Research Article

Nucleic acid–protein interfaces studied by MAS solid-state NMR spectroscopy

Philipp Innig Aguion a, Alexander Marchanka a,b, Teresa Carlomagno c,*

a Institute for Organic Chemistry and Centre of Biomolecular Drug Research (BMWZ), Leibniz University Hannover, Schneiderberg 38, 30167 Hannover, Germany
b Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstr. 1, 69117 Heidelberg, Germany
c School of Biosciences/College of Life and Environmental Sciences, Institute of Cancer and Genomic Sciences/College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

ARTICLE INFO

Keywords:
sNMR
Nucleic acid–protein complexes
Intermolecular interfaces

ABSTRACT

Solid-state NMR (ssNMR) has become a well-established technique to study large and insoluble protein assemblies. However, its application to nucleic acid–protein complexes has remained scarce, mainly due to the challenges presented by overlapping nucleic acid signals. In the past decade, several efforts have led to the first structure determination of an RNA molecule by ssNMR. With the establishment of these tools, it has become possible to address the problem of structure determination of nucleic acid–protein complexes by ssNMR. Here we review first and more recent ssNMR methodologies that study nucleic acid–protein interfaces by means of chemical shift and peak intensity perturbations, direct distance measurements and paramagnetic effects. At the end, we review the first structure of an RNA–protein complex that has been determined from ssNMR-derived intermolecular restraints.

Introduction

Besides its long-established role as carrier of genetic information in protein translation, RNA acts in many other cellular contexts, with new roles being discovered regularly (Cech and Steitz, 2014). The vast majority of the RNAs made from the human genome have distinct functions from protein coding (non-coding RNAs, ncRNAs), but many of these functions remain unknown. Both coding RNA (mRNA) and ncRNAs may act in complex with specific RNA-binding proteins (RBPs), which contain well-defined RNA recognition motifs (Corley et al., 2020). Being involved in such high number of biological functions, RNA molecules hold potentials as both therapeutic agents and targets.

As for proteins, the development of RNAs either as therapeutic targets or as protein-targeting agents requires understanding their three-dimensional structures and their interaction modes with proteins. Methods to characterize RNA–protein interactions develop rapidly (Schlundt et al., 2017). X-ray crystallography and cryo electron microscopy (cryo-EM) can be applied to high molecular-weight complexes and have been the methods of choice to study many large ribonucleoprotein complexes (RNPs) in the past years (Ben-Shem et al., 2011; Ghanim et al., 2021; Jackson et al., 2014; Khatter et al., 2015; Nguyen et al., 2016). These techniques work well for RNPs with well-defined structures but fall short when addressing conformational heterogeneity or dynamic processes. RNA molecules can adopt different folds depending on both the environmental conditions and the interacting partners and are often flexible in isolation. The presence of disordered RNA regions can make crystallization of RNPs quite challenging (Blanco and Montoya, 2011), while cryo-EM analysis remains blind to disordered molecular regions. In addition, both X-ray crystallography and cryo-EM are unable to represent the dynamics of the studied system at atomic level, which, especially in enzymes, provides the crucial link between structure and function. Finally, several intermolecular interactions with functional regulatory roles are transient in nature; transient complexes are difficult to crystallize and may dissociate during the preparation of cryo-EM grids, rendering both X-ray crystallography and cryo-EM inapplicable.

NMR spectroscopy is a structural biology technique that is able to provide structural information in the presence of disorder and dynamics. As such, NMR is very useful to study RNPs, and more generally nucleic acid–protein complexes, containing flexible regions, and reveals whether and how these disordered regions contribute to binding specificity and/or modulate affinity. NMR spectroscopy also enables structural studies of transiently forming complexes in a wide range of affinities (Campagne et al., 2011; Carlomagno, 2014; Dominguez et al., 2016).
Table 1

| Nuclei | Spin quantum number | Natural abundance (%) | Gyromagnetic ratio $\gamma$ (10$^6$ rad T$^{-1}$ s$^{-1}$) | NMR transition frequency at 18.8 T (MHz) |
|--------|---------------------|----------------------|------------------------------------------------|-----------------------------------|
| $^1$H  | $\frac{1}{2}$       | 99.98                | 26.7519                                          | 800                               |
| $^1$H  | 1                   | 0.015                | 4.1066                                           | 123                               |
| $^{13}$C | $\frac{1}{2}$     | 1.1                  | 6.7283                                           | 200                               |
| $^{15}$N | $\frac{1}{2}$     | 0.37                 | $\sim$2.7126                                      | 80                                |
| $^{19}$F | $\frac{1}{2}$     | 100                  | 25.1815                                          | 753                               |
| $^{31}$P | $\frac{1}{2}$     | 100                  | 10.8394                                          | 324                               |

2011; Simon et al., 2011; Yadav and Lukavsky, 2016). Finally, NMR allows the direct observation of hydrogen atoms, which remain inaccessible by X-ray crystallography or cryo-EM but play a crucial role in nucleic acids interactions (Yip et al., 2020).

Notoriously, NMR studies in solution are limited to particles of less than ~ 50 kDa, due to the direct dependency of line-broadening on the molecular size. For proteins, this limit has been considerably extended by the methyl-TROSY methodology (transverse relaxation-optimized spectroscopy) (Kay, 2011; Rosenzweig and Kay, 2014; Tugarinov et al., 2003) in combination with selective $^{13}$CH$^3$ methyl group labeling of highly deuterated proteins. Methyl-TROSY NMR exploits the excellent relaxation properties of methyl groups, which, when present as only hydrogen-bearing groups in otherwise deuterated proteins, retain feasible NMR line widths even in particles as large as 1 MDa (Lapinaita et al., 2013; Mas et al., 2018, 2013; Sprangers and Kay, 2007; Grazier et al., 2020). Unfortunately, RNAs do not contain any methyl groups and NMR studies of large RNAs rely on a challenging combination of two-dimensional $^1$H–$^1$H correlation experiments measured on several selectively-deuterated samples (Brown et al., 2020; Keane et al., 2015).

Recently, methyl-TROSY has been applied to high-molecular weight DNA, where methyl groups were engineered at the C5 and N6 positions of cytosines (5mC) and adenines (6mA) (Abramov et al., 2020). However, because artificial methylation of both DNA and RNA can affect their structures, dynamics and interactions with binding partners (Choy et al., 2010; Helm, 2006; Kawai et al., 1992; Ngo et al., 2016; Williams et al., 2001), nucleic acids methylation cannot be generally applicable to study large nucleic acids by NMR. In fact, methylation of both DNA and RNA occurs naturally in the cell, where it exerts a regulatory function by modulating both the structure and the interactome of nucleic acids.

Solid-state NMR spectroscopy (ssNMR) is another form of biomolecular NMR spectroscopy, which has been applied extensively to insoluble and non-crystalline particles, such as membrane proteins (Cady et al., 2010; Lange et al., 2006; Park et al., 2012; Shahid et al., 2012; Shi et al., 2009) and amyloid fibrils (Colvin et al., 2016; Hoop et al., 2016; Tycko, 2011; Van Melckebeke et al., 2010). ssNMR linewidths do not depend on the molecular weight, making the application of ssNMR to large particles feasible, provided that there is enough signal to compensate for the small number of large particles that can be fitted into a rotor of a given size. In addition, because ssNMR lines are intrinsically broader than solution NMR lines, selective isotope labeling is often crucial to resolve spectral overlaps and achieve site-specific assignment also of moderately sized molecules. Despite these limitations, ssNMR has been applied to large viral assemblies (Andreas et al., 2016; Goldburt et al., 2007; Lusky et al., 2021; Morag et al., 2015, 2014; Sergeyev et al., 2011; Yu and Schaefer, 2008) and site specific structural information was obtained for the 46-residue-long major coat protein subunit of the filamentous bacteriophage Pf1, as part of the 36 MDa virion (Goldburt et al., 2007), thanks to the fact that the 7300 subunits of the virion all adopt the same conformation. Over the years, the ssNMR toolbox has been extended for the application to RNA in isolation (Lepper et al., 2004; Lusky et al., 2021; Riedel et al., 2006, 2005a, 2005b; Yang et al., 2017), RNA bound to short peptides (Huang et al., 2010, 2011, 2017, 2012).
mapping reports on changes in chemical shifts occurring in the protein upon binding of the nucleic acid, while temperature mapping measures the temperature dependence of $^1$H chemical shifts, which in turn depends on the involvement of the atom in hydrogen bonds. Comparison of temperature coefficients in the free and nucleic acid-bound states can reveal which protein residues are involved in nucleic acid binding. Quantification of peak intensities can also reveal which atoms are close to the partner molecule in the complex, because of binding-induced line-broadening effects.

### Table 2

Selection of ssNMR studies of intermolecular interactions in nucleic acid–protein complexes since 2005.

| Year   | Reference                        | Complex type                          | Site specific information | Protein (aa) | NA (nt) | ssNMR technique |
|--------|----------------------------------|---------------------------------------|---------------------------|--------------|---------|-----------------|
| 2005   | (Öhnen et al., 2005)             | 1:1 dsRNA–peptide                     | Yes                       | 11           | 29      | $^{31}$P, $^{15}$F REDOR |
| 2008   | (Yu and Schaefer, 2008)          | dsDNA–protein (intact bacteriophage)  | No                        | not known    | 342,000 | $^{15}$N, $^{31}$P, $^{2}$H, $^{1}$N REDOR |
| 2010   | (Jehle et al., 2010)             | 1:1 dsRNA–protein                     | Yes                       | 123          | 26      | $^{31}$P, $^{15}$N TEDOR |
|        | (Huang et al., 2010)             | 1:1 dsRNA–peptide                     | Yes                       | 11           | 29      | $^{13}$C, $^{15}$N REDOR |
| 2011   | (Huang et al., 2011)             | 1:1 dsRNA–peptide                     | Yes                       | 11           | 29      | $^{13}$C, $^{15}$N, $^{31}$P, $^{2}$H, $^{1}$N REDOR |
|        | (Sergeyev et al., 2011)          | ssDNA–protein (intact bacteriophage)  | Partially (only for coat protein) | 46           | 7349    | $^{13}$C, $^{15}$N DARR |
|        | (Bechinger et al., 2011)         | Oligomeric dsDNA–peptide              | No                        | 27           | not known | $^{15}$N, $^{31}$P REDOR |
| 2013   | (Asami et al., 2013)             | 1:1 dsRNA–protein                     | Yes                       | 123          | 26      | Intensity mapping (Dipolar-coupling-mediated line broadening) |
|        | (Morag et al., 2014)             | ssDNA–protein (intact bacteriophage)  | Partially (only for coat protein) | 50           | 8233    | $^{13}$C, $^{15}$C, CORD, DARR, PHHC |
| 2016   | (Wiegand et al., 2016)           | 2:12:12 ssDNA–protein–AMP–PNP        | Yes                       | 521          | 20      | CSP (1D $^1$H $^{31}$P $^{13}$C, $^{15}$C DARR) |
| 2017   | (Wiegand et al., 2017)           | 2:12:12 ssDNA–protein–ADP            | No                        | 521          | 20      | 2D PC-C, 2D PHHC-C |
| 2019   | (Wiegand et al., 2019)           | 2:12:12 ssDNA–protein–AMP–PCP/ADP:AMP/ADP:ATP/ADP | Yes | 521 | 20 | CSP (1D $^1$H $^{31}$P $^{13}$C, $^{15}$C DARR, 2D CP $^{13}$C–$^{15}$N), Intensity mapping ($^{13}$C, $^{15}$N DARR), NHNP CSP ($^{13}$C, $^{15}$N DARR, 2D CP $^{13}$C–$^{15}$N), CHHP, NHNP |
|        | (Boudet et al., 2019)            | 1:1:2 ssDNA–protein–protein–ATP      | Yes                       | 331          | 9       | |
| 2020   | (Wiegand et al., 2020)           | 2:12:12 ssDNA–protein–ADP:AMP–PCP    | Yes                       | 521          | 20      | NHNP, CHHP, CSP (CP-hNH), Intensity mapping ($^{13}$P, $^{2}$H, $^{1}$N DARR) |
|        | (Lacabanne et al., 2020)         | 1:1:2 ssDNA–protein–ATP              | Yes                       | 331          | 9       | CSP (CP-hNH, CP-hCH), Intensity mapping (CP-hNH) |
|        | (Ahmed et al., 2020)             | Monomeric dsRNA–protein               | Yes                       | 123          | 26      | PRE, CSP ($^{13}$C, $^{15}$C DARR, 2D CP $^{13}$C–$^{15}$N) |
| 2021   | (Zehnder et al., 2021)           | 2:12:12 ssDNA–protein–ADP:AMP–PCP    | Yes                       | 521          | 20      | PRE |
|        | (Malar et al., 2021)             | 2:12:12 ssDNA–protein–ADP:AMP–PCP    | Yes                       | 521          | 20      | CP-hPH, CSP temperature mapping (CP-hNH) |

**Fig. 2.** Overview of ssNMR methods utilized to date to probe nucleic acid–protein interfaces. Information derived from paramagnetic effects, cross-interface dipolar correlations, chemical shifts and peak intensities are highly complementary and can be utilized as restraints in a molecular docking protocol that builds the nucleic acid–protein complex from the structures of the individual components. Pseudo contact shifts (PCSs) measure changes in chemical shifts due to the influence of a paramagnetic ion with anisotropic tensor. Paramagnetic relaxation enhancement (PRE) measures the increase in relaxation of nuclei in the vicinity of an unpaired electron. Both effects can be quantified and translated into nucleus-electron distances. Cross-interface dipolar correlations directly measure intermolecular distances through a variety of dipolar recoupling techniques. The ssNMR toolbox for the measurement of intermolecular distances covers transferred echo double resonance (TEDOR)/rotational echo double resonance (REDOR)-based recoupling experiments, $^{13}$C–$^{15}$C correlation experiments, such as dipolar assisted rotational resonance (DARR) and combined E2E-driven (CORD), spin-diffusion based CHHP/NHHP experiments, and $^1$H-detected CP-based hPH experiments. Chemical shift perturbation (CSP) experiments.
RNA as part of RNP complexes (Aguion et al., 2021; Ahmed et al., 2020; Marchanka et al., 2013, 2015, 2018b) and DNA–protein complexes (Boudet et al., 2019; Lacabanne et al., 2020; Malär et al., 2021b; Wiegand et al., 2020b, 2019, 2017b, 2016; Zehnder et al., 2021).

The main challenge in NMR of RNA both in solution and in solid-state is the overlap of the signals due to the limited chemical diversity of the nucleotides. This is especially true in canonical, and thus homogeneous, tertiary structure elements, such as A-form helices. This challenge can be addressed using nucleotide-type selective and/or segmental isotope labeling (Duss et al., 2010; Nelissen et al., 2008; Tzakos et al., 2007), as well as site-specific labeling (Lu et al., 2010; Marchanka et al., 2018a), which reduce spectral crowding. As mentioned above, selective labeling becomes crucial in ssNMR, where the signal overlaps are significant also for RNAs of medium size in the absence of significant structural diversity (i.e. in helical regions).

Fig. 3. CSP mapping to determine nucleic acid–protein binding interfaces in ssNMR. (A) Excerpts of overlaid 2D CP NCACX correlation spectra of free (black) and RNA-bound (green) Pf L7Ae protein, with peaks showing noticeable CSPs upon RNA binding. (B) Same as in (C) for 2D $^{13}$C–$^{13}$C DARR correlation spectra. Figures (A–B) are reproduced from (Ahmed et al., 2020) ©2020 with permission from John Wiley and Sons. (C–D) Temperature-dependent proton chemical-shift perturbations. Residue-specific temperature coefficients and corresponding excerpts of the hNH spectra of DnaB helicase complexed with ADP·AlF₄⁻ and ssDNA. Residues in (C) show nearly no dependence of their chemical shifts on temperature and are thus likely involved in hydrogen bonds. The chemical shifts of residues in (D) have a larger dependence on temperature and are thus not involved in hydrogen bonds. Figures (C–D) are reproduced from (Malär et al., 2021b) ©2021 with permission from Springer Nature.

The second challenge in NMR of RNA is the unequal proton distribution. Nucleic acids have a high proton density in the ribose ring but only few protons in the nucleobases and no protons at the backbone phosphate. This leads to a limited number of $^1$H–$^1$H distance restraints available to determine the conformation at both the backbone and the Watson-Crick and Hoogsteen sides of the nucleobases. Fortunately, in ssNMR, the number of distance restraints that can be collected does not directly correlate with the number of protons, as distance restraints can be measured via both heteronuclear and homonuclear correlations mediated by dipolar couplings. Moreover, the distance range of restraints measured in ssNMR experiments such as rotational echo double resonance (REDOR) (Gullion and Schaefer, 1989a, 1989b) or proton-driven spin diffusion (PDS) (Szeverenyi et al., 1982) can exceed the range of those obtained in solution NMR by NOESY experiments (Huang et al., 2011, 2010; Marchanka and Carlomagno, 2019; Olsen et al., 2005, 2003; Studelska et al., 1996), providing a useful tool for the refinement.
of global conformations. For example, REDOR experiments have provided long-range distance restraints up to 16 Å in proteins (Studeiski et al., 1996) and 13 Å in nucleic acids (Olsen et al., 2003). ssNMR can adopt an important role in structural biology of RNP complexes. However, the disadvantages caused by signals overlap have long discouraged the application of ssNMR to RNA-containing particles. In the past decade our lab has developed a suite of ssNMR experiments that achieve assignment of RNA $^{13}$C,$^{15}$N (Marchanka et al., 2015, 2013; Marchanka and Carlomagno, 2019) and $^1$H resonances (Agion et al., 2021; Agion and Marchanka, 2021; Marchanka et al., 2018b) as well as de novo RNA structure determination using distance restraints obtained solely from ssNMR experiments (Marchanka et al., 2015). These experiments, together with those developed in several other laboratories for the structure determination of proteins using $^{13}$C,$^{15}$N detection (Castellani et al., 2003; Zhao, 2012) at low MAS frequencies and $^1$H detection (Andreas et al., 2016; Schubeis et al., 2018) at fast MAS frequencies (Penzel et al., 2019; Schledorn et al., 2020), allow the structure determination of the individual RNA and protein components of RNP complexes by ssNMR. Once the structures of the individual components are established, the identification of the intermolecular interfaces as well as the measurement of intermolecular distances yield the structure of the complex.

In this review we present recent advances of ssNMR spectroscopy to determine intermolecular contacts in RNP complexes and discuss advantages and challenges of the individual strategies. Due to the similar nature of the interfaces, we also review ssNMR studies of intermolecular interactions in DNA–protein complexes. We present conventional ssNMR methods that rely on the detection of low $^1$H nuclei, such as $^{13}$C, $^{15}$N and $^{31}$P, as well as novel $^1$H-detected ssNMR experiments under fast magic-angle-spinning (MAS) rates. Finally, we discuss the first example of structure determination of an RNA–protein complex guided solely by ssNMR-derived restraints.

Sample preparation

NMR of nucleic acids is more challenging than that of proteins due to the poor chemical shift dispersion of their NMR signals. Thus, even for relatively small nucleic acids, advanced isotope labeling strategies may be required to obtain site specific structural information. Nucleic acids for ssNMR studies can be prepared by either chemical (Beaucage and Reece, 2009; Roy and Caruthers, 2013) or enzymatic synthesis; however, the methods available to produce isotope-labeled RNA are considerably more advanced than those available for DNA. An extensive description of isotope labeling strategies for RNA (Marchanka et al., 2018a) and ssDNA/dsDNA (Nelissen et al., 2016) by either chemical or enzymatic synthesis can be found in the literature.

As opposed to X-ray crystallography, ssNMR does not require crystals of any particular size and is therefore applicable to particles with substantial flexible regions, which are difficult to force in well-ordered, large crystals. Common sample preparation techniques in ssNMR include micro (nano)-crystallization (Bertini et al., 2010a; Franks et al., 2005; Huang et al., 2012; Marchanka et al., 2013; McDermott et al., 2012; Wang et al., 1994), freezing in the presence of a cryoprotectant (Siemer et al., 2013), or sedimentation of soluble macromolecules into the ssNMR rotor using ultracentrifugation (Barbet-Massin et al., 2015; Bertini et al., 2011; Gardiennet et al., 2012; Lacabanne et al., 2020; Wiegand et al., 2016, 2020a). Micro-crystallization, ethanol precipitation and sedimentation could all yield sufficiently narrow linewidths to allow for site-specific assignments in individual samples (Agion and Marchanka, 2021); however, ethanol precipitation is incompatible with the protein component of RNP complexes, while sedimentation has been successfully applied to RNP complexes, such as the prokaryotic ribosome (Barbet-Massin et al., 2015), but no data is available with respect to RNA linewidths in these samples. In contrast, both lyophilization and flash-freezing have been demonstrated to lead to inhomogeneous line broadening (Huang et al., 2011; Olsen et al., 2005; Siemer et al., 2012), impairing site-specific assignments. Nevertheless, structural information can be obtained also for these samples, in
ssNMR of nucleic acid–protein complexes prepared using either sedimentation or micro-crystallization have been reported to be stable over a long time. A DNA–protein–ATP complex prepared by sedimentation and stored at −20 °C for 3 years has been shown to yield virtually the same 13C–13C DARR spectrum as the freshly prepared sample (Lacabanne et al., 2020; Wiegand et al., 2020a). A ssNMR RNP complex sample prepared by micro-crystallization in our laboratory (Aguion et al., 2021) shows identical 1H–13C and 1H–15N fingerprint spectra after storage at +4 °C for two years, with only minimal loss of signal intensities (Fig. 1).

Finally, ssNMR 1H-detected experiments at MAS rates above 100 kHz require low sample quantities (300–800 μg for rotors of 0.7–0.8-mm size) (Aguion et al., 2021; Lacabanne et al., 2020; Marchanka et al., 2018b), thus limiting the cost and time-demand of sample preparation.

Characterization of nucleic acid–protein interfaces by MAS ssNMR

The approaches used so far to characterize intermolecular contacts in nucleic acid–protein complexes can be divided in three classes, whereby many of the published studies use a combination of these approaches.

(i) Similar to solution-state NMR, the involvement of a molecular surface in interactions with a binding partner can be detected by either chemical shift perturbations (CSPs) or intensity changes of the ssNMR peaks of the surface atoms, when comparing the free and the bound-state of the molecule (Ahmed et al., 2020; Boudet et al., 2019; Lacabanne et al., 2020; Malar et al., 2021b; Wiegand et al., 2020b, 2019, 2016; Williamson, 2013). For example, the formation of an hydrogen bond at a nucleic acid–protein interface causes a downfield shift of the involved 1H atom (Wagner et al., 1983). Changes in peak intensities can report on changes in the dynamics of one of the binding partners upon complex formation.
Fig. 6. Homonuclear 2D $^{13}$C-$^{13}$C DARR spectrum of uniformly $^{13}$C,$^{15}$N-labeled PF1 bacteriophage virion. The spectrum shows cross peaks between the ssDNA ribose and base atoms and a specific tyrosine residue (Y40) of the viral coat protein, as highlighted by orange boxes. The figure is reproduced from (Sergeyev et al., 2011) ©2011 with permission from American Chemical Society.

(Hacabanne et al., 2020; Wiegand et al., 2019). In another study, Asami et al. measured a $^{15}$N–$^1$H correlation spectrum of a protein in complex with either $^1$H or $^2$H-RNA (Asami et al., 2013) and quantified the difference in protein peaks’ intensities that distinguished the RNA–protein interface, owing to the line-broadening caused by the RNA $^1$H in the $^1$H–RNA–protein complex. These effects are specific to the surface of the protein in contact with the RNA, while CSPs can also occur in regions other than the intermolecular surface, due to allosteric effects. In any case, both CSPs and changes in peak intensities can be used as ambiguous restraints in docking protocols.

(ii) Intermolecular distances can be measured directly through intermolecular dipolar correlation experiments. In solution-state NMR, intermolecular distances are measured through $^{13}$C,$^{15}$N-edited, $^{13}$C,$^{15}$N-filtered $^1$H–$^1$H NOESY experiments (Breeze, 2000; Zwalien et al., 1997). In ssNMR, a plethora of methods yield dipolar correlations, such as cross-polarization (CP) (Hartmann and Hahn, 1962), rotational echo double resonance (REDOR) (Gullion and Schaefer, 1989a, 1989b) and the closely related transferred echo double resonance (TEDOR) (Hing et al., 1992), as well as dipolar-assisted rotational resonance (DARR) (Takegoshi et al., 2001, 1999), proton-driven spin diffusion (PDSD) (Szeverenyi et al., 1982), proton spin diffusion (PSD) (Lange et al., 2002; Wilhelm et al., 1998), radio frequency-driven dipolar recoupling (RFDR) (Bennett et al., 1992; Sodickson et al., 1993), or combined R2$^*$-driven (CORD) (Hou et al., 2013) experiments. One of the most popular approaches for the measurement of nucleic acid–protein distances utilizes TEDOR-based $^{31}$P–$^{13}$C and and $^{31}$P–$^{15}$N correlations (Bechinger et al., 2011; Huang et al., 2011, 2010; Jehle et al., 2010; Olsen et al., 2005; Yu and Schaefer, 2008). $^{31}$P has a high gyromagnetic ratio (Table 1) and is present in nucleic acids exclusively and with 100 % isotopic abundance; thus, in the presence of a $^{13}$C,$^{15}$N-labeled protein, these correlation experiments are sensitive and report exclusively on intermolecular contacts between the protein and the nucleic acid.

(iii) The third approach utilizes paramagnetic relaxation enhancement (PRE) effects, which rely on the interaction between nuclei and unpaired electrons. Because the electron spin gyromagnetic ratio is approximately 600-times larger than that of the $^1$H nucleus, these effects are large and can be used to measure longer distances than internuclear dipolar correlations. In solution, PRE measurements have been often applied to measure long-range intra- and intermolecular distances (up to 30 Å) in nucleic acid–protein complexes (Amran et al., 2014; Grazialeti et al., 2020; Hennig et al., 2015; Lapinaite et al., 2013; Leeper et al., 2010; Mackereith et al., 2011; Martin-Tumasz et al., 2010). For a detailed description of paramagnetic NMR in solution and solid-state, readers are referred to a comprehensive review by Pell and coworkers (Pell et al., 2019). In ssNMR, PRE measurements were first employed to identify residues of metalloproteins in proximity to the paramagnetic metal (Balayssac et al., 2007b, 2007a; Pintacuda et al., 2007); recently, PRE experiments have also been used to study nucleic acid–protein complexes, (Ahmed et al., 2020; Wiegand et al., 2017a, 2017b; Zehnder et al., 2021). Notably, Ahmed et al. report the first structure of an RNP complex obtained solely from ssNMR data.

Fig. 2 gives an overview of the methods developed and utilized to date to probe nucleic acid–protein interfaces, which will be discussed in detail in the next chapters. A representative set of ssNMR studies of nucleic acid–protein complexes is given in Table 2 in chronological order.

Chemical shift perturbations and intensity changes

Both CSPs and peak intensities changes can be used to reveal intermolecular surfaces. CSP mapping can be achieved with all types of NMR spectra. In $^{13}$C,$^{15}$N-detected ssNMR, chemical-shift differences measured in 2D $^{13}$C–$^{13}$C DARR spectra (Ahmed et al., 2020; Boudet et al., 2019; Wiegand et al., 2016, 2019) and 2D CP $^{13}$C–$^{15}$N spectra (Ahmed et al., 2020; Boudet et al., 2019; Wiegand et al., 2019) of $^{13}$C,$^{15}$N-labeled proteins in the free and complexed forms were used to identify surfaces involved in nucleic-acid binding (Fig. 3A–B). Wiegand et al. used 1D $^1$H–$^{31}$P CP spectra to detect binding of a protein to (dT)$_{20}$ ssDNA, whose $^{31}$P peaks shifted from ~1 ppm in the free form to 0–1 ppm in the protein-bound form (Wiegand et al., 2016, 2019). However, nucleic acid CSPs have never been used to reveal the nucleic acid residues involved in protein recognition. This task is indeed challenging, especially for RNAs, as these do not always adopt a single conformation in the free form and thus the observed CSPs may report on both folding and binding.

Recently, $^1$H-detected ssNMR methods, such as CP hNH and hCH experiments (Zhou et al., 2007a, 2007b; Zhou and Rienstra, 2008), have been utilized to map CSPs of a $^{13}$C,$^{15}$N-labeled protein in a DNA–protein complex upon nucleotide binding to yield a DNA–protein–nucleotide complex (Lacabanne et al., 2020; Wiegand et al., 2020a). $^1$H-detection in MAS ssNMR is expected to greatly enhance the resolution and sensitivity of CSP mapping.

Another chemical-shift-based approach was developed by Malär et al. to detect protein side chains involved in hydrogen bonds with the nucleic acid and the nucleotide in a ssDNA–protein–nucleotide complex (Malär et al., 2021b). The method measures the temperature dependence of $^1$H chemical shifts and correlates a weak dependence with the involvement of the $^1$H in a hydrogen bond. This correlation was often
demonstrated in solution (Baxter and Williamson, 1997; Cierpicki et al., 2002; Cierpicki and Otlewski, 2001) and more recently also in ssNMR (Malár et al., 2021a). Mapping the dependence of chemical shifts on temperature in a CP hNH experiment, Malár et al. were able to confirm that the chemical shifts of 1H protein atoms previously identified as forming hydrogen bonds with the nucleotide and the DNA vary very little with temperature and certainly much less than those not involved in hydrogen bonds (Fig. 3C–D). These experiments, when applied to proteins in both the free and nucleic acid-bound states, can reveal which protein amino acids are involved in nucleic acid recognition.

The first study that employed quantification of ssNMR peak intensities to determine intermolecular surfaces in an RNP complex was communicated by Asami et al. (Asami et al., 2013), who also pioneered 1H-detection in MAS ssNMR of uniformly 13C,15N labeled RNA at moderate MAS frequencies of 20 kHz. The 1H,15N cross peaks of protein and RNA detected in a hNH experiment (Zhou et al., 2007a, 2007b; Zhou...
and Riestra, 2008) do not significantly overlap, with the exception of those belonging to RNA amino groups and protein arginine sidechains (Fig. 4A). This allowed monitoring both components of the complex simultaneously. To obtain narrow linewidths in the hNH experiment at low MAS speed, both RNA and protein were exchanged to hydrogen-2, while the labile hydrogen-2 was back-exchanged to hydrogen-1 to a 10% extent. To identify the surface of the protein that binds the RNA, Asami et al. measured protein 1H,15N correlations of two samples, containing 2H,15N-labeled protein and either 2H,15N-labeled RNA or 15N,13C-labeled RNA. By comparing the intensities of the 1H,15N protein cross peaks in the two samples they could identify which protein amino acids form the RNA–protein interface, as the 2H,15N peaks of these amino acids had reduced intensities in the sample containing deuterated RNA, due to the line broadening of the protein hydrogen-2 caused by the dipolar interaction with the nearby RNA hydrogen-1 atoms. Spectra simulations using the program SIMPSON (Bak et al., 2000) demonstrated that a measurable effect occurred when the distance between the protein hydrogen-2 and any RNA hydrogen was less than 6–8 Å, while the intensities of the protein amide peaks further than 8 Å from the RNA were identical in the two spectra (Fig. 4B). In principle, the ratio of signal intensities in the two samples can be quantified and converted in a distance value; however, as the experiment does not reveal which RNA hydrogen causes the line broadening of a given protein hydrogen-2, these distances cannot be used as unambiguous restraints in structure calculations.

Other studies detected side-chain-specific nucleic acid–protein interactions by monitoring the appearance or disappearance of arginine and lysine side chain peaks in 1H,13C CP correlation spectra or CP INH spectra upon binding of either nucleic acids or nucleotides (Lacabanne et al., 2020; Wiegand et al., 2019). The appearance of the peaks of these side chains was due to the loss of local dynamics connected with binding. Wiegand et al. also detected the appearance of a 31P,31P correlation peak in the 31P,31P DARR spectrum of the (dT)20 ssDNA upon binding to the DnaB–ADP:AlF4 binary complex, as a result of the rigidification of two of the 20 ssDNA nucleotides when they formed the ssDNA–protein–nucleotide complex (Wiegand et al., 2020b).

Cross-interface dipolar correlations

MAS ssNMR techniques to measure internuclear distances were first applied to an RNA–peptide complex by Olsen et al., who, however, did not measure intermolecular distances but detected peptide binding through the change in the value of a 19F–31P distance in the RNA (Olsen et al., 2005). The 19F atom was introduced as a 2-fluorine at an individual position in nucleic acids, either in the ribose-5′-fluorine at an individual position in nucleic acids, either in the ribose 2′-F, or in the base, as 5-fluorouridine (5FU) and 5-fluorocytidine (5FC), has become popular because of the large chemical shift dispersion (Marchanka et al., 2006) and the high sensitivity of its chemical shift to the structural and chemical environment (Hennig et al., 2007; Marchanka et al., 2018a; Scott et al., 2004). Moreover, substitution of 1H by 19F in nucleic acids can be quantified and converted in a distance value; however, as the experiment does not reveal which RNA hydrogen causes the line broadening of a given protein hydrogen-2, these distances cannot be used as unambiguous restraints in structure calculations.

Other studies detected side-chain-specific nucleic acid–protein interactions by monitoring the appearance or disappearance of arginine and lysine side chain peaks in 1H,13C CP correlation spectra or CP INH spectra upon binding of either nucleic acids or nucleotides (Lacabanne et al., 2020; Wiegand et al., 2019). The appearance of the peaks of these side chains was due to the loss of local dynamics connected with binding. Wiegand et al. also detected the appearance of a 31P,31P correlation peak in the 31P,31P DARR spectrum of the (dT)20 ssDNA upon binding to the DnaB–ADP:AlF4 binary complex, as a result of the rigidification of two of the 20 ssDNA nucleotides when they formed the ssDNA–protein–nucleotide complex (Wiegand et al., 2020b).

Cross-interface dipolar correlations

MAS ssNMR techniques to measure internuclear distances were first applied to an RNA–peptide complex by Olsen et al., who, however, did not measure intermolecular distances but detected peptide binding through the change in the value of a 19F–31P distance in the RNA (Olsen et al., 2005). The 19F atom was introduced as a 2-fluorine at an individual position in nucleic acids, either in the ribose 2′-F, or in the base, as 5-fluorouridine (5FU) and 5-fluorocytidine (5FC), has become popular because of the large chemical shift dispersion (Marchanka et al., 2006) and the high sensitivity of its chemical shift to the structural and chemical environment (Hennig et al., 2007; Marchanka et al., 2018a; Scott et al., 2004). Moreover, substitution of 1H by 19F in nucleic acids can be quantified and converted in a distance value; however, as the experiment does not reveal which RNA hydrogen causes the line broadening of a given protein hydrogen-2, these distances cannot be used as unambiguous restraints in structure calculations.

Other studies detected side-chain-specific nucleic acid–protein interactions by monitoring the appearance or disappearance of arginine and lysine side chain peaks in 1H,13C CP correlation spectra or CP INH spectra upon binding of either nucleic acids or nucleotides (Lacabanne et al., 2020; Wiegand et al., 2019). The appearance of the peaks of these side chains was due to the loss of local dynamics connected with binding. Wiegand et al. also detected the appearance of a 31P,31P correlation peak in the 31P,31P DARR spectrum of the (dT)20 ssDNA upon binding to the DnaB–ADP:AlF4 binary complex, as a result of the rigidification of two of the 20 ssDNA nucleotides when they formed the ssDNA–protein–nucleotide complex (Wiegand et al., 2020b).
acids has a negligible effect on both structure (Hennig et al., 2007; Merritt et al., 1999) and intermolecular interactions (Olsen et al., 2005).

The substitution of an individual phosphate group with a thiophosphate has also been widely used in NMR spectroscopy of nucleic acids to unambiguously identify individual 31P atoms thanks to the downfield shift by 50–60 ppm induced by the sulfur atom on the 31P chemical shift (Huang et al., 2011, 2010; Merritt et al., 1999; Olsen et al., 2005, 2003). Also in this case, the pS substitution results in little-to-no perturbation to the global structure, as demonstrated by several high-resolution NMR studies for DNA:RNA hybrids (Bachelin et al., 1998; Gonzalez et al., 1995; Merritt et al., 1999) and crystal structures of wild-type and pS-substituted DNA duplexes (Cruse et al., 1986). However, alterations of RNA conformation caused by pS substitutions have been reported as well (Smith and Nikonowicz, 2000). In addition, the different charge distribution and polarizability of thiophosphate in comparison to phosphate alters the strength of both ionic interactions and hydrogen bonds and thus, potentially, nucleic acid–protein affinities (Milligan and Uhlenbeck, 1989; Schnitzer and Von Ahsen, 1997). Consequently, it is safer to use methods that do not require chemical modification of the nucleic acid to measure internuclear distances in the context of nucleic acid–protein complexes. With the recent developments in the ssNMR experimental design and sample preparation techniques, it is nowadays possible to obtain site-specific assignments of 31P resonances in the backbone of 13C,15N labeled RNA by, for example, 13C,31P-TEDOR and 13C,31P-PEDOR,13C,15N-PDSD experiments (Marchanka et al., 2015; Marchanka and Carlomagno, 2019). In addition, 1H-detected hPH experiments serve as useful complement to confirm 31P resonance assignments (Malal et al., 2021b).

13C,31P, 15N,31P, 13C,15P, and 15N,13P REDOR experiments on an RNA-peptide complex (Huang et al., 2011, 2010), as well as 31P,15N TEDOR experiments on an RNA–protein complex (Jehle et al., 2010) were used to measure the first intermolecular RNA–peptide and RNA–protein distances through the dipolar coupling between the RNA 19F and 31P nuclei and the protein/peptide 13C and 15N nuclei (Fig. 5D–E). These experiments were performed on complexes containing 13C,15N-labeled protein and unlabeled or selectively fluorinated RNA and yielded exclusively intermolecular cross peaks. Huang et al. utilized a selectively labeled peptide, where a single arginine was 13C,15N enriched, and 5FU/ pS-modified RNA (Huang et al., 2011, 2010). Besides the problems discussed above with regards to the impact of chemical modifications on RNA structure and interactions, the use of selectively labeled molecules requires these molecules to be available by chemical synthesis, to allow the easy incorporation of labeled or modified building blocks at any position in the primary sequence. This requirement limits the applicability of the strategy. In a more general approach, Jehle et al. measured 31P,15N RNA backbone-protein backbone distances between a uniformly 13C,15N labeled protein and an unlabeled RNA (Jehle et al., 2010). As the assignment of both 31P RNA (Marchanka et al., 2015; Marchanka and Carlomagno, 2019) and 15N protein (Andreas et al., 2016; Castellani et al., 2003; Schubeis et al., 2018; Zhao, 2012) peaks can be obtained by ssNMR experiments, the method is broadly applicable.

Other studies used 15N,31P REDOR experiments to probe intermolecular contacts between either bacteriophage T4 DNA or salmon sperm DNA 31P atoms and 15N atoms of lysine side chains in either a DNA–protein (Yu and Schaefer, 2008) or a DNA–peptide complex (Behinger et al., 2011). In both cases, the interacting atoms were not assigned, and the experiments were used solely to prove binding.

Following a different approach, Sergeyev et al. (Sergeyev et al., 2011) were able to detect cross-peaks between the uniformly 13C,15N labeled ssDNA of the PF1 bacteriophage virion and a specific tyrosine residue of the viral coat protein using dynamic nuclear polarization (DNP) enhanced homonuclear 13C,15C DARR correlation experiments (Abragam and Goldman, 1978; Abkay and Oschkinat, 2016; Su et al., 2015; Takegoshi et al., 2001, 1999). The aromatic stacking interactions between the DNA bases and the tyrosine had been previously predicted but had remained unproven, until they were detected by ssNMR (Fig. 6).

In another study (Morag et al., 2014), fd bacteriophages were prepared on a 13C,15N enriched medium supplemented with non-isotopically labeled aromatic amino acid precursors to obtain particles with uniformly labeled DNA and selectively-unlabeled proteins. Assignment of atom-types of the fd bacteriophage circular ssDNA was obtained using a 13C,15C CORD correlation experiment (Hou et al., 2013). However, due to the large size of the DNA, site-specific resonance assignment was not attempted. The contacts between the ssDNA and the virion capsid proteins were measured in 13C,15C CORD and 15N,13C DARR experiments. The intermolecular correlations could be distinguished from the intramolecular ones due to the partially different chemical shift ranges of 13C atoms in DNA and proteins. DNA phosphate–capsid protein interactions were detected utilizing homonuclear 1H–1H PSD mixing (Lange et al., 2002; Wilhelm et al., 1998) coupled with P–H and H–C CP transfers (Hartmann and Hahn, 1962) in a PHHC experiment. The PHHC experiment yielded intramolecular correlations between the DNA phosphate backbone and 13C nuclei of the DNA riboses and bases as well as intermolecular correlations between the DNA phosphate backbone and side chains of amino acids of the coat protein, in particular lysines (Fig. 7A). This study is remarkable, due to the large size of the particles under investigation. A similar approach could be used to reveal RNA–protein interfaces in RNP complexes, possibly in combination with site-specific assignments. However, neither 12C,13C DARR nor PHHC experiments have been applied to measure RNA–protein distances so far.

The CHHP experiment, which is conceptually the same as the PHHC experiment, but with detection of 13C and 19F chemical shifts in τ1 and τ2, respectively, and the NHHP experiment, where polarization is transferred between the protein 15N nuclei and the DNA 31P nuclei through 1H–1H spin diffusion, were utilized in other studies (Boudet et al., 2019; Wiegard et al., 2019, 2020) to observe intermolecular 13C,31P and 15N,31P contacts between a 13C,15N-labeled protein and DNA (Fig. 7B, C).

All these 31P,13C,15N heteronuclear correlation experiments suffer
from low signal-to-noise ratio (SNR), in particular for large complexes. The problem can be addressed using DNP, which can significantly enhance the SNR of NMR spectra exploiting the transfer of the large electron polarization to nuclear spins. However, biomolecular MAS-DNP studies are typically performed at cryogenic temperatures, which lead to significant broadening of the NMR lines (Bauer et al., 2017; Siemer et al., 2012; Siemer and McDermott, 2008). Nevertheless, in DNP-enhanced $^{31}P$–$^{13}C$ CP and PHHC transfer experiments, Wiegand et al. measured intermolecular contacts between a $^{13}C$, $^{15}N$ labeled protein and either ssDNA or ADP in a ssDNA–protein–ADP complex (Wiegand et al., 2017b). In both cases, a $^{13}C$–$^{13}C$ DARR mixing step was added after the intermolecular transfer step to reduce spectral overlap, resulting in 2D ($^{13}C$,$^{15}N$) (PC)–$C$ and (ii) 2D (PHHC)–$C$ correlation spectra. Due to the line broadening induced by cryogenic temperatures, the experiments yielded information only on the protein amino-acid type in contact with either ADP or ssDNA, but allowed neither site-specific protein assignments nor discrimination between the $^{31}P$ chemical shifts of the ssDNA and ADP. Of note, the $^{1}H$–$^{1}H$ spin-diffusion-driven PHHC–$C$ correlation is less selective than the direct CP-driven PC–$C$ correlation but can yield longer $^{31}P$–$^{13}C$ distances of the order of 7–10 Å.

Finally, Malár et al. have recently probed spatial proximities between protein $^{13}C$ nuclei and nucleic acid/nucleotide $^{31}P$ nuclei in a ssDNA–protein–ADP complex utilizing a $^{1}H$-detected $^{31}P$–$^{13}C$ correlation experiment under high MAS speed of 105 kHz (Fig. 8) (Malár et al., 2021b). The CP hPH experiment can be performed with unlabeled material and appears to be much more sensitive for probing spatial proximity of hydrogen atoms to phosphate groups than the previously described $^{1}H$–$^{1}H$ PSD-based CHHP and NHHP experiments.

**Paramagnetic relaxation enhancement**

The first implementation of PREs in ssNMR of a protein–nucleotide complex took advantage of the fact that ATP-binding is often accompanied by binding of Mg$^{2+}$ ions, which can be substituted by paramagnetic ions such as Mn$^{2+}$ (Bonneau and Legault, 2014; Tamaki et al., 2016) and Co$^{2+}$ (Balayssac et al., 2008; Bertini et al., 2010b) without affecting protein function (Otting, 2010). The presence of a paramagnetic center can cause an increase of the relaxation rates (paramagnetic relaxation enhancement, PRE) and/or a change in the chemical shifts (pseudo-contact shifts, PCS) of close-by nuclei (Jaroniec, 2012). These effects depend quantitatively on the distance between the nucleus and the paramagnetic center. Through the analysis of signal intensities in 2D $^{1}H$–$^{1}H$ DARR spectra of the diamagnetic (in the presence of Mn$^{2+}$) and paramagnetic (in the presence of Mn$^{2+}$) protein–nucleotide complex, Wiegand et al. identified protein residues close to the nucleotide binding site (Wiegand et al., 2017a). Recently, Zehnder et al. extended this approach to a ssDNA–protein–nucleotide complex, in which Mn$^{2+}$ and Co$^{2+}$ ions were used as paramagnetic centers (Zehnder et al., 2021). 2D $^{13}C$–$^{13}C$ DARR, as well as 2D NCA and 3D NCAB spectra (Baldus et al., 1998) were recorded for diamagnetic and paramagnetic ssDNA–protein–nucleotide complexes (Fig. 9A) to localize the metal ion in the complex. Moreover, PCSs of the (d7)20 ssDNA observed in 1D $^{1}H$–$^{31}P$ CP spectra allowed to localize the DNA phosphate groups in the complex (Fig. 9B).

If the molecule of interest does not have a well-defined, individual metal binding site, paramagnetic centers can be introduced in nucleic acid–protein complexes by coupling nitroxide-based tags, one-by-one, to individual cysteine residues engineered at specific protein sites (Nadad et al., 2007). The PRE effects caused by the paramagnetic tag on the RNA can be used to measure RNA–protein distances in the RNP complex (Ahmed et al., 2020). Ahmed et al. measured PRE effects induced in an RNP complex by protein-coupled paramagnetic tags on $^{13}C$, $^{15}N$-labeled RNA in 2D $^{13}C$–$^{15}N$-labeled proteins (Hobday et al., 2002, 1999) (Fig. 9C). Unlike CSPs, which are sensitive also to allosteric effects, PRE-effects are specific reporters of intermolecular contacts. To avoid PRE effects generated by interparticle contacts in the precipitate, the paramagnetic-labeled RNP complex was diluted in 1:3 molar ratio with the diamagnetic, unlabeled RNP complex.

To obtain enough RNA-protein distance restraints by this methodology, the paramagnetic tag should be placed close to the RNA binding site but in such ways as not to affect binding (Fig. 9D). CSP data, discussed previously, can guide in the recognition of the RNA binding surface and in the design of the cysteine mutants. The integrity of the complex, after introduction of the nitroxide tag, should be verified by careful inspection of the NMR spectra and of other biophysical properties of the complex. As far as nucleic acids are concerned, paramagnetic tags can be covalently attached to either the phosphate backbone or the ribose or modified bases. In alternative, modified paramagnetic nucleotides can be introduced into the RNA by solid-phase synthesis (Miao et al., 2021). Although paramagnetically tagged RNA has not been used in ssNMR so far, these techniques can be easily applied to increase the number of paramagnetic sites that can be probed in nucleic acid–protein complexes.

**Structure determination of an RNP complex by ssNMR only**

Ahmed et al. succeeded in determining the structure of the complex formed by the Pf protein L7Ae and the 26mer Box C/D RNA using exclusively ssNMR-derived restraints (Ahmed et al., 2020). They achieved this goal through a docking protocol, which started from the individual structures of protein-bound 26mer RNA determined by ssNMR (Marchanka et al., 2015) and RNA-bound L7Ae protein (Xue et al., 2010). The RNA-bound protein structure was determined by X-ray crystallography, but in principle it would have been accessible by ssNMR through well-established methodology (Schubels et al., 2018; Zhao, 2012). The docking protocol was implemented in the program Haddock (Domínguez et al., 2003) and used a combination of ssNMR-derived CSP and PRE data to guide complex assembly.

PRE-based restraints were derived from the quantification of PRE effects observed on nucleotide-type specific $^{13}C$, $^{15}N$ labeled (A$^{\text{lab}}$ and U$^{\text{lab}}$) RNA in 2D $^{13}C$–$^{15}N$ SPC-5 spectra and induced by 3 different paramagnetic tags coupled to the unlabeled protein. PRE effects were quantified as peak volume ratios in spectra measured for the same sample in either the paramagnetic (with the nitroxide radical coupled to a cysteine residue) or the diamagnetic state (after addition of ascorbic acid to reduce the radical). These ratios ($\frac{V_{\text{para}}}{V_{\text{dia}}}$) were converted into distance restraints in a semi-quantitative manner, whereby the intramolecular $\frac{V_{\text{para}}}{V_{\text{dia}}}$ ratios of the protein peaks, whose electron-nucleus distances were known from the protein structure, were used to calibrate a linear regression between the $\frac{V_{\text{para}}}{V_{\text{dia}}}$ and the electron-nucleus distance. Although not rigorously correct, this linear regression, with the appropriate tolerance bounds, was good enough to deliver broad distance ranges that defined the relative position of the protein and the RNA to a good level of precision.

CSP-derived ambiguous restraints were measured for the protein in 2D $^{13}C$–$^{13}C$ DARR and 2D CP $^{13}C$–$^{15}N$ spectra of uniformly $^{13}C$, $^{15}N$ labeled protein in the free and RNA-bound states.

In total 72 restraints were used for the docking and lead a 26mer Box C/D RNA–L7Ae structure that is very similar to that of an orthologous Box C/D RNA–L7Ae complex determined by crystallography, thus verifying the accuracy of the ssNMR-derived structure (Fig. 10). This exemplary study demonstrates the ability of ssNMR to yield nucleic acid–protein complex structures at atomic resolution.

**Conclusions and outlook**

The ssNMR methods discussed here build a suite of complementary experiments that identify interaction interfaces and yield intermolecular distance restraints in nucleic acid–protein complexes. They provide a powerful tool to determine the structural basis of intermolecular recognition for those nucleic acids–protein complexes that are not
amenable to X-ray crystallography, cryo-EM or solution-state NMR. The structure of the individual components of the complex can be obtained by ssNMR as well as with previously published and often reviewed methodology (Marchanka and Carломagno, 2019; Schubeis et al., 2018; Zhao, 2012).

In 2021 the program AlphaFold (Jumper et al., 2021) revolutionized structural biology of folded protein domains and their complexes, by demonstrating an unprecedented accuracy in the prediction of protein folding and interactions based on database knowledge. Nucleic acid structures, in particular RNA, are difficult to predict, as they largely depend on the environment and on the binding partners. Thus, it is unclear whether AlphaFold will ever be expanded to this class of polymers. In view of this, NMR spectroscopy, with its power to illuminate intermolecular interactions involving flexible molecules, gains unique relevance.

To date, many studies of nucleic acid–protein complexes have focused on obtaining site specific information for the proteins in the complex. The nucleic acid component has been often neglected due to lack of spectral resolution and limited access to advanced isotope labeling techniques. Consequently, we expect MAS ssNMR of nucleic acid structures, in particular RNA, are difficult to predict, as they largely depend on the environment and on the binding partners. Thus, it is unclear whether AlphaFold will ever be expanded to this class of polymers. In view of this, NMR spectroscopy, with its power to illuminate intermolecular interactions involving flexible molecules, gains unique relevance.

Declarations of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has received support from the DFG (Deutsche Forschungsgemeinschaft) grant CA294/21-1 to T.C.

References

Abraham, A., Marchanka, T., 2019. Eukaryotic ribosome at 3.0 Å resolution. Nature. 463, 689–694. https://doi.org/10.1038/463689a
Bacchus, F., Schmitz, C., Robijn, F., Grzesiek, S., Wüthrich, K., 2011. High-resolution solid-state NMR studies of metalloproteins. J. Am. Chem. Soc. 133, 1040–1040. https://doi.org/10.1021/ja106426p
Bauer, T., Dotta, C., Balacescu, L., Gad, J., Huneker, A., Böckmann, A., Meier, B.H., 2017. Line Broadening in Low Temperature Solid State NMR Spectra of Fibers. J. Biomol. NMR. 67, 51–61. https://doi.org/10.1007/s10856-016-0083-4
Baxter, N.J., Williamson, M.P., 1997. Temperature dependence of 1H chemical shifts in protein-DNA interactions. J. R. Soc. Interface 8, 1065–1073. https://doi.org/10.1098/rsif.2010.0543
Bechinger, B., Vidovic, V., Bertani, P., Richer, A., 2011. A new family of peptide-nucleic acid nanostructures with potent transfection activities. J. Pept. Sci. 17, 88–93. https://doi.org/10.1002/jps.21318
Bennett, A.E., Ok, J.H., Griffin, R.G., Vega, S., 1992. Chemical shift correlation spectroscopy in rotating-frame frequency-driven dipolar recoupling and longitudinal exchange. J. Chem. Phys. 96, 8624–8627. https://doi.org/10.1063/1.462267
Benn-Shem, A., Garreau de Loubresse, N., Melikov, S., Jenner, L., Yusupova/Gulnara, Yusupov, M., 2011. The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. Science. 334, 1524–1529.
Berti, I., Bhamuik, A., De Paepe, G., Griffin, R.G., Lelli, M., Lewandowski, J.R., Luchinat, C., 2010a. High-resolution solid-state NMR structure of a 17.6 kDa protein. J. Am. Chem. Soc. 132, 1032–1040. https://doi.org/10.1021/ja906426p
Berti, I., Emley, L., Lelli, M., Luchinat, C., Mao, J., Pintacuda, G., 2010b. Ultrafast MAS solid-state NMR permits extensive 13C and 1H detection in paramagnetic metalloproteins. J. Am. Chem. Soc. 132, 5558–5559. https://doi.org/10.1021/ja10399g
Berti, I., Luchinat, C., Parghi, G., Ravera, E., Reif, B., Turano, P., 2011. Solid-state NMR of proteins sedimented by ultracentrifugation. Proc. Natl. Acad. Sci. U. S. A. 108, 10396–10399. https://doi.org/10.1073/pnas.1103854108
Blanco, F.J., Montoya, G., 2011. Transient DNA-NMR protein interactions. FERS J. 278, 1643–1650. https://doi.org/10.1111/j.1742-4658.2011.08095.x
Bonneau, E., Legault, P., 2014. NMR localization of divalent cations at the active site of the neurospora V sulfiyube reveals insights into RNA-metal ion interactions. Biochemistry 53, 579–590. https://doi.org/10.1021/bi401484a
Bouret, J., Devillier, J.C., Wiegand, T., Salmon, L., Meier, B.H., Lipps, G., Allain, F.H.T., 2011. A new family of peptide-nucleic acid nanostructures with potent transfection activities. J. Pept. Sci. 17, 88–93. https://doi.org/10.1002/jps.21318
Brown, J.D., Kharytonych, S., Chaudry, I., Iyer, A.S., Carter, H., Becker, G., Desal, Y., Glang, L., Choi, S.H., Singh, K., Lopresti, M.W., Orellana, M., Rodriguez, T., Oboh, U., Hijji, J., Gingham, F.G., Stewart, K., Francis, D., Edwards, B., Chen, P., Case, D.A., Teletinsky, A., Summers, M.F., 2020. Structural basis for transcriptional start site control of HIV-1 RNA fate. Science. 368, 413–417. https://doi.org/10.1126/science.aaz7959
Cady, S.D., Schmidt-Rohr, K., Wang, J., Soto, C.S., Degrado, W.F., Hong, M., 2010. Structural basis for transcriptional start site control of HIV-1 RNA fate. Science. 368, 413–417. https://doi.org/10.1126/science.aaz7959
Campagne, S., Gervais, V., Milton, A., 2011. Nuclear magnetic resonance analysis of protein-DNA interactions. J. R. Soc. Interface 8, 1065–1078. https://doi.org/10.1098/rsif.2010.0543
Carломagno, T., 2014. Present and future of NMR for RNA-protein complexes: A perspective of integrated structural biology. J. Magn. Reson. 241, 126–136. https://doi.org/10.1016/j.jmr.2013.10.007
Castellarini, F., Van Rossum, B.J., De Jong, A., Rebbein, K., Oschkinat, H., 2003. Determination of solid-state NMR structures of proteins by means of three-dimensional 15N–13C–13C dipolar correlation spectroscopy and chemical shift analysis. Biochemistry 42, 11476–11483. https://doi.org/10.1021/bi026490z
Cech, T.R., Steitz, J.A., 2014. The noncoding RNA revolution - Trashing old rules to forge new ones. Cell 157, 77–94. https://doi.org/10.1016/j.cell.2014.03.008
Choy, J.S., Wei, S., Lee, J.Y., Tan, S., Chu, S., Lee, T.H., 2010. DNA methylation increases C:1012911329730.
Cierpicki, T., Otlewski, J., 2001. Amide proton temperature coefficients as hydrogen bond indicators in proteins. J. Biomol. NMR 21, 249–262. https://doi.org/10.1021/bi401484a
Cochran, C., Henschel, A., Kaczka, J., Hacq, J.G., Dervan, P.B., Kessler, H., 1998. Structure of a steroregular phosphorothioate DNA/RNA duplex. Nat. Struct. Biol. 5, 271–276.
Coles, M., Rasmussen, J.T., Nielsen, N.C., 2000. SIMPSON: A General Simulation Program for Solid-State NMR Spectroscopy. J. Magn. Reson. 147, 296–330. https://doi.org/10.1006/jmre.2000.2179

Marchanka, A., Kreutz, C., Carlomagno, T., 2018a. Isotope labeling for studying RNA by solid-state NMR spectroscopy. J. Biomol. NMR 71, 151–164. https://doi.org/10.1007/s10858-018-10095-6.

Marchanka, A., Stanek, J., Pintacuda, G., Carlomagno, T., 2018b. Rapid access to RNA protonate-determined solid NMR at MHz frequencies. 1. J. Biomol. NMR 73, 19–29. https://doi.org/10.1007/s10858-018-0219-9.

Pintacuda, G., Giraud, N., Pierattelli, R., Bockmann, A., Bertini, I., Emde, L., 2006. Solvent effect on NMR properties of a paramagnetic chaperonin in action reveals how nucleotide binding regulates the functional cycle. J. Biomol. NMR 34, 897–905. https://doi.org/10.1007/s10858-005-1614-6.

Riedel, K., Leppert, J., Ohlenenschläger, O., Gorlach, M., Ramachandran, R., 2005a. Characterisation of hydrogen bonding networks in RNAs via magic angle spinning solid state NMR spectroscopy. J. Biomol. NMR 31, 231–336. https://doi.org/10.1007/s10858-005-1614-6.

Riedel, K., Leppert, J., Ohlenenschläger, O., Gorlach, M., Ramachandran, R., 2005b. TEDOR with aladiative inversion pulses: Resonance assignments of 13C/15N labelled proteins. J. Biomol. NMR 31, 499–506. https://doi.org/10.1007/s10858-004-4666-x.

Riedel, K., Herbst, C., Häfner, S., Leppert, J., Ohlenenschläger, O., Swanson, M.S., Gorlach, M., Ramachandran, R., 2006. Constraints on the structure of (CUG)77 RNA from magic-angle-spinning solid-state NMR spectroscopy. Angew. Chem. Int. Ed. 45, 6260–6263. https://doi.org/10.1002/anie.200600769.

Rosenzweig, R., Kay, L.E., 2014. Bringing dynamic molecular machines into focus by methyl-TROSY NMR. Annu. Rev. Biochem. 83, 291–315. https://doi.org/10.1146/annurev-biochem-060712-144631.

Roy, S., Caruthers, M., 2013. Synthesis of DNA/RNA and their analog via phosphoramide and H-phosphonate chemistries. Molecules 18, 14268–14284. https://doi.org/10.3390/molecules181114268.

Schroder, M., Malar, A.A., Torosyan, A., Pencel, S., Klose, D., Ozs, A., Org., M.L., Wang, S., Lecoq, L., Cadalbert, R., Samoson, A., Bockmann, A., Meier, B.H., 2020. Protein NMR Spectroscopy at 150 kHz Magic-Angle-Spinning Continues To Improve Resolution and Mass Sensitivity. ChemBioChem 21, 2540–2548. https://doi.org/10.1002/chem.202002064.

Schludt, A., Tants, J.N., Sattler, M., 2017. Integrated structural biology to unravel molecular mechanisms of protein-RNA recognition. Methods 118–119, 119–136. https://doi.org/10.1016/j.ymeth.2017.03.015.

Schülen, W., Von Ahlen, U., 1997. Identification of specific R-pHosphate oxygenes in the tirannocoid loop required for ribosomal P-site binding. Proc. Natl. Acad. Sci. U. S. A. 94, 12823–12828. https://doi.org/10.1073/pnas.94.24.12823.

Schüle, T., Le, T., Andreas, B.B., Pintacuda, G., 2018. 1H magic angle spinning NMR evolves as a powerful new tool for membrane protein in vitro solution. J. Magn. Reson. 287, 140–152. https://doi.org/10.1016/j.jmr.2017.11.014.

Scott, L.G., Geierstanger, B.H., Williamson, J.R., Hennig, M., 2004. Enzymatic synthesis and 19F NMR studies of 2-fluoroadenine- substituted RNA. J. Am. Chem. Soc. 126, 17729–17737. https://doi.org/10.1021/ja047556g.

Sergeyev, I.V., Day. L.A., Goldburt, A., McDermott, A.E., 2011. Chemical shifts for the unusual DNA structure in PII bacteriophage from dynamic-nuclear-polarization-enhanced solid-state NMR spectroscopy. J. Am. Chem. Soc. 133, 20208–20217. https://doi.org/10.1021/ja206712x.

Shahid, S.A., Markovic, S., Linke, D., Van Rossum, B.J., 2009. Assignment and secondary structure of the YadA membrane protein by solid-state NMR. Sci. Rep. 2, 1–10. https://doi.org/10.1038/srep00127.

Shi, L., Lake, E.M.R., Ahmed, M.A.M., Brown, L.S., Ladizhansky, V., 2009. Solid-state NMR study of proteorhodopsin in the lipid environment: Secondary structure and dynamics. Biochim. Biophys. Acta - Biomembr. 1786, 2563–2574. https://doi.org/10.1016/j.bbamem.2009.09.014.

Siemer, A.B., Huang, K.Y., McDermott, A.E., 2012. Protein Linewidth and Solvent Dynamics in Frozen Solution NMR. PLoS ONE 7. https://doi.org/10.1371/journal.pone.0047242.

Siemer, A.B., McDermott, A.E., 2008. Solid-state NMR on a type III antifreeze protein in the presence of ice. J. Am. Chem. Soc. 130, 17394–17399. https://doi.org/10.1021/ja078493v.

Simon, B., Kirkpatrick, J.P., Eckhardt, S., Reuter, M., Rocha, E.A., Andrade-Navarro, M., 2013. Chemokine receptor CXCR1 in phospholipid bilayers. Nature 491, 779–783. https://doi.org/10.1038/nature11341.

Sprangers, R., Kay, L.E., 2007. Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature 445, 618–622. https://doi.org/10.1038/nature05512.

Studebaker, D.R., Klug, C.A., Besen, D.D., McDowell, L.M., Schafer, J., 1996. Long-range distance measurements of protein binding sites by rotational-echo double-resonance NMR. J. Am. Chem. Soc. 118, 5476–5477. https://doi.org/10.1021/ja951059a.

Su, Y., Andreas, L., Griffen, R.G., 2015. Magic angle spinning NMR of proteins: High-frequency dynamic nuclear polarization and 1H detection. Annu. Rev. Biochem. 84, 497–531. https://doi.org/10.1146/annurev-biochem-060414-010506.

Szeverenyi, N.M., Sullivan, M.J., Maciel, G.E., 1982. Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic angle spinning. J. Magn. Reson. 47, 462–475. https://doi.org/10.1016/0022-2364(82)90213-5.

Togashi, K., Nakamura, S., Terao, T., 1999. 13C–13C polarization transfer by resonant interference uncoupling under magic-angle spinning in solid-state NMR. Chem. Phys. Lett. 307, 295–302. https://doi.org/10.1016/S0009-2614(99)00353-3.

Takagoshi, K., Nakamura, S., Terao, T., 2001. 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. Chem. Phys. Lett. 344, 631–637. https://doi.org/10.1016/S0009-2614(00)00782-1,
