with atomic resolution, dynamics of the protein and its complexes	For a standard analysis: < 50 kDa
Solid-state nuclear magnetic resonance spectroscopy	Spectrometers	No	Sample should be labeled with $^{15}N/^{13}C$	50–500 µl *****	5–20 µg/µl	No	+++	Structure with atomic resolution, dynamics of the protein and its complexes	No
Cryo-electron microscopy	Electron gun, cryo-EM available only in special centers	Yes	Glycerol must be avoided	1–3 µl	1–10 µg/µl	No	+++	Structure with near-atomic resolution	> 100 kDa
Atomic force microscopy	Scanning probe	No	No	1–5 µl	a single molecule is enough	No	+	Shape of single molecules, the mechanical properties of complexes	No
Small angle X-ray scattering	X-ray tube or synchrotron radiation available only in a large scale facility	Yes	No	10–100 µl	1–10 µg/µl	No	+	Low-resolution structure, dynamics of the protein and its complexes, qualitative description of the flexibility, structural parameters like radius of gyration, volume and long range fluctuations	No
Small angle neutron scattering	Nuclear reactor, spallation source available only in a large scale facility	No	For phase contrast measurements, the sample should be partially deuterated, buffer with a different D_2O content	10–100 µl	1–10 µg/µl	Yes	++	Low-resolution structure, dynamics of the protein and its complexes, qualitative description of the flexibility, structural parameters like radius of gyration, volume and long range fluctuations	No
Fluctuation X-ray scattering	X-ray free- electron laser	Yes	No	10–100 µl	1–20 µg/µl	No	+	Low-resolution structure, dynamics of the protein and its complexes, qualitative description of the flexibility, structural parameters like radius of gyration, volume and long range fluctuations	No
Circular dichroism spectroscopy	UV-lamp or synchrotron radiation available only in a large-scale facility	No	Limitation with the organic component of the buffer	10–100 µl	0.1–1 µg/µl	No	+	Quantitative content of the secondary structure, molecular dynamics at the level of the secondary structure, determination of thermodynamics parameters	No
Infrared spectroscopy	Silicon carbide rod or synchrotron radiation available only in a large-scale facility	No	Buffer with a different D_2O content	5–100 µl	5–30 µg/µl	Yes	+	Quantitative content of the secondary structure, molecular dynamics at the level of the secondary structure, determination of thermodynamics parameters	No

* Difficulty of sample preparation, on a three-point scale: simple +, moderate ++, difficult +++
** Volume and concentration of the sample are difficult to estimate because each complex needs specific buffer conditions and the sample concentration.
*** Techniques in the testing phase, parameters are approximate.
**** The presented volumes of the tested samples are approximate and depend on the used equipment.
***** Does not apply for solid state samples.

unambiguously identified and whether the two molecules (each only a few nm large) are next to each other (and thus interact) can be verified [128]. MINFLUX nanoscopy relies on a fluorescence on–off transition for the separation of neighboring emitters, similar to all fluorescence nanoscopy analyses [129]. MINFLUX nanoscopy was applied to image a pore protein of the injectisome of the enteropathogen *Yersinia enterocolitica* [130]; the MICOS complex, a large protein complex within the mitochondrial inner membrane [131]; in addition, the precise stepping motion of the motor protein kinesin-1 was imaged as it walks on microtubules in living cells [132]. Undoubtably, MINFLUX exhibits great potential for future research on proteins and their complexes with nucleic acids.

Below, we also reveal some numerical values for the structures of protein-nucleic acid complexes deposited in the PDB database or papers on the structure of protein-nucleic acid complexes (and nucleoproteins) published in the PubMed database.

Despite the great progress that has been made in structural research, the basic technique is still crystallography; as of February 2023, ~173,405 records for protein, nucleic acids and their complexes were found in the PDB database. As of February 2023, ~13,750 NMR records for protein, nucleic acids and their complexes were found in the PDB database. Regarding the cryo-EM structures, ~12,630 protein-nucleic acid complex structures were deposited in the PDB database, which represents ~7% of all structures. However, it must be emphasized that the cryo-EM technique was developed mostly during the last decade. In addition, between 1988 and 2022, over 11,500 papers demonstrating the imaging of proteins with AFM were published in the PubMed database, but among them only ~750 papers concerned protein-nucleic acid complex.

The SAXS technique has been developed since the 1960 s, but its greatest expansion occurred only at the beginning of the 21st century, when synchrotrons were launched or modernized. In the 1990 s, ~2540 papers were published in the PubMed database, in which SAXS methods were used to study proteins. The increase in the use of SAXS methods in protein research has been noticeable in the last two decades; between 2001 and 2010, ~9240 papers appeared in the PubMed database, and between 2011 and 2022, over 23,070 papers were published.

SANS requires special equipment and large-scale facilities, which are not easily available. Due to these restrictions, the SANS technique is not commonly used. In the years 1969–2022, in the PubMed database, 1264 records for SANS-based protein analysis appeared; within this time, only 171 papers were published from 1969 to 2000. However, we can observe a continuous increase in the use of this method. In the first decade of the 21st century, 266 papers were published in the PubMed database, while in the period from 2010 to 2022, this number grew to 827.

Likewise, the development of modern sources of radiation and automation contributed to the dissemination of CD and IR spectroscopy for the study of protein-nucleic acid complexes. In the case of CD spectroscopy, in the last decade of the twentieth century, ~5690 papers were published in the PubMed database, while from 2001 to 2022, ~29,240 papers appeared. In the case of IR spectroscopy, from 2001 to 2010, ~9290 papers were published in the PubMed database, while in the years 2011–2022, the number of published papers doubled and reached ~18,750.

The significant increase in the number of publications in the field of structural studies of biomolecular complexes is an excellent indicator of the growing interest in using the techniques presented in this manuscript to structurally and functionally analyze protein-nucleic acid complexes.

Finally, it must be underlined that a number of computational tools and methods that facilitate the study of the protein–nucleic acid interactions have already been developed [133]. In addition, there are computational tools for predicting structures of proteins [134] and nucleic acids [135]. However, despite significant recent advances in protein or nucleic acid structure prediction, the prediction of the structures of protein-nucleic acid complexes is still a largely unsolved problem. When predicting interactions between proteins and nucleic acids, two kinds of computational problems should be considered: (i) binding site prediction (taking into account the protein information, prediction which locations on the protein surface are the binding sites for RNAs or DNAs) and (ii) binding preference prediction (the RNA/DNA sequences that can bind to a specific protein have already been determined experimentally for some protein-nucleic acid complexes) [133]. For studying the interaction more comprehensively, it is advised to consider simultaneously the protein and nucleic acid information, including both the sequence and structural information, and predict both binding sites and binding preferences [133].

Because of the limitation of protein structure data, the majority of the computational methods rely only on the sequence data to perform the prediction. One of the most widely used tools that predicts a protein's 3D structure solely from its amino acid sequence is AlphaFold, an artificial intelligence system developed by DeepMind [136]. Considering the recent progress of AlphaFold, and the large amount of the structure data generated by this tool [136], the study of the interactions between proteins and nucleic acids will probably be supported soon by deep learning methods operating directly on the structure data [133]. There is no doubt that the recent huge progress in structural bioinformatics will translate in the near future into significant progress in the field of structural biology of protein-nucleic acid complexes. For example, the RoseTTAFoldNA tool is being recently developed to generate 3D structure models for protein-DNA and protein-RNA complexes [137]. Hopefully, the data generated by computational methods will help to fill in missing information in the 3D structures of protein-nucleic acid complexes that were already determined by the experimental methods.

CRediT authorship contribution statement

Kamil Szpotkowski: Conceptualization, Figure preparation (Figs. 1–4), Table preparation, Contribution to organizing the contents of the article, Writing the draft of the manuscript, Writing – review & editing. **Klaudia Wójcik:** Figure preparation (Fig. 2), Contribution to writing the draft of the manuscript, Writing – review & editing. **Anna Kurzyńska-Kokorniak:** Conceptualization, Organizing the contents of the article, Writing – review & editing, Responsibility for the final form of the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

The manuscript entitled *"Structural studies of protein–nucleic acid complexes: a brief overview of the selected techniques", by Kamil Szpotkowski, Klaudia Wójcik and Anna Kurzyńska-Kokorniak, has not been published and is not under consideration for publication elsewhere. We have no conflicts of interest to disclose.*

Acknowledgments

This work was supported by the National Science Centre, Poland [OPUS 2021/41/B/NZ2/03781 to A.K-K., and MINIATURA 2021/05/X/NZ1/01778 to K.S.]. We would like to thank Dr Kinga Ciechanowska for providing materials for Fig. 4 preparation.

References

[1] Akopian D, Shen K, Zhang X, et al. Signal recognition particle: an essential protein-targeting machine. Annu Rev Biochem 2013;Vol 82(82):693–721. https://doi.org/10.1146/annurev-biochem-072711-164732

[2] Tzlil S, Kindt JT, Gelbart WM, et al. Forces and pressures in DNA packaging and release from viral capsids. Biophys J 2003;84(3):1616–27. https://doi.org/10.1016/s0006-3495(03)74971-6

[3] Hogan DJ, Riordan DP, Gerber AP, et al. Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. Plos Biol 2008;6(10):2297–313. https://doi.org/10.1371/journal.pbio.0060255
[4] Sotomayor-Vivas C, Hernandez-Lemus E, Dorantes-Gilardi R. Linking protein structural and functional change to mutation using amino acid networks. Plos One 2022;17(1). https://doi.org/10.1371/journal.pone.0261829
[5] Jonchhe S, Pandey S, Emura T, et al. Decreased water activity in nanoconfinement contributes to the folding of G-quadruplex and i-motif structures. Proc Natl Acad Sci USA 2018;115(38):9539–44. https://doi.org/10.1073/pnas.1805939115
[6] Camilloni C, Bonetti D, Morrone A, et al. Towards a structural biology of the hydrophobic effect in protein folding. Sci Rep 2016(6). https://doi.org/10.1038/srep28285
[7] Kendrew JC, Bodo G, Dintzis HM, et al. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature 1958;181(4610):662–6. https://doi.org/10.1038/181662a0
[8] Perutz MF, Rossmann MG, Cullis AF, et al. Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis. Nature 1960;185(4711):416–22. https://doi.org/10.1038/185416a0
[9] Aggarwal AK, Rodgers DW, Drottar M, et al. Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. Science 1988;242(4880):899–907. https://doi.org/10.1126/science.3187531
[10] Yonath A. Approaching atomic resolution in crystallography of ribosomes. Annu Rev Biophys Biomol Struct 1992;21:77–93. https://doi.org/10.1146/annurev.bb.21.060192.000453
[11] Maveyraud L, Mourey L. Protein X-ray crystallography and drug discovery. Molecules 2020;25(5). https://doi.org/10.3390/molecules25051030
[12] McPherson A, Gavira JA. Introduction to protein crystallization. Acta Crystallogr F Struct Biol Commun 2014;70(Pt 1):2–20. https://doi.org/10.1107/S2053230×13033141
[13] Collins KD. Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. Methods 2004;34(3):300–11. https://doi.org/10.1016/j.ymeth.2004.03.021
[14] Hollis T. Crystallization of protein-DNA complexes. Methods Mol Biol 2007;363:225–37. https://doi.org/10.1007/978-1-59745-209-0_11
[15] Rosenbach H, Victor J, Borggräfe J, et al. Expanding crystallization tools for nucleic acid complexes using U1A protein variants. J Struct Biol 2020;210(2):107480. https://doi.org/10.1016/j.jsb.2020.107480
[16] Keenan RJ, Siehl DL, Gorton R, et al. DNA shuffling as a tool for protein crystallization. Proc Natl Acad Sci USA 2005;102(25):8887–92. https://doi.org/10.1073/pnas.0502497102
[17] Blakeley MP, Langan P, Niimura N, et al. Neutron crystallography: opportunities, challenges, and limitations. Curr Opin Struct Biol 2008;18(5):593–600. https://doi.org/10.1016/j.sbi.2008.06.009
[18] Adams P, Langan P. Opportunities and challenges with the growth of neutron crystallography. Acta Crystallogr D Biol Crystallogr 2010;66(Pt 11):1121–3. https://doi.org/10.1107/S0907444910039387
[19] Fenn TD, Schnieders MJ, Mustyakimov M, et al. Reintroducing electrostatics into macromolecular crystallographic refinement: application to neutron crystallography and DNA hydration. Structure 2011;19(4):523–33. https://doi.org/10.1016/j.str.2011.01.015
[20] Fan GY, Maldonado F, Zhang Y, et al. In vivo calcineurin crystals formed using the baculovirus expression system. Microsc Res Tech 1996;34(1):77–86. https://doi.org/10.1002/(sici)1097-0029(19960501)34:1<77::aid-jemt11>3.3.co;2-g
[21] Hand E. X-ray free-electron lasers fire up. Nature 2009;461(7265):708–9. https://doi.org/10.1038/461708a
[22] Coe J, Fromme P. Serial femtosecond crystallography opens new avenues for Structural Biology. Protein Pept Lett 2016;23(3):255–72. https://doi.org/10.2174/0929866523666160120152937
[23] Ekeberg T, Svenda M, Seibert MM, et al. Single-shot diffraction data from the Mimivirus particle using an X-ray free-electron laser. Sci Data 2016;3:160060. https://doi.org/10.1038/sdata.2016.60
[24] Li H, Nazari R, Abbey B, et al. Diffraction data from aerosolized Coliphage PR772 virus particles imaged with the Linac Coherent Light Source. Sci Data 2020;7(1):404. https://doi.org/10.1038/s41597-020-00745-2
[25] von Ardenne B, Mechelke M, Grubmüller H. Structure determination from single molecule X-ray scattering with three photons per image. Nat Commun 2018;9(1):2375. https://doi.org/10.1038/s41467-018-04830-4
[26] Bothwell JHF, Griffin JL. An introduction to biological nuclear magnetic resonance spectroscopy. Biol Rev 2011;86(2):493–510. https://doi.org/10.1111/j.1469-185X.2010.00157.x
[27] Ikeya T, Ban D, Lee D, et al. Solution NMR views of dynamical ordering of biomacromolecules. Biochim Et Biophys Acta-Gen Subj 2018;1862(2):287–306. https://doi.org/10.1016/j.bbagen.2017.08.020
[28] Lee JH, Okuno Y, Cavagnero S. Sensitivity enhancement in solution NMR: Emerging ideas and new frontiers. J Magn Reson 2014;241:18–31. https://doi.org/10.1016/j.jmr.2014.01.005
[29] Camacho-Zarco AR, Schnapka V, Guseva S, et al. NMR Provides Unique Insight into the Functional Dynamics and Interactions of Intrinsically Disordered Proteins. Chem Rev 2022;122(10):9331–56. https://doi.org/10.1021/acs.chemrev.1c01023
[30] Lacabanne D, Meier BH, Bockmann A. Selective labeling and unlabeling strategies in protein solid-state NMR spectroscopy. J Biomol Nmr 2018;71(3):141–50. https://doi.org/10.1007/s10858-017-0156-z
[31] Cavalli A, Salvatella X, Dobson CM, et al. Protein structure determination from NMR chemical shifts. Proc Natl Acad Sci USA 2007;104(23):9615–20. https://doi.org/10.1073/pnas.0610313104
[32] Ahlawat S, Mote KR, Lakomek NA, et al. Solid-State NMR: Methods for Biological Solids. Chem Rev 2022;122(10):9643–737. https://doi.org/10.1021/acs.chemrev.1c00852
[33] Opella SJ, Marassi FM. Applications of NMR to membrane proteins. Arch Biochem Biophys 2017;628:92–101. https://doi.org/10.1016/j.abb.2017.05.011
[34] Kang C. Applications of in-cell NMR in structural biology and drug discovery. Int J Mol Sci 2019;20(1). https://doi.org/10.3390/ijms20010139
[35] Carlomagno T. Present and future of NMR for RNA-protein complexes: A perspective of integrated structural biology. J Magn Reson 2014;241:126–36. https://doi.org/10.1016/j.jmr.2013.10.007
[36] Ahmed M, Marchanka A, Carlomagno T. Structure of a Protein-RNA Complex by Solid-State NMR Spectroscopy. Angew Chem-Int Ed 2020;59(17):6866–73. https://doi.org/10.1002/anie.201915465
[37] Schutz S, Sprangers R. Methyl TROSY spectroscopy: A versatile NMR approach to study challenging biological systems. Prog Nucl Magn Reson Spectrosc 2020;116:56–84. https://doi.org/10.1016/j.pnmrs.2019.09.004
[38] Arthanari H, Takeuchi K, Dubey A, et al. Emerging solution NMR methods to illuminate the structural and dynamic properties of proteins. Curr Opin Struct Biol 2019;58:294–304. https://doi.org/10.1016/j.sbi.2019.06.005
[39] Lapinaite A, Simon B, Skjaerven L, et al. The structure of the box C/D enzyme reveals regulation of RNA methylation. (+). Nature 2013;502(7472):519. https://doi.org/10.1038/nature12581
[40] Mandala VS, Williams JK, Hong M. Structure and Dynamics of Membrane Proteins from Solid-State NMR. Annu Rev Biophys 2018;Vol 47(47):201–22. https://doi.org/10.1146/annurev-biophys-070816-033712
[41] Lecoq L, Fogeron ML, Meier BH, et al. Solid-state NMR for studying the structure and dynamics of viral assemblies. Virus-Basel 2020;12(10). https://doi.org/10.3390/v12101069
[42] Tycko R. Solid-state NMR studies of amyloid fibril structure. Annu Rev Phys Chem 2011;Vol 62(62):279–99. https://doi.org/10.1146/annurev-physchem-032210-103539
[43] Marchanka A, Simon B, Althoff-Ospelt G, et al. RNA structure determination by solid-state NMR spectroscopy. Nat Commun 2015;6. https://doi.org/10.1038/ncomms8024
[44] Ahmed M, Marchanka A, Carlomagno T. Structure of a Protein-RNA Complex by Solid-State NMR Spectroscopy. Angew Chem Int Ed Engl 59. 2020. p. 6866–73. https://doi.org/10.1002/anie.201915465
[45] Sun SJ, Han Y, Paramasivam S, et al. Solid-State NMR Spectroscopy of Protein Complexes. In: Shekhtman A, Burz DS, editors. Protein Nmr Techniques. Third edition.,,2012. p. 303–31.
[46] Aguion PI, Marchanka A, Carlomagno T. Nucleic acid-protein interfaces studied by MAS solid-state NMR spectroscopy. J Struct Biol-X 2022(6). https://doi.org/10.1016/j.yjsbx.2022.100072
[47] Zehnder J, Cadalbert R, Terradot L, et al. Paramagnetic Solid-State NMR to Localize the Metal-Ion Cofactor in an Oligomeric DnaB Helicase. Chem-a Eur J 2021;27(28):7745–55. https://doi.org/10.1002/chem.202100462
[48] Lacabanne D, Boudet J, Malar AA, et al. Protein Side-Chain-DNA Contacts Probed by Fast Magic-Angle Spinning NMR. J Phys Chem B 2020;124(49):11089–97. https://doi.org/10.1021/acs.jpcb.0c08150
[49] Abramov G, Velyvis A, Rennella E, et al. A methyl-TROSY approach for NMR studies of high-molecular-weight DNA with application to the nucleosome core particle. Proc Natl Acad Sci USA 2020;117(23):12836–46. https://doi.org/10.1073/pnas.2004317117
[50] Dudas EF, Bodor A. Quantitative, diffusion NMR based analytical tool to distinguish folded, disordered, and denatured biomolecules. Anal Chem 2019;91(8):4929–33. https://doi.org/10.1021/acs.analchem.8b05617
[51] Milne JL, Borgnia MJ, Bartesaghi A, et al. Cryo-electron microscopy–a primer for the non-microscopist. FEBS J 2013;280(1):28–45. https://doi.org/10.1111/febs.12078
[52] Renaud JP, Chari A, Ciferri C, et al. Cryo-EM in drug discovery: achievements, limitations and prospects. Nat Rev Drug Discov 2018;17(7):471–92. https://doi.org/10.1038/nrd.2018.77
[53] Derosier DJ, Klug A. Reconstruction Of 3 Dimensional Structures From Electron Micrographs. -& Nature 1968;217(5124):130. https://doi.org/10.1038/217130a0
[54] Xu BJ, Liu L. Developments, applications, and prospects of cryo-electron microscopy. Protein Sci 2020;29(4):872–82. https://doi.org/10.1002/pro.3805
[55] Henderson R, Baldwin JM, Ceska TA, et al. Model For The Structure Of Bacteriorhodopsin Based On High-Resolution Electron Cryomicroscopy. J Mol Biol 1990;213(4):899–929. https://doi.org/10.1016/s0022-2836(05)80271-2
[56] Yip KM, Fischer N, Paknia E, et al. Atomic-resolution protein structure determination by cryo-EM. Nature 2020;587(7832):157. https://doi.org/10.1038/s41586-020-2833-4
[57] Liu ZM, Wang J, Cheng H, et al. Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. (+). Cell 2018;173(5):1191. https://doi.org/10.1016/j.cell.2018.03.080
[58] Wei X, Ke H, Wen A, et al. Structural basis of microRNA processing by Dicer-like 1. Nat Plants 2021;7(10):1389–96. https://doi.org/10.1038/s41477-021-01000-1
[59] Jouravleva K, Golovenko D, Demo G, et al. Structural basis of microRNA biogenesis by Dicer-1 and its partner protein Loqs-PB. Mol Cell 2022;82(21):4049. https://doi.org/10.1016/j.molcel.2022.09.002
[60] Su SC, Wang J, Deng T, et al. Structural insights into dsRNA processing by Drosophila Dicer-2-Logs-PD. Nature 2022;607(7918):399. https://doi.org/10.1038/s41586-022-04911-x

[61] Hillen HS, Kokic G, Farnung L, et al. Structure of replicating SARS-CoV-2 polymerase. (+). Nature 2020;584(7819):154. https://doi.org/10.1038/s41586-020-2368-8
[62] Wang Q, Wu JQ, Wang HF, et al. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. (+). Cell 2020;182(2):417. https://doi.org/10.1016/j.cell.2020.05.034
[63] Chen J, Malone B, Llewellyn E, et al. Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. (+). Cell 2020;182(6):1560. https://doi.org/10.1016/j.cell.2020.07.033
[64] Yin WC, Mao CY, Luan XD, et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. (+). Science 2020;368(6498):1499. https://doi.org/10.1126/science.abc1560
[65] Bravo JPK, Dangerfield TL, Taylor DW, et al. Remdesivir is a delayed translocation inhibitor of SARS-CoV-2 replication. (+). Mol Cell 2021;81(7):1548. https://doi.org/10.1016/j.molcel.2021.01.035
[66] Peng Q, Peng RC, Yuan B, et al. Structural Basis of SARS-CoV-2 Polymerase Inhibition by Favipiravir. 2021;2(1):Innovation. https://doi.org/10.1016/j.xinn.2021.100080
[67] Sharif H, Li Y, Dong Y, et al. Cryo-EM structure of the DNA-PK holoenzyme. Proc Natl Acad Sci USA 2017;114(28):7367–72. https://doi.org/10.1073/pnas.1707386114
[68] Williams DR, Lee KJ, Shi J, et al. Cryo-EM structure of the DNA-Dependent protein kinase catalytic subunit at subnanometer resolution reveals alpha helices and insight into DNA binding. Structure 2008;16(3):468–77. https://doi.org/10.1016/j.str.2007.12.014
[69] Binnig G, Quate CF, Gerber C. Atomic force microscope. Phys Rev Lett 1986;56(9):930–3. https://doi.org/10.1103/PhysRevLett.56.930
[70] Main KHS, Provan JI, Haynes PJ, et al. Atomic force microscopy-A tool for structural and translational DNA research. APL Bioeng 2021;5(3):031504. https://doi.org/10.1063/5.0054294
[71] Whited AM, Park PS. Atomic force microscopy: a multifaceted tool to study membrane proteins and their interactions with ligands. Biochim Biophys Acta 2014;1838(1 Pt A):56–68. https://doi.org/10.1016/j.bbamem.2013.04.011
[72] Rousso I, Deshpande A. Applications of Atomic Force Microscopy in HIV-1 Research. Viruses 2022;14(3). https://doi.org/10.3390/v14030648
[73] Dufrêne YF. Recent progress in the application of atomic force microscopy imaging and force spectroscopy to microbiology. Curr Opin Microbiol 2003;6(3):317–23. https://doi.org/10.1016/s1369-5274(03)00058-4
[74] Deng X, Xiong F, Li X, et al. Application of atomic force microscopy in cancer research. J Nanobiotechnol. 2018;16(1):102. https://doi.org/10.1186/s12951-018-0428-0
[75] Heath GR, Kots E, Robertson JL, et al. Localization atomic force microscopy. Nature 2021;594(7863):385–90. https://doi.org/10.1038/s41586-021-03551-x
[76] Ido S, Kobayashi K, Oyabu N, et al. Structured water molecules on membrane proteins resolved by atomic force microscopy. Nano Lett 2022;22(6):2391–7. https://doi.org/10.1021/acs.nanolett.2c00029
[77] Dufrêne YF, Ando T, Garcia R, et al. Imaging modes of atomic force microscopy for application in molecular and cell biology. Nat Nanotechnol 2017;12(4):295–307. https://doi.org/10.1038/nnano.2017.45
[78] Suzuki Y, Endo M, Sugiyama H. Studying RNAP-promoter interactions using atomic force microscopy. Methods 2015;86:4–9. https://doi.org/10.1016/j.ymeth.2015.05.018
[79] Pan Y, Shlyakhtenko LS, Lyubchenko YL. High-speed atomic force microscopy directly visualizes conformational dynamics of the HIV Vif protein in complex with three host proteins. J Biol Chem 2020;295(34):11995–2001. https://doi.org/10.1074/jbc.RA120.014442
[80] Volokhina I, Chumakov M. Study of the VirE2-ssT-DNA complex formation by scanning probe microscopy and gel electrophoresis- T-complex visualization. Microsc Micro 2007;13(1):51–4. https://doi.org/10.1017/S1431927607070158
[81] Yang J, Takeyasu K, Shao Z. Atomic force microscopy of DNA molecules. FEBS Lett 1992;301(2):173–6. https://doi.org/10.1016/0014-5793(92)81241-d
[82] Yang Y, Sass LE, Du C, et al. Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions. Nucleic Acids Res 2005;33(13):4322–34. https://doi.org/10.1093/nar/gki708
[83] Yeh JI, Levine AS, Du S, et al. Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. Proc Natl Acad Sci USA 2012;109(41):E2737–46. https://doi.org/10.1073/pnas.1110067109
[84] Shibata M, Nishimasu H, Kodera N, et al. Real-space and real-time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy. Nat Commun 2017;8(1):1430. https://doi.org/10.1038/s41467-017-01466-8
[85] Beckwitt EC, Kong M, Van Houten B. Studying protein-DNA interactions using atomic force microscopy. Semin Cell Dev Biol 2018;73:220–30. https://doi.org/10.1016/j.semcdb.2017.06.028
[86] Grawert TW, Svergun DI. Structural modeling using solution small-angle X-ray scattering (SAXS). J Mol Biol 2020;432(9):3078–92. https://doi.org/10.1016/j.jmb.2020.01.030
[87] Kikhney AG, Svergun DI. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. Febs Lett 2015;589(19):2570–7. https://doi.org/10.1016/j.febslet.2015.08.027
[88] Tuukkanen AT, Spilotros A, Svergun DI. Progress in small-angle scattering from biological solutions at high-brilliance synchrotrons. Iucrj 2017;4:518–28. https://doi.org/10.1107/s2052252517008740
[89] Czapinska H, Kowalska M, Zagorskaite E, et al. Activity and structure of EcoKMcrA. Nucleic Acids Res 2018;46(18):9829–41. https://doi.org/10.1093/nar/gky731
[90] Janecki DM, Swiatkowska A, Szpotkowska J, et al. Poly(C)-binding Protein 2 Regulates the p53 Expression via Interactions with the 5 '-Terminal Region of p53 mRNA. Int J Mol Sci 2021;22(24). https://doi.org/10.3390/ijms222413306
[91] Bernadó P, Svergun DI. Analysis of intrinsically disordered proteins by small-angle X-ray scattering. Methods Mol Biol 2012;896:107–22. https://doi.org/10.1007/978-1-4614-3704-8_7
[92] Cordeiro TN, Herranz-Trillo F, Urbanek A, et al. Small-angle scattering studies of intrinsically disordered proteins and their complexes. Curr Opin Struct Biol 2017;42:15–23. https://doi.org/10.1016/j.sbi.2016.10.011
[93] Marcaida MJ, Kauzlaric A, Duperrex A, et al. The Human RNA Helicase DDX21 Presents a Dimerization Interface Necessary for Helicase Activity. Iscience 2020;23(12). https://doi.org/10.1016/j.isci.2020.101811
[94] Blanchet CE, Spilotros A, Schwemmer F, et al. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). J Appl Crystallogr 2015;48:431–43. https://doi.org/10.1107/s160057671500254x
[95] Graewert MA, Franke D, Jeffries CM, et al. Automated Pipeline for Purification, Biophysical and X-Ray Analysis of Biomacromolecular Solutions. Sci Rep 2015(5). https://doi.org/10.1038/srep10734
[96] Meier M, Patel TR, Booy EP, et al. Binding of G-quadruplexes to the N-terminal Recognition Domain of the RNA Helicase Associated with AU-rich Element (RHAU). J Biol Chem 2013;288(49):35014–27. https://doi.org/10.1074/jbc.M113.512970
[97] Ariyo EO, Booy EP, Patel TR, et al. Biophysical Characterization of G-Quadruplex Recognition in the PITX1 mRNA by the Specificity Domain of the Helicase RHAU. Plos One 2015;10(12). https://doi.org/10.1371/journal.pone.0144510
[98] Kurta RP, Donatelli JJ, Yoon CH, et al. Correlations in Scattered X-Ray Laser Pulses Reveal Nanoscale Structural Features of Viruses. Phys Rev Lett 2017;119(15):158102. https://doi.org/10.1103/PhysRevLett.119.158102
[99] Pande K, Donatelli JJ, Malmerberg E, et al. Free-electron laser data for multiple-particle fluctuation scattering analysis. Sci Data 2018;5:180201. https://doi.org/10.1038/sdata.2018.201
[100] Hollamby MJ. Practical applications of small-angle neutron scattering. Phys Chem Chem Phys 2013;15(26):10566–79. https://doi.org/10.1039/c3cp50293g
[101] Krueger S. Small-angle neutron scattering contrast variation studies of biological complexes: Challenges and triumphs. Curr Opin Struct Biol 2022(74). https://doi.org/10.1016/j.sbi.2022.102375
[102] Lapinaite A, Carlomagno T, Gabel F. Small-Angle Neutron Scattering of RNA-Protein Complexes. Rna Spectrosc: Methods Protoc 2020;2113:165–88. https://doi.org/10.1007/978-1-0716-0278-2_13
[103] Bressler I, Kohlbrecher J, Thunemann AF. SASfit: a tool for small-angle scattering data analysis using a library of analytical expressions. J Appl Crystallogr 2015;48:1587–98. https://doi.org/10.1107/s1600576715016544
[104] Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. Biochim Biophys Acta 2005;1751(2):119–39. https://doi.org/10.1016/j.bbapap.2005.06.005
[105] Winogradoff D, Li PY, Joshi H, et al. Chiral systems made from DNA. Adv Sci (Weinh) 2021;8(5):2003113. https://doi.org/10.1002/advs.202003113
[106] Micsonai A, Wien F, Kernya L, et al. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc Natl Acad Sci USA 2015;112(24):E3095–103. https://doi.org/10.1073/pnas.1500851112
[107] Wallace BA, Janes RW. Synchrotron radiation circular dichroism (SRCD) spectroscopy: an enhanced method for examining protein conformations and protein interactions. Biochem Soc Trans 2010;38(4):861–73. https://doi.org/10.1042/BST0380861
[108] Miles AJ, Janes RW, Wallace BA. Tools and methods for circular dichroism spectroscopy of proteins: a tutorial review. Chem Soc Rev 2021;50(15):8400–13. https://doi.org/10.1039/d0cs00558d
[109] Kelly SM, Price NC. The use of circular dichroism in the investigation of protein structure and function. Curr Protein Pept Sci 2000;1(4):349–84. https://doi.org/10.2174/1389203003381315
[110] Ma Y, Iida K, Nagasawa K. Topologies of G-quadruplex: Biological functions and regulation by ligands. Biochem Biophys Res Commun 2020;531(1):3–17. https://doi.org/10.1016/j.bbrc.2019.12.103
[111] Kypr J, Kejnovská I, Renciuk D, et al. Circular dichroism and conformational polymorphism of DNA. Nucleic Acids Res 2009;37(6):1713–25. https://doi.org/10.1093/nar/gkp026
[112] Gudanis D, Popenda L, Szpotkowski K, et al. Structural characterization of a dimer of RNA duplexes composed of 8-bromoguanosine modified CGG trinucleotide repeats: a novel architecture of RNA quadruplexes. Nucleic Acids Res 2016;44(5):2409–16. https://doi.org/10.1093/nar/gkv1534
[113] Hache F, Changenet P. Multiscale conformational dynamics probed by time-resolved circular dichroism from seconds to picoseconds. Chirality 2021;33(11):747–57. https://doi.org/10.1002/chir.23359
[114] Chen E, Wood MJ, Fink AL, et al. Time-resolved circular dichroism studies of protein folding intermediates of cytochrome c. Biochemistry 1998;37(16):5589–98. https://doi.org/10.1021/bi972369f
[115] Miles AJ, Ramalli SG, Wallace BA. DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. Protein Sci 2022;31(1):37–46. https://doi.org/10.1002/pro.4153
[116] Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers 2008;89(5):392–400. https://doi.org/10.1002/bip.20853

[117] Barth A. Infrared spectroscopy of proteins. Biochim Biophys Acta 2007;1767(9):1073–101. https://doi.org/10.1016/j.bbabio.2007.06.004
[118] Geinguenaud F, Militello V, Arluison V. Application of FTIR Spectroscopy to Analyze RNA Structure. Methods Mol Biol 2020;2113:119–33. https://doi.org/10.1007/978-1-0716-0278-2_10
[119] Lorenz-Fonfria VA. Infrared Difference Spectroscopy of Proteins: From Bands to Bonds. Chem Rev 2020;120(7):3466–576. https://doi.org/10.1021/acs.chemrev.9b00449
[120] Donchet A, Oliva J, Labaronne A, et al. The structure of the nucleoprotein of Influenza D shows that all Orthomyxoviridae nucleoproteins have a similar NP. Sci Rep 2019;9(1):600. https://doi.org/10.1038/s41598-018-37306-y
[121] Amenabar I, Poly S, Nuansing W, et al. Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy. Nat Commun 2013;4:2890. https://doi.org/10.1038/ncomms3890
[122] Yong YC, Wang YZ, Zhong JJ. Nano-spectroscopic imaging of proteins with near-field scanning optical microscopy (NSOM). Curr Opin Biotechnol 2018;54:106–13. https://doi.org/10.1016/j.copbio.2018.01.022
[123] Yang Z, Tang D, Hu J, et al. Near-field nanoscopic terahertz imaging of single proteins. Small 2021;17(3):e2005814. https://doi.org/10.1002/smll.202005814
[124] Fisher TE, Oberhauser AF, Carrion-Vazquez M, et al. The study of protein mechanics with the atomic force microscope. Trends Biochem Sci 1999;24(10):379–84. https://doi.org/10.1016/s0968-0004(99)01453-x
[125] Zinzula L, Basquin J, Bohn S, et al. High-resolution structure and biophysical characterization of the nucleocapsid phosphoprotein dimerization domain from the Covid-19 severe acute respiratory syndrome coronavirus 2. Biochem Biophys Res Commun 2021;538:54–62. https://doi.org/10.1016/j.bbrc.2020.09.131
[126] Dinesh DC, Chalupska D, Silhan J, et al. Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. PLoS Pathog 2020;16(12):e1009100. https://doi.org/10.1371/journal.ppat.1009100
[127] Zeng W, Liu G, Ma H, et al. Biochemical characterization of SARS-CoV-2 nucleocapsid protein. Biochem Biophys Res Commun 2020;527(3):618–23. https://doi.org/10.1016/j.bbrc.2020.04.136
[128] Schmidt R, Weihs T, Wurm CA, et al. MINFLUX nanometer-scale 3D imaging and microsecond-range tracking on a common fluorescence microscope. Nat Commun 2021;12(1). https://doi.org/10.1038/s41467-021-21652-z
[129] Ostersehlt LM, Jans DC, Wittek A, et al. DNA-PAINT MINFLUX nanoscopy. Nat Methods 2022;19(9):1072–5. https://doi.org/10.1038/s41592-022-01577-1
[130] Carsten A, Rudolph M, Weihs T, et al. MINFLUX imaging of a bacterial molecular machine at nanometer resolution. Methods Appl Fluor 2022;11(1). https://doi.org/10.1088/2050-6120/aca880
[131] Pape JK, Stephan T, Balzarotti F, et al. Multicolor 3D MINFLUX nanoscopy of mitochondrial MICOS proteins. Proc Natl Acad Sci USA 2020;117(34):20607–14. https://doi.org/10.1073/pnas.2009364117
[132] Gerasimaitė RT, Bucevičius J, Kiszka KA, et al. Blinking Fluorescent Probes for Tubulin Nanoscopy in Living and Fixed Cells. ACS Chem Biol 2021;16(11):2130–6. https://doi.org/10.1021/acschembio.1c00538
[133] Wei J, Chen S, Zong L, et al. Protein-RNA interaction prediction with deep learning: structure matters. Brief Bioinform 2022;23(1). https://doi.org/10.1093/bib/bbab540
[134] Deng H, Jia Y, Zhang Y. Protein structure prediction. Int J Mod Phys B 2018;32(18). https://doi.org/10.1142/S021797921840009X
[135] Popenda M, Szachniuk M, Antczak M, et al. Automated 3D structure composition for large RNAs. Nucleic Acids Res 2012;40(14):e112. https://doi.org/10.1093/nar/gks339
[136] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596(7873):583–9. https://doi.org/10.1038/s41586-021-03819-2
[137] Minkyung B. Accurate prediction of nucleic acid and protein nucleic cid complex using RoseTTAFoldNA, 2022.