REVIEW

Structural biology of supramolecular assemblies by magic-angle spinning NMR spectroscopy

Caitlin M. Quinn1,2 and Tatyana Polenova1,2*

1 Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716, USA
2 Pittsburgh Center for HIV Protein Interactions, University of Pittsburgh School of Medicine, Pittsburgh, PA 15266, USA

Quarterly Reviews of Biophysics (2017), 50, e1, page 1 of 44 doi:10.1017/S0033583516000159

Abstract. In recent years, exciting developments in instrument technology and experimental methodology have advanced the field of magic-angle spinning (MAS) nuclear magnetic resonance (NMR) to new heights. Contemporary MAS NMR yields atomic-level insights into structure and dynamics of an astounding range of biological systems, many of which cannot be studied by other methods. With the advent of fast MAS, proton detection, and novel pulse sequences, large supramolecular assemblies, such as cytoskeletal proteins and intact viruses, are now accessible for detailed analysis. In this review, we will discuss the current MAS NMR methodologies that enable characterization of complex biomolecular systems and will present examples of applications to several classes of assemblies comprising bacterial and mammalian cytoskeleton as well as human immunodeficiency virus 1 and bacteriophage viruses. The body of work reviewed herein is representative of the recent advancements in the field, with respect to the complexity of the systems studied, the quality of the data, and the significance to the biology.

1. Introduction 2

2. Current methodology for structural and dynamics analysis of biological assemblies by MAS NMR 2
   2.1. Isotopic labeling 4
   2.2. Resonance assignments and structure determination 4
      2.2.1. Through-space multidimensional correlation spectroscopy 5
      2.2.2. Through-bond multidimensional correlation spectroscopy 6
      2.2.3. Proton detection and fast MAS 6
      2.2.4. Protein structure determination by MAS NMR 7
   2.3. MAS NMR for the study of protein dynamics 9
      2.3.1. Microsecond to nanosecond timescale dynamics 10
      2.3.2. Millisecond to microsecond timescale dynamics 12
   2.4. Intermolecular interactions 13
      2.4.1. Chemical shift perturbations 13
      2.4.2. Dipolar-edited correlation spectroscopy 14

3. MAS NMR of cytoskeleton-associated proteins 15
   3.1. MTs and MT-associated proteins (MAPs) 16
      3.1.1. Structure of CAP-Gly domain of dynactin 16
      3.1.2. Interface of CAP-Gly with MTs 17

* Author for correspondence: T. Polenova, Department of Chemistry and Biochemistry, University of Delaware, 036 Brown Labs, Newark, DE 19716, USA.

Email: tpolenov@udel.edu; Tel.: 302-831-1968


Downloaded from https://www.cambridge.org/core. IP address: 54.70.40.11, on 31 Jan 2020 at 11:51:05, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0033583516000159
3.1.3. Dynamics of CAP-Gly 17
3.2. Bactofilins 18

4. MAS NMR of viral assemblies and intact viral particles 19
4.1. HIV-1 capsid and maturation intermediates 20
4.1.1. Structural characterization of HIV-1 capsid assemblies 21
4.1.2. Conformational dynamics of HIV-1 capsid assemblies by MAS NMR 23
4.1.3. MAS NMR of HIV-1 maturation intermediates 26
4.2. Bacteriophages 28
4.2.1. Structural characterization of bacteriophage capsid proteins 28
4.2.2. Characterization of bacteriophage capsid dynamics 29
4.2.3. Characterization of protein–nucleic acid interactions in bacteriophages 30
4.3. Other viral assemblies 31

5. Conclusions and future perspectives 32

Acknowledgements 32

References 32

1. Introduction
In the past decade, the field of magic-angle spinning (MAS) nuclear magnetic resonance (NMR) has made significant strides. This technique has advanced to the level where we can now determine structures and characterize dynamics of complex systems, including large protein assemblies, at atomic resolution. A decade ago, this effort was in its infancy with the demonstration of the proof of principle that structures of small proteins can be solved de novo. Now we are tackling a wide range of biologically pressing problems, where traditional techniques yield only limited insights or are powerless. Recent instrument technology and methodological advancements have been conducive to the study of increasingly complex biological systems. Such advancements include the development of fast MAS capabilities (up to ~110 kHz at present) and very high magnetic fields (up to 1 GHz at present with 1.2 GHz magnets currently in production) that yield unprecedented gains in sensitivity and resolution and enable proton detection (Holland et al. 2010; Lewandowski et al. 2011a; Zhou et al. 2007b).

As a biophysical method, MAS NMR offers many advantages over other techniques. There are no theoretical size limitations (though challenges with respect to sensitivity and resolution arise with increasing molecular weight), no solubility limitations, and no requirements for well-formed crystals or long-range order. MAS NMR can achieve atomic-level resolution and also tackle very large systems such as whole cells and intact viral particles. MAS NMR can probe both structure and dynamics at or close to physiologically relevant experimental conditions including temperature and pH. These advantages allow for the characterization of highly complex biological systems to address compelling questions in biology. MAS NMR can provide unique insights into an astounding range of biological systems, including proteins embedded in native membrane environments (Brown & Ladizhansky, 2015; Naito et al. 2015), aggregates of misfolded or disordered proteins (Comellas & Rienstra, 2013; Tycko, 2011), biomaterials (Goobes, 2014), and metalloproteins (Jaroniec, 2012; Knight et al. 2012, 2013). MAS NMR is also well suited for the study of biological assemblies comprised of multiple components or multiple copies of the same molecule, including entire viruses and cells (Goldburt, 2013; Loquet et al. 2013b; Weingarth & Baldus, 2013; Yan et al. 2013b). In this review, we will discuss the current MAS NMR methodology for structural and dynamics studies of biological systems with specific focus on applications to supramolecular assemblies represented by proteins associated with the cytoskeleton and viral assemblies.

2. Current methodology for structural and dynamics analysis of biological assemblies by MAS NMR
The general work flow for MAS NMR studies of biological systems (Fig. 1) first entails preparation of samples isotopically enriched with NMR active nuclei (namely, $^{13}$C and $^{15}$N). Proteins are subsequently prepared for MAS NMR studies by crystallization (Martin & Zilm, 2003), assembly, sedimentation (Bertini et al. 2011b, 2013), or similar approaches, and packed into rotors. Optimization is key as sample conditions can have a significant impact on spectral quality. Before advanced structural or dynamics studies can be executed, site-specific resonance assignments must be obtained. This is accomplished by acquiring a suite of multidimensional spectra (typically two-dimensional (2D) and three-dimensional (3D)), and establishing through-
space (dipolar) and/or through-bond (scalar) intra- and inter-residue homonuclear correlations and heteronuclear correlations (HETCOR).

MAS NMR can access information of great interest in biology, including protein structure and dynamics, as well as protein–protein and protein–ligand interactions. Protein structure determination by MAS NMR requires the quantification of structural restraints, such as short- and long-range (separated by more than four residues) internuclear distances, as well as backbone torsion angles. Structural restraints are integrated into simulated annealing protocols for structure calculation. MAS NMR has been increasingly combined with other biophysical methods such as cryo-electron microscopy (cryo-EM) for

Fig. 1. Workflow for studies of biological supramolecular assemblies by MAS NMR. Preparation of homogeneous, isotopically labeled samples and resonance assignments are the first steps of any structural biology study by MAS NMR. Resonance assignments and other experiments exploit two types of inter-nuclear correlations: through-space (dipolar-based), which selects for rigid residues, and through-bond (scalar or J coupling based), which selects for dynamic residues. Biological questions that can be addressed by MAS NMR include structure determination, protein dynamics, and intermolecular interactions. Protein structure determination generally entails first obtaining long-range, inter-nuclear distance correlations, often combined with other structural restraints, and subsequently input into simulated annealing protocols for structure calculation. Two approaches commonly used for the determination of site-specific millisecond to nanosecond protein dynamics are relaxation dispersion and measurement of reduced anisotropic interactions (e.g., chemical shift anisotropy or dipolar interactions). Finally, MAS NMR can characterize protein–protein and protein–ligand interactions. Methods for observing these intermolecular interfaces include chemical shift perturbations, dipolar filtered experiments such as dREDOR, and quantitative distance measurements with REDOR/TEDOR-based experiments. Isotopic labeling schematic reprinted with permission from Higman et al. (2009). Copyright 2009 Springer. Sedimented solute NMR (SedNMR) figure adapted with permission from Bertini et al. (2013). Copyright 2013 American Chemical Society. CA-SP1 A92E TEM image and through-space and through-bond correlation experiments reprinted with permission from Han et al. (2013). Copyright 2013 American Chemical Society. Structure determination and chemical shift perturbation figures adapted with permission Yan et al. (2013a). Copyright 2013 Elsevier. Anisotropic spin interactions and protein dynamics/structure figures adapted with permission from Lu et al. (2015a). Copyright 2015 National Academy of Sciences. dREDOR figure and CAP-Gly/MT complex TEM adapted with permission from Yan et al. (2015a). Copyright 2015 National Academy of Sciences. TEDOR/REDOR distances figure reprinted with permission from Nieuwkoop & Rienstra (2010). Copyright 2010 American Chemical Society. Relaxation dispersion figure reprinted with permission from Lewandowski et al. (2011b). Copyright 2011 American Chemical Society.
structure determination of supramolecular assemblies. Anisotropic interactions, such as magnitudes and orientations of dipolar and chemical shift tensors are highly sensitive to both structure and dynamics. A wide range of methods exists to study protein dynamics with MAS NMR over timescales from picoseconds to seconds. Chemical shift perturbations and dipolar-edited correlation methods can be used to characterize protein–protein and protein–ligand interactions.

Two essential considerations for successful MAS NMR experiments are sensitivity and resolution, which can be affected by numerous factors such as protein size and dynamics, sample homogeneity, and nuclear spin interactions including dipolar and J (scalar) couplings. Challenges related to resolution and sensitivity can be alleviated or overcome using advanced hardware (i.e., faster spinning probes and higher magnetic fields) and appropriate choice of isotopic labeling schemes. Spectroscopic methods employed are a fundamental factor to maximize sensitivity and resolution, including choice of magnetization transfer method (i.e., through-space versus through-bond) and detection method (1H versus heteronuclear detection). In the following sections, we provide an overview of methods commonly employed for the study of biological systems including supramolecular assemblies by MAS NMR.

2.1 Isotopic labeling

Isotopic enrichment with magnetically active 13C and 15N is essential for the study of proteins by NMR. Beyond uniform isotopic labeling with 13C-glucose and 15NH4Cl, there are many alternative labeling schemes for selective incorporation of isotopes into desired sites. Spectral crowding is a substantial challenge in MAS NMR, and beyond 3D- and 4D spectra, higher magnetic fields, and fast MAS, isotope editing is used to alleviate the congestion. Sparse as well as selective isotopic labeling methods are often employed for the determination of long-range 13C–13C distance restraints and torsion angles. The common protocols include preparation of recombinant proteins from minimal media containing [2-13C]glycerol, [1,3-13C]glycerol, [1,6–13C]glucose, and [2-13C]glucose as the sole carbon source (Higman et al. 2009; Hong, 1999; LeMaster & Kushlan, 1996). Selective labeling with [2-13C]glycerol and [1,3-13C]glycerol was essential to the first protein structure determination by MAS NMR (Castellani et al. 2002). These labeling schemes exploit bacterial metabolic pathways to achieve known patterns of amino acid labeling (Goldbourt et al. 2007a). These selective labeling schemes also serve to reduce line broadening by reducing strong 13C–13C dipolar couplings and J-couplings. For further spectral simplification, amino acid specific labels can be incorporated (Mcintosh & Dahlquist, 1990), which also allow for the study of critical protein properties, such as amino acid protonation state, as well as the determination of select distance restraints with fewer ambiguities. With His-to-Gln mutations and selective labeling of His37 of M2(21–97), Hong and co-workers showed that the protonation state of this transmembrane domain residue is perturbed by the presence of the cytoplasmic domain, suggesting a mechanism of 1H conduction (Liao et al. 2015). Perdeuteration with back exchange of amide protons enables the acquisition of high-resolution proton-detected spectra and the determination of 1H–1H distance restraints by reducing the very strong 1H–1H dipolar couplings (Chevelkov et al. 2006; Reif et al. 2003; Zhou et al. 2007b) and is further discussed below. Additional selective 13C and 15N labeling schemes for aliphatic groups of Ala, Val, Leu, and Ile, developed by Kay and co-workers (Rosen et al. 1996) and first applied in the solid state by Reif and co-workers (Agarwal et al. 2006), can also be used for structural restraints as demonstrated for structure determination of ubiquitin (Agarwal et al. 2014). To characterize intermolecular interfaces and distances, differential labeling schemes have been developed. In this family of labeling schemes one region of the protein, monomer in an assembly, or binding partner contains one set of labels (e.g., 13C or 13C,15N), while its interaction partner has different labeling (e.g., 15N). In these differentially labeled samples, intermolecular interactions are then measured by experiments where magnetization is selectively transferred across the intermolecular interface, demonstrated for distance determination of select 13C–13N spin pairs in gramicidin A in early work (Fu et al. 2000), and later applied to protein studies by Baldus (Etzkorn et al. 2004) and Polenova (Marulanda et al. 2004; Yang et al. 2008). Generally, there are many isotopic labeling approaches available to an experimentalist, and an appropriate combination of isotopic labeling schemes is selected to address specific questions.

2.2 Resonance assignments and structure determination

Performing resonance assignments entails obtaining homo- and heteronuclear intra-residue and sequential inter-residue correlations and is the necessary first step to any study of protein structure or dynamics by MAS NMR. Early work of note includes complete or near complete resonance assignments of BPTI (58 residues, McDermott et al. 2000), SH3 (62 residues, Pauli et al. 2001), ubiquitin (76 residues, (Igumenova et al. 2004a, b), and Crh (85 residues, (Bockmann et al. 2003)). Isotropic chemical shifts yield information on secondary structure, protonation states, and dynamics (Williamson, 1990; Wishart & Sykes, 1994). For structure determination, long-range distance restraints must also be obtained. Determining resonance assignments and distance restraints requires collecting a suite of multidimensional spectra using dipolar and/or scalar based correlations. From homonuclear correlations and HETCOR experiments, the spin system belonging to a given amino
acid is first identified from experiments including 2D $^{13}$C–$^{13}$C experiments and 2D/3D $^{15}$N–$^{13}$C NCACX experiments. Inter-residue correlation experiments such as 2D/3D NCOCX are then used to establish sequential, residue specific assignments. These experiments are further detailed in Section 2.2.1. In large systems, assignments could be challenging due to spectral congestion and typically require a large number of experiments in conjunction with sparse isotopic labeling discussed above, as demonstrated for assignment of the 189 residue protein DsbA by Rienstra and co-workers (Sperling et al. 2010). Modern technological advancements including fast MAS (frequencies of 40–110 kHz), which provides both sensitivity and resolution enhancement (Barbet-Massin et al. 2014b; Bertini et al. 2010; Laage et al. 2009; Parthasarathy et al. 2013; Samoson et al. 2005), and proton detection (Chevelkov et al. 2006; Paulson et al. 2003; Zhou et al. 2007a) enabled the development of new experiments for time-efficient resonance assignments and recording distance restraints. Fast MAS and proton detection are further discussed in Section 2.2.3.

### 2.2.1 Through-space multidimensional correlation spectroscopy

Through-space correlation experiments rely on distance-dependent internuclear dipolar couplings ($D_{ij} \propto \gamma_i \gamma_j / r^3$). Observed correlations can be short or long range, depending on the chosen pulse sequence and experimental parameters (e.g., mixing time). Early through-space correlation experiments were optimized for MAS frequencies of 10–30 kHz. With advances in probe technology and faster spinning speeds, methods have been developed to achieve efficient polarization transfer at higher MAS rates. Common through-space homonuclear correlation experiments optimized for the slower spinning regime (10–30 kHz) include dipolar-assisted rotational resonance (DARR) (Takegoshi et al. 2001), RF-assisted diffusion (RAD) (Morcombe et al. 2004), proton-driven spin diffusion (PDSD) (Szeverenyi et al. 1982), and dipolar recoupling enhanced by amplitude modulation (DREAM) (Verel et al. 2001) to obtain $^{13}$C–$^{13}$C correlations. Some applications of note on biological assemblies include PDSD for resonance assignments and structure determination of the type III secretion system (T3SS) needle (Demers et al. 2014; Loquet et al. 2011), BacA filament (Shi et al. 2015; Vasa et al. 2015), HET-s amyloid (Wasmer et al. 2008), and DARR for detection of the Pf1 bacteriophage DNA signals (Sergeyev et al. 2011) and characterization of the human immunodeficiency virus 1 (HIV-1) capsid and CA-SP1 maturation intermediate (Han et al. 2010, 2013). For determination of heteronuclear NCA, NCO, NCACX, and NCOCX correlations, $^{15}$N–$^{13}$C double cross-polarization (DCP), first presented by Schaefer et al. (1979), is commonly employed. Baldus et al. developed frequency-selective DCP (known as SPECIFIC-CP) (Baldus et al. 1998) for selective NCA or NCO excitation, as demonstrated for resonance assignments of SH3 (Pauli et al. 2001). SPECIFIC-CP has been shown to be broadly applicable (Luca et al. 2003). Other recoupling sequences such as dipolar insensitive nuclei enhanced by polarization transfer (INEPT) for selective C–H excitation at both moderate (De Vita & Frydman, 2001; Wickramasinghe et al. 2008) and fast (Holland et al. 2010) spinning speeds have been also reported.

At MAS frequencies faster than 30 kHz, the conventional spin diffusion-based experiments for recording homonuclear correlations are no longer efficient. Under these conditions, DREAM and fpRFD (finite pulse rf driven recoupling (Ishii, 2001)) are efficient for recording one- and two-bond correlations. Another family of experiments that is particularly useful for recording long-range $^{13}$C–$^{13}$C distance restraints is COMbined R2$_n$-Driven (CORD) dipolar recoupling sequences, where the magnetization transfer is driven by rotor-synchronized R2$_n$ symmetry-based recoupling (Hou et al. 2011a, 2013a; Lu et al. 2015b). The RN$n$ and CN$n$ symmetry recoupling schemes were originally presented by Levitt and co-workers (Carravetta et al. 2000). CORD utilizes a super-cycled R2$_n$ recoupling to achieve broadband homonuclear correlations with high polarization transfer efficiency at both moderate and fast MAS rates while not suffering from dipolar truncation effects. Proton-assisted recoupling (PAR) is another method that performs well at fast MAS to obtain long distance $^{13}$C–$^{13}$C (De Paepe et al. 2008; Lewandowski et al. 2009b) or $^{13}$N–$^{15}$N (Lewandowski et al. 2009a) correlations. PAR is based on third spin assisted recoupling (TSAR), in which two spins are connected via dipolar couplings with a third spin leading to zero quantum (ZQ) polarization transfer. Distances of ~6–7 Å can be observed with PAR and CORD.

Beyond DCP, several methods have been developed for the acquisition of long range $^{15}$N–$^{13}$C correlations. PAIN-CP (proton-assisted insensitive nuclei cross-polarization) is a third-spin-assisted heteronuclear polarization transfer (Agarwal et al. 2013; De Paepe et al. 2011; Lewandowski et al. 2007) first presented by Griffin and co-workers, which like its homonuclear counterpart discussed above, utilizes neighboring proton spins to enhance magnetization transfer efficiency with appropriate choice of $^{13}$C, $^{15}$N, and $^1$H rf fields. Transferred echo double resonance (TEDOR) (Hing et al. 1992) is a REDOR (rotational echo double resonance (Gullion & Schaefer, 1989)) derived scheme that can also be used to detect $^{15}$N–$^{13}$C distances up to ~8 Å. In REDOR-based pulse sequences, the dipolar coupling between two spins is reintroduced by a train of rotor-synchronized 180° pulses (Gullion & Schaefer, 1989). The resulting dephasing of magnetization is proportional to the magnitude of the dipolar coupling (and hence distance between the two spins). A variation of the TEDOR pulse sequence developed by Jaroniec et al. (2002), z-filtered TEDOR, is shown in Fig. 2e. The inclusion of a $z$-filter is needed to eliminate artifacts due to $^{13}$C–$^{13}$C J couplings in uniformly labeled systems. TEDOR-derived distance restraints have been
applied to a range of systems including structure determination of microcrystalline GB1 by Rienstra and co-workers (Nieuwkoop et al. 2009) and L7Ae-bound Box C/D RNA by Carломagno and co-workers (Marchanka et al. 2015). Pulse sequences, schematics, and model compound spectra for several through-space correlation methods are presented in Fig. 2.

2.2.2 Through-bond multidimensional correlation spectroscopy

Complementary to through-space, dipolar-based correlation experiments, scalar-based through-bond transfer mechanisms can be exploited to obtain inter-nuclear correlations. Through-bond experiments utilize the electron-mediated J coupling between neighboring atoms. This transfer mechanism can be especially valuable in cases of dynamics (Heise et al. 2005) and at fast MAS frequencies, situations where dipolar couplings are partially or fully averaged. J-based experiments are also ideal at faster spinning frequencies due to the lower required decoupling power, and can allow for the necessary longer coherence evolution times (Bertini et al. 2011a). Experiments such as heteronuclear (Elena et al. 2005) or homonuclear (Linser et al. 2008) INEPT (Morris & Freeman, 1979), homonuclear total through-bond correlation spectroscopy (TOBSY) (Hardy et al. 2001), homonuclear constant-time uniform sign cross-peak COSY (CTUC-COSY), (Chen et al. 2006), as well as solid-state INADEQUATE (Lesage et al. 1997) and refocused INADEQUATE (Lesage et al. 1999) have been used to complement dipolar-based correlation spectroscopy in the study of protein assemblies (Fig. 3). With these methods, sufficient sensitivity is attained despite the relatively small size of the J-couplings (e.g. 50 Hz 13C–13C J coupling versus 2 kHz dipolar coupling). TOBSY experiments utilize the POST-C7 symmetry sequence (Hohwy et al. 1998) to achieve efficient, scalar-based polarization transfer. CTUC-COSY offers excellent sensitivity by converting both zero-quantum and double-quantum magnetization, and has been applied to detect dynamic regions of the Y145Stop human prion protein (Helmus et al. 2010) and α-Synuclein fibrils (Comellas et al. 2011), as well as to obtain pure one-bond correlations in 13C–13C spectra of CAP-Gly (Sun et al. 2009). INADEQUATE experiments in the solid state use double-quantum coherence transfer identical to solution NMR. More recent modifications of solid-state INADEQUATE have included the addition of a z-filter (Cadars et al. 2007) and FSLG (frequency-switched Lee–Goldberg) homonuclear 1H–1H decoupling (Baltisberger et al. 2011) to reduce artifacts, and development of band-selective INADEQUATE using the spin state selective excitation (S2E) scheme, which has been demonstrated at 60 kHz MAS (Bertini et al. 2011a). The use of scalar transfers in proton-detected experiments at fast MAS has recently been demonstrated in the solid state as well including 13C–13C INEPT transfer for resonance assignments of superoxide dismutase (SOD) (Knight et al. 2011). Pintacuda and co-workers reported the application of ‘out-and-back’ 13C–13C scalar-based transfer for resonance assignments with fast MAS (frequencies of 60 kHz and higher (Barbet-Massin et al. 2013)). These proton-detected 3D experiments can be applied to both fully protonated samples as well as perdeuterated samples with 100% H/H2 back exchange, and were demonstrated on AP205 bacteriophagase as well as numerous other diverse classes of proteins (Barbet-Massin et al. 2014b).

2.2.3 Proton detection and fast MAS

In contrast to solution NMR where dipolar couplings are averaged out by molecular tumbling, the strong 1H–1H dipolar couplings present in solid-state NMR (SSNMR) samples lead to very broad 1H lines. As a consequence, MAS NMR experiments have customarily been acquired with direct detection of low γ nuclei such as 13C and 15N, which greatly limits sensitivity. Proton detection takes advantage of the high gyromagnetic ratio of protons for increased sensitivity and with advances in hardware is increasingly applied in SSNMR. Early work by Reif, Griffin, and Zilm demonstrated that with perdeuteration to reduce 1H–1H dipolar couplings and 100% amide 1H-back exchange, proton-detected HETCOR experiments could be applied in the solid state and that the anticipated sensitivity gains are realized, while dipolar truncation is avoided (Paulson et al. 2003; Reif & Griffin, 2003; Reif et al. 2001). Subsequent work demonstrated that increased 1H resolution can be achieved with higher levels of deuteration (i.e., only 10–40% 1H back exchange) (Akbey et al. 2010) and/or faster MAS frequencies (Chevelkov et al. 2006; Samoson et al. 2001). Linser et al. first demonstrated the application of 1H detection to amyloids and membrane proteins (Linser et al. 2011b). With the advent of fast MAS (>40 kHz), proton-detection even on fully protonated proteins, first demonstrated by Rienstra and co-workers (Zhou et al. 2007a), has become feasible with improvements in resolution and sensitivity scaling with the MAS rate (Agarwal et al. 2014; Lewandowski et al. 2011a; Marchetti et al. 2012). Further, the sensitivity gains of 1H detection enable the use of very small sample amounts (Agarwal et al. 2014; Dannatt et al. 2015). Recent works of note in the application of proton detection include studies of RNA–protein interfaces by Asami et al. (2013), structure determination of SOD by Knight et al. (2012), and measurements of heteronuclear dipolar couplings in Pf1 bacteriophage by Opella and co-workers (Park et al. 2013). Additional capabilities of 1H detection include obtaining direct information on hydrogen-bond length from 1H chemical shifts (Zhou & Rienstra, 2008). Pintacuda and co-workers have recently demonstrated a suite of 3D proton-detected experiments to enable rapid data acquisition and assignment, with data sets of sufficient quality for the automated assignment routines to be applicable
Barbet-Massin et al. (2014b). They demonstrated the use of these sequences on several challenging systems, including assemblies of AP205 bacteriophage and Measles virus (MeV) nucleocapsid (Barbet-Massin et al. 2014a, b).

2.2.4 Protein structure determination by MAS NMR

Structure determination by SSNMR was first demonstrated on small peptides oriented in lipid bilayers with static methods by Cross and co-workers (Ketcham et al. 1996; Wang et al. 2001) and Opella and co-workers (Opella et al. 1999) MAS NMR structure determination of a protein was first reported by Oshkinat and co-workers for the α-spectrin SH3 domain (Castellani et al. 2002). Technical and methodological advances have enabled the application of MAS NMR to structure determination of increasingly complex systems (Fig. 4a). MAS NMR is particularly valuable for the high-resolution structure determination of supramolecular assemblies, which are often insoluble or non-crystalline. Protein structure determination by MAS NMR requires determination of a sufficient number of quantitative or semi-quantitative structural restraints including distance restraints obtained from homonuclear correlation and HETCOR spectra, using long-range magnetization transfer techniques described above, which can probe interatomic distances of up to ~7 Å (Fig. 4b). 

13C–13C distances are the most frequently utilized and often make use of selectively labeled samples for spectral simplification and semi-quantitative cross-peak intensity analysis. Additional distance restraints that have been utilized include 15N–13C (Nieuwkoop et al. 2009), 13C–13C (Hu et al. 2012; Lewandowski et al. 2009a), and increasingly 1H–1H (Andreas et al. 2016; Linser et al. 2011a; Zhou et al. 2007b) distances. Very recently, Pintacuda and co-workers presented the first protein structures determined on fully protonated samples with 1H detection (Andreas et al. 2016). They acquired RFDOR-based 1H–1H distance restraints with ≥100 kHz MAS to determine the structures of two proteins: GB1 and the AP205 nucleocapsid assembly. Less than 0.5 mg of U–15N, 13C protein and 2 weeks of experiment time and ‘unsupervised’ structure determination were sufficient to derive the protein structures.
In addition to inter-atomic distance restraints, anisotropic spin interactions including dipolar and chemical shift tensor magnitudes and orientations are a powerful tool for protein structure determination. These interactions exhibit secondary structure, orientation, and amino acid type dependence that can be exploited in structure determination, demonstrated extensively by Rienstra and co-workers on GB1 (Franks et al. 2008; Wylie et al. 2009, 2011). In structure calculations, dipolar couplings and CSA tensors can constrain backbone torsion angles (Ladizhansky et al. 2003; Rienstra et al. 2002) as demonstrated on Aβ11–25 with recoupling of chemical shift anisotropy (ROCSA) measurements (Chan & Tycko, 2003). Recent applications include the use of 1H–15N and 1H–13Cα dipolar couplings determined from separated local field (SLF) measurements (Das et al. 2013) as orientation restraints (with dihedral angles derived from isotropic chemical shifts) in determining the structure of CXCR1, a chemokine receptor involved in inflammatory response ((Park et al. 2012), Fig. 5). Anisotropic spin interactions can also provide valuable dynamics information, as further discussed in Section 2.3. Additional structural restraints that may be incorporated into a structure calculation include predicted torsion angles based on backbone chemical shifts from TALOS (Shen & Bax, 2015; Shen et al. 2009), hydrogen bonding, and paramagnetic relaxation enhancements (PREs, (Nadaud et al. 2007)). PREs utilize the enhanced R1 relaxation of residues in close proximity to a paramagnetic center as a structural restraint and were applied to structure determination of SOD (Knight et al. 2012) and the membrane protein Anabaena sensory rhodopsin (ASR) (Wang et al. 2013). Beyond intra-subunit restraints, long-range distances, such as those acquired in zf-TEDOR experiments, can contribute inter-subunit restraints for structure determination of supramolecular assemblies (Nieuwkoop & Rienstra, 2010).

Structural restraints are incorporated in simulated annealing calculations in a program such as Xplor-NIH (Schwieters et al. 2003, 2006) or CYANA (Guntert, 2004), with optimization protocols, which include molecular dynamics (MD) and Monte
Carlo simulations. Recently, multiple laboratories have demonstrated the use of de novo structure prediction based on isotropic chemical shifts and amino acid sequence with CS-ROSETTA (Das et al. 2009; Shen et al. 2008), without requiring distance restraints. CS-ROSETTA has been incorporated into structure determination of the biological assemblies T3SS (Demers et al. 2014; Loquet et al. 2012) and the M13 bacteriophage (Morag et al. 2015). (Whether this approach is generally applicable to a wide range of systems remains to be investigated.) Rosetta enables the modeling of symmetric macromolecular assemblies and has been a key development for the atomic-resolution structure determination of these large and complex systems (DiMaio et al. 2011).

The capabilities of MAS NMR for structure determination have been further expanded in recent years by the application of integrated structure determination, wherein MAS NMR restraints are combined with other biophysical methods, such as EM, solution NMR, and MD simulations. While exact approaches may differ, in general, the secondary structure as determined from secondary chemical shifts or high-resolution monomeric structure is mapped into the lower resolution electron density map (typically by rigid body modeling), and this structure is further refined with simulated annealing, using structure restraints such as cryo-EM structure factors and NMR distance restraints. Fig. 6 illustrates the iterative protocol. Lower-resolution microscopy can provide information on the symmetry and macromolecular organization, while MAS NMR data provide atomic-level structural details including inter-subunit contacts. This approach is proving to be particularly auspicious for the study of macromolecular assemblies including structure determination of αβ-crystallin (with small-angle X-ray scattering (SAXS), MD (Jehele et al. 2010)), T3SS (with cryo-EM (Demers et al. 2014; Loquet et al. 2012)), DsbB (with X-ray crystallography, MD (Tang et al. 2013)), FimA (with solution NMR, STEM (scanning transmission electron microscopy) (Habenstein et al. 2015)), and the mouse ASC inflammasome (with cryo-EM, Fig. 6 (Sborgi et al. 2015)).

2.3 MAS NMR for the study of protein dynamics

Protein dynamics are an essential attribute of biological function including intra-cellular transport (Desai & Mitchison, 1997) and inter-cellular signaling (Alenghat & Golan, 2013, as well as detrimental pathologies, such as in the case of amyloids (Chiti & Dobson, 2006). Relevant motions include both faster, small-amplitude motions such as backbone fluctuations and larger amplitude motions such as whole domain reorganization (Tzeng & Kalodimos, 2012). In contrast to other techniques that are used for characterization of biomolecular dynamics, such as SAXS, FRET (fluorescence resonance energy transfer), and AFM (atomic force microscopy), NMR (both solution and MAS) yields information on multiple sites within a protein simultaneously. Furthermore, nuclear spin interactions, including the chemical shift, dipolar, and quadrupolar tensors, are sensitive probes of dynamics over many decades of motional timescales, from picoseconds to seconds, making NMR a unique technique for probing motions over the entire range of functionally relevant timescales, often in a single sample as demonstrated for GB1 (Lewandowski et al. 2015) and thioredoxin (Yang et al. 2009). MAS NMR is particularly well suited for probing protein dynamics in large biological assemblies and has shed light on a number of intriguing biological questions, such as gating of membrane proteins (Hu et al. 2010; Wang & Ladizhansky, 2014; Weingarth et al. 2014; Wylie et al. 2014), mechanisms of enzyme catalysis (Caulkins et al. 2015; Rozovsky & McDermott, 2001; Schanda et al. 2014; Ullrich
and the regulation of protein–protein interactions in supramolecular assemblies (Hoop et al. 2014; Opella et al. 2008; Yan et al. 2015b). Unlike in solution NMR, the anisotropic tensorial spin interactions are recorded in MAS NMR rather than the motionally averaged residual interactions. Dipolar, CSA, and quadrupolar tensors contain orientational information and thus bear a wealth of information on the motional symmetry and amplitudes, which can be inferred only indirectly from the isotropic chemical shifts or residual dipolar interactions. Nuclear spin relaxation is also extensively used as a probe of dynamics over a wide range of conditions, and yields unprecedented insights into hierarchical protein dynamics, as was recently demonstrated (Lewandowski et al. 2015). For more extensive review of MAS NMR for the study of protein dynamics see the following review articles (Krushelnitsky et al. 2013; Watt & Rienstra, 2014).

Recently, several groups have presented comprehensive studies of protein dynamics for systems of interest using a combination of the methods described below to gain insight into protein dynamics across multiple timescales. These works include dipolar order parameter (DOP) and 15N 1ρ studies of ASR (Good et al. 2014), R1 and R1ρ studies of GB1 (Lewandowski et al. 2010, 2011b, 2015), SH3 (Zinkevich et al. 2013) and SOD (Knight et al. 2012), 1H–15N DOP, 15N CSA, and peak intensity experiments of thioredoxin (Yang et al. 2009), DOP, 15N CSA, and peak intensity experiments on HIV-1 capsid (Byeon et al. 2012; Lu et al. 2015a), DOP and peak intensity measurements of CAP-Gly (Yan et al. 2015b), and DOP, R1, R1ρ, and R2 studies of ubiquitin (Haller & Schanda, 2013; Schanda et al. 2010).

2.3.1 Microsecond to nanosecond timescale dynamics

Dynamic processes on the microsecond-to-nanosecond timescale include backbone fluctuations and rotation of side chain methyl groups. These motions can be observed with 1H–15N and 1H–13C dipolar as well as 13C, 15N, and 1H chemical shift anisotropy (CSA) tensor measurements, and R1 relaxation experiments. Dynamic processes on this timescale result in narrowing of the tensors below their rigid-limit value (Torchia & Szabo, 1985). The rigid limit 1H–15NH and 1H–13Cα dipolar coupling constants are 11.34 kHz (Yao et al. 2008) and 22.7 kHz (Alkaraghouli & Koetzle, 1975), respectively. Reduced 1H–13C and 1H–15N dipolar coupling strengths due to dynamics are also reflected in peak intensities in 1H–13C or 1H–15N cross polarization experiments as demonstrated by the differential 1H–13N CP buildup for the soluble and transmembrane domains of the N-terminal FMN-binding domain of NADPH-cytochrome P450 oxidoreductase, a redox partner of cytochrome P450 (Huang et al. 2014). For further characterization of microsecond-to-nanosecond dynamics, several methods that rely on quantitative measurement of T1 relaxation, dipolar couplings, and CSA tensors have been developed. Reducing interference from strong 1H–1H couplings and spin diffusion have been important components in the development of methods to study microsecond to nanosecond dynamics.

Longitudinal spin-lattice relaxation (R1) is used to probe protein backbone mobility on the nanosecond timescale. Pseudo-quantitative R1 measurements were first conducted on Crh by Emsley and co-workers (Giraud et al. 2004). Quantitative R1 measurements have subsequently been applied to several microcrystalline systems including GB1 (Lewandowski et al. 2010), SH3 (Zinkevich et al. 2013), and ubiquitin (Haller & Schanda, 2013; Schanda et al. 2010), as
well as the transmembrane protein ASR (Good et al. 2014), the metalloprotein SOD (Knight et al. 2012), and an amyloid-like fragment of the yeast prion protein Sup35p (Lewandowski et al. 2011c). While $^{15}$NR $R_1$ measurements are relatively straightforward, $^{13}$C-$^{13}$C $R_1$ measurements require fast MAS (>60 kHz, (Lewandowski et al. 2010)) or selective labeling (Asami et al. 2015), in order to reduce $^{13}$C–$^{13}$C proton driven spin diffusion.

Dipolar chemical shift correlation (DIPSHIFT), first presented by Griffin and co-workers (Munowitz et al. 1981, 1982) and extended to slower dynamics by DeAzevedo et al. (2008) is a common technique for the measurement of $^1$H–$^{15}$N and $^1$H–$^{13}$C dipolar couplings and characterization of microsecond-to-nanosecond timescale dynamics. In traditional DIPSHIFT experiments, the magnetization evolves under the influence of the heteronuclear dipolar coupling while $^1$H–$^1$H couplings are suppressed with phase-modulated Lee–Goldburg decoupling (PMLG) (Vinogradov et al. 1999). Alternatively, DIPSHIFT-based RN-symmetry recoupling experiments can be used for the measurement of $^1$H–$^{15}$N and $^1$H–$^{13}$C dipolar couplings (Fig. 7, (Hou et al. 2011b)). These sequences selectively reintroduce heteronuclear dipolar couplings while reducing interference from homonuclear dipolar couplings (Carravetta et al. 2000). In addition, these pulse sequences are suitable for fast MAS frequencies and can be used in fully protonated systems (Hou et al. 2011b). In these R $\nu$ symmetry sequences, the heteronuclear dipolar coupling is reintroduced with a rotor-synchronized RN $\nu$ radio frequency pulse train applied on the proton channel. An NCA, NCO, or $^{13}$C–$^{13}$C correlation dimension is incorporated for site-specific determination of dynamics. Recently a modification of the RN-DIPSHIFT experiment, phase-alternating R-symmetry (PARS), was developed (Hou et al. 2014). This sequence incorporates a phase-shifted RN symmetry block applied on $^1$H, with $\pi$ pulses applied on the X channel and efficiently suppresses effects from the $^1$H CSA. Further, the series of X channel pulses refocuses the chemical shift, eliminating the need for a Hahn echo and giving the experiment inherently higher sensitivity than RN-DIPSHIFT. Under fast MAS conditions (>60 kHz), cross-polarization with variable contact time (CPVC) is another promising approach for characterization of motions on these timescales (Paluch et al. 2013, 2015b; Zhang et al. 2015). This approach has been recently demonstrated for recording motions in aromatic groups in GB1 and dynein light-chain LC8 proteins (Paluch et al. 2015a).

CN (Chan & Tycko, 2003) and RN (Hou et al. 2010, 2012, 2013b) symmetry sequences for measurement of chemical shift tensors have also been established (dubbed ROCSA and RNCSA, respectively). Like their dipolar counterparts, these experiments can be used under fast MAS and in fully protonated, uniformly $^{13}$C enriched systems, with effective suppression of homonuclear dipolar couplings. Several variations of these RN symmetry pulse sequences for the study of...
microsecond-to-nanosecond dynamics are presented in Fig. 7. RN-DIPSHIFT, PARS, and RNCSA experiments have been successfully applied to a range of supramolecular assemblies, with several studies highlighted below.

2.3.2 Millisecond to microsecond timescale dynamics

Biological processes on the millisecond-to-microsecond timescale include domain motions and enzyme catalysis. Rotating frame ($R_{1p}$) and transverse ($R_{2*}$) spin-lattice relaxation and $^{15}$N–$^{13}$C dipolar couplings are sensitive to dynamics on this timescale. Quantitative spin-lattice relaxation methods can measure exchange rates, population distributions, and chemical shift differences among exchanging sites. An important consideration for relaxation-based dynamics studies is interference from relaxation mechanisms unrelated to dynamics. These issues can be overcome by the use of deuteration and/or fast MAS (Lewandowski et al. 2011b; Quinn & McDermott, 2012; Tollinger et al. 2012).

Rotating frame-based experiments measure relaxation resulting from spatial reorientation of a CSA or dipolar tensor and are sensitive to microsecond dynamics. Lewandowski et al. quantified site-specific backbone dynamics of GB1 with $^{15}$N and $^{13}$C $R_{1p}$ relaxation measurements (Fig. 8a,b, (Lamley et al. 2015; Lewandowski et al. 2011b)). The method has been recently applied to the membrane protein ASR by Ladizhansky and co-workers (Good et al. 2014). Krushelnitsky and co-workers presented a suite of relaxation studies on SH3 over a range of timescales that included $R_{1p}$ relaxation, as well as $^1$H–$^{15}$N dipolar couplings and $R_1$ relaxation (Zinkevich et al. 2013). As technology and methodology in the field of MAS NMR continue to advance, these experiments are being applied to increasingly complex systems.

Carr–Purcell–Meiboom–Gill (CPMG) measurements, sensitive to sub-millisecond motions, have been well established for the study of protein dynamics in solution state NMR (Epstein et al. 1995; Farrow et al. 1994; Lorieau et al. 2010; Mandel et al. 1995; Meiboom & Gill, 1958; Volkman et al. 2001; Zhang et al. 2006). These experiments have recently been extended to applications in the solid state, as demonstrated by Schanda and co-workers on microcrystalline ubiquitin (Tollinger et al. 2012). In this study, differential line broadening of zero- and double-quantum coherences as an initial indicator of dynamics was also utilized (Dittmer & Bodenhausen, 2004). Dipolar evolution time from REDOR dephasing curves can also serve as a measure of millisecond timescale dynamics. In this case, reduced effective $^{15}$N–$^{13}$C dipolar couplings at dynamic sites result in the absence of REDOR dephasing, as observed for the important linker residue Tyr145 in HIV-1 capsid assemblies (Byeon et al. 2012). Reduced $^{15}$N–$^{13}$C dipolar couplings and hence the presence of millisecond timescale dynamics can also be reflected in reduced N–C cross-peak intensities (Helmus et al. 2008).

Fig. 7. RN-symmetry based sequences for the measurement of dipolar and chemical shift anisotropy lineshapes. (Hou et al. 2014) (a) conventional RN-based DIPSHIFT, (b) $^1$H CSA recoupling with or without heteronuclear decoupling, (c) PARS, (d) constant time PARS, (e) 3D PARS for dipolar lineshape measurements. Reproduced with permission from Hou et al. (2014). Copyright 2014 AIP Publishing.
Exchange experiments such as center-band only detection of exchange (CODEX) can monitor slow dynamics on the timescale of milliseconds-to-seconds. In a CODEX experiment, dipolar (Krushelnitsky et al. 2009; Li & McDermott, 2009) or CSA (DeAzevedo et al. 1999) recoupling is applied before and after a mixing period. Dynamic residues will undergo only partial refocusing after the mixing period and exhibit reduced peak intensities. CODEX has been applied to characterize site-specific backbone and side-chain dynamics of α-spectrin SH3 (Fig. 8c–e, (Krushelnitsky et al. 2009)), as well as small organic molecules (Li & McDermott, 2012).

2.4 Intermolecular interactions

Protein–protein and protein–ligand interactions are integral in the regulation of cellular processes. Understanding these interactions can be very advantageous for guided design of therapeutics. Protein–protein interactions are a hallmark of biological assemblies. MAS NMR is a powerful method for the study of intermolecular interactions as many protein–protein and protein–ligand complexes, including membrane proteins, amyloids, and viruses exist in non-crystalline or insoluble environments.

2.4.1 Chemical shift perturbations

Analysis of chemical shift perturbations, including intensity changes and line broadening, between free and bound states of a system of interest can indicate sites affected by the presence of a binding partner as demonstrated in early work of ligand binding to Bcl-xL with MAS NMR from Zech and McDermott (Zech et al. 2004). Chemical shift perturbations indicate not only direct protein–protein (Liu et al. 2016) or protein–ligand interactions (Schutz et al. 2011), but also allosteric structural changes that occur as a result of binding. An extension of chemical shift perturbations, changes in the CSA tensor...
magnitude can also serve as a probe of intermolecular interactions, such as applied to studies of cytb(S) in complex with cytP4502B4 (Pandey et al. 2013). Further, MAS NMR methods exist to characterize the bona fide intermolecular binding interface. For example, PRE in sites distal to the spin label can also be used as an indicator of intermolecular interactions (Wang et al. 2012). Dipolar edited correlation techniques have been particularly productive for the study of direct intermolecular interactions in biological macromolecular assemblies.

2.4.2 Dipolar-edited correlation spectroscopy

One approach to identifying intermolecular interfaces is the use of differential isotopic labeling of the binding partners, where one binding partner is $^{13}$C or $^{13}$C, $^{15}$N labeled and the other is $^{15}$N labeled, first demonstrated by Polenova (Marulanda et al. 2004; Yang et al. 2008) and Baldus (Etzkorn et al. 2004; Weingarth & Baldus, 2013). Long range polarization transfer across the intermolecular interface is achieved using heteronuclear mixing sequences such as REDOR (Gullion & Schaefer, 1989) or TEDOR (Hing et al. 1992), NHHC (Lange et al. 2002), PAIN (Lewandowski et al. 2007), or a combination of these methods such as REDOR–PAINCP (Fig. 9, panel 2 (Yang et al. 2008)). These techniques have been applied for the observation of intermonomer interactions in microcrystalline Crh (Etzkorn et al. 2004) and structural characterization of α-synuclein amyloid fibrils (Lv et al. 2012) with NHHC correlation experiments, PAIN-derived intermolecular correlations in thiodenodoxin reassemblies (Yang et al. 2008), and quaternary structure of GB1 crystals with TEDOR (Nieuwkoop & Rienstra, 2010). An alternate approach is to use complementary, selective $^{13}$C labeling such as a mixture of [1-$^{13}$C]glucose and [2-$^{13}$C]glucose labeled protein to observe intermolecular $^{13}$C–$^{13}$C correlations. This scheme has been applied to α-synuclein fibrils (Loquet et al. 2010).

A considerable drawback to the approach described above is the requirement that both species are isotopically labeled. Many systems of interest cannot be readily prepared with isotopic labels. Thus methods to observe binding interfaces where one binding partner is natural abundance are essential. These interface correlations can be achieved in REDOR-filter-based experiments. In these experiments, REDOR dephasing is applied to $^1$H–$^{13}$C and/or $^1$H–$^{15}$N dipolar couplings to eliminate magnetization arising from directly bonded protons. Subsequent $^1$H polarization transfer arises from the source without $^{13}$C/$^{15}$N labels and magnetization is transferred to the magnetically active nuclei of the isotopically labeled protein at the binding interface. This approach has been demonstrated for the study of natural abundance rhodopsin in complex with $^{13}$C-labeled 11-cis-retinal (Kühne et al. 2005) and assemblies of U–$^{13}$C, $^{15}$N-CAP-Gly with natural abundance polymerized microtubules (MTs) assemblies (Yan et al. 2015a). A significant drawback of this approach in the latter study is the presence of residual deuteration at ca. 1%, resulting in unwanted intramolecular transfers. To overcome this limitation, double-REDOR filter approach (dREDOR) was demonstrated on assemblies of U–$^{13}$C, $^{15}$N-CAP-Gly with natural abundance polymerized MTs (Yan et al. 2015a). In this method, a double $^1$H–$^{13}$C/$^1$H–$^{15}$N REDOR filter is applied to dephase the $^1$H signals from the U–$^{13}$C, $^{15}$N-CAP-Gly, with the subsequent transfer of $^1$H magnetization from the MTs back to U–$^{13}$C, $^{15}$N-CAP-Gly across the intermolecular interface.

In addition to the observation of protein–protein or protein–ligand interactions, REDOR-filter-based experiments, such as MELODI-HETCOR (Fig. 9, panel 1 (Yao et al. 2001)) have also been applied to the study of water–protein interactions, including the identification of a hydrated arginine for an antimicrobial peptide in a lipid bilayer (Li et al. 2010) and the characterization of Pf1 bacteriophage hydration (Purusottam et al. 2013; Sergeyev et al. 2014).

In a variation of these dipolar edited methods, Asami et al. identified residues of the protein L7Ae that form the binding interface with box C/D RNA in the protein-RNA complex using $^1$H, $^{15}$N-labeled protein and $^1$H, $^{13}$C, $^{15}$N- or $^2$H, $^{13}$C, $^{15}$N-labeled RNA. In this experiment, protein residues interacting with $^1$H-RNA exhibited line broadening and weaker peak intensities relative to the corresponding peaks in the $^2$H-RNA complex due to $^1$H–dipolar interactions (Asami et al. 2013).

Beyond qualitative identification of intermolecular interfaces, the distance dependence of dipolar couplings ($1/r^4$) can be used to quantify protein–protein and protein–ligand intermolecular distances. In addition to applications in structure calculation and dynamics studies, REDOR dephasing or TEDOR build-up curves can be used to quantify intermolecular distances. Protein–protein intermolecular distances can lend insight into higher-order structural features, as has been demonstrated for amyloid fibrils by Tycko, Dobson, and others (Fitzpatrick et al. 2013; Petkova et al. 2006; Van der Wel et al. 2010) (Fig. 10) and membrane protein architecture by Griffin and co-workers (Andreas et al. 2015). REDOR distance measurements are also used for the studies of ligand–protein interactions, including binding of amantadine and its derivatives in the M2 channel by Cross (Wright et al. 2016) and Hong (Cady et al. 2010) and retinal binding in bacteriorhodopsin (Helmle et al. 2000). REDOR-derived distances can also indicate changes in protein or nucleic acid conformation upon ligand binding, demonstrated for tat peptide-bound TAR RNA of HIV-1 (Olsen et al. 2005).
Various classes of supramolecular assemblies studied by MAS NMR include amyloid systems (reviewed in (Comellas & Rienstra, 2013; Tycko, 2011)), the Shigella type-III secretion system (TSS3) (Demers et al. 2014; Loquet et al. 2012, 2013a), the Escherichia coli pilus protein FimA (Habenstein et al. 2015), and the MAVS (mitochondrial antiviral signaling) protein (He et al. 2015, 2016). This review focuses on two particular classes of supramolecular assemblies: cytoskeletal proteins and viruses.

3. MAS NMR of cytoskeleton-associated proteins

The cytoskeleton is an essential cellular component in all domains of life. Functions of the cytoskeleton include maintenance of cell shape, motility, intracellular transport, endocytosis, and cell signaling (Fischer & Fowler, 2015). In eukaryotes, the cellular cytoskeleton is composed of three main filament types: microfilaments (actin filaments), MTs (Nogales, 2000), and intermediate filaments. While most filaments of the prokaryotic cytoskeleton have eukaryotic analogues, there are filament types that are unique to prokaryotes (Lowe et al. 2004). Function of the cytoskeleton is crucially dependent on interactions with...
binding partners, including the motor proteins dynein, kinesin, and myosin (Vale, 2003). The disruption of these interactions by small molecules is a key mechanism in therapeutics for the treatment of cancers (Wood & Bergnes, 2004) and neurodegenerative diseases (Gunawardena, 2013).

3.1 MTs and MT-associated proteins (MAPs)

MTs and MAPs perform several vital physiological functions in the cell including mitosis and transport of signaling molecules (Vale, 2003). MTs are an important target of chemotherapeutics due to their essential roles in cell division. MTs are highly dynamic and continually polymerizing and de-polymerizing in the cellular matrix (Howard & Hyman, 2003). Despite extensive structural and biochemical characterization, many open questions remain with respect to the function of MTs and their associated proteins, including the atomic-level understanding of protein–protein interactions, of the role of protein dynamics in different conformational states, and of how protein–protein interactions and dynamics come together to orchestrate cellular processes. MAS NMR can lend insight into the structure and dynamics of MT–MAP complexes at atomic resolution. To date, in-depth MAS NMR studies have been performed on only a handful of systems, including dynactin’s CAP-Gly domain assembled with MTs (Fig. 11a,b), bactofilin (Fig. 11c), and MTs interacting with small molecules, such as paclitaxel (Taxol) (Li et al. 2000; Paik et al. 2007), epothilone B (Kumar et al. 2010), and their derivatives. In the following, we review the work on the first two cytoskeletal assemblies.

3.1.1 Structure of CAP-Gly domain of dynactin

Dynactin, an activator of the dynein motor assembly, is a protein complex involved in intracellular transport (Caviston & Holzbaur, 2006). Dynactin regulates dynein transport along MTs, and mutations within its p150Glu subunit lead to neurodegenerative disorders, such as Huntington’s disease, Charcot–Marie Tooth disease, amyotrophic lateral sclerosis, distal spinal bulbar muscular atrophy, and Perry syndrome (Chen et al. 2014). Within the p150Glu subunit of dynactin, CAP-Gly (cytoskeleton-associated protein glycine-rich) is an 89 residue MT-binding domain (Vaughan et al. 2002; Waterman-Storer et al. 1995). Dynactin CAP-Gly is the first MAP assembled with MTs, whose structure and dynamics have been investigated in depth by MAS NMR, yielding atomic-resolution insights unavailable from other techniques and establishing a proof of principle for investigations of other cytoskeleton-associated assemblies (Ahmed et al. 2010; Sun et al. 2009; Yan et al. 2013a, 2015a, b). Recently, the atomic-level resolution structure of CAP-Gly bound to polymeric MTs was reported (Yan et al. 2015a) (PDB ID code 2MPX), the first structure of any cytoskeleton-associated protein assembled with cytoskeleton.

To determine the structure of CAP-Gly in complex with MTs, three different isotopic labeling schemes were used: U-15N,13C; U-15N, [2-13C]glucose; and U-15N, [1,6-13C]glucose. The structure was determined using hundreds of medium- and long-range
distance restraints collected from $^{13}$C–$^{13}$C CORD and $^{15}$N–$^{13}$C PAIN-CP experiments, as well as hydrogen-bonding restraints and torsion angles from TALOS+. The equivalent resolution in the structure was 1.9–2.5 Å with a tightly constrained ensemble of lowest-energy conformers. A very similar approach was previously used to determine the structure of free CAP-Gly (PDB ID code 2M02) (Yan et al. 2013a). The structure of CAP-Gly assembled on MTs indicates that, while the overall secondary structure is retained, the flexible loops of CAP-Gly have remarkably different conformations when associated with MTs (Fig. 12). Loop $\beta_3/\beta_4$ adopts a more open conformation in the free state of CAP-Gly, and rearranges to a more closed conformation when bound to MTs. The different sidechain orientations of residues in this loop may play a role in CAP-Gly’s structural plasticity and ability to interact with different binding partners.

EB1 is another MAP that, like dynactin, localizes at the plus end of the growing MT. EB1 is thought to have a role in MT dynamics; specifically it may promote MT elongation (Rogers et al. 2002). EB1 interacts with the p150Glued subunit of dynactin and it is hypothesized that the two proteins form a plus end complex to regulate MT dynamics (Ligon et al. 2003). The p150Glued subunit may play a role in recruiting EB1 to the MTs. Residues of CAP-Gly that are perturbed by binding of EB1 were identified by chemical shift changes. Chemical shift perturbations indicate that free CAP-Gly exists in multiple conformers, but is conformationally homogeneous when bound to MTs (Yan et al. 2013a).

### 3.1.2 Interface of CAP-Gly with MTs

The main challenge for NMR characterization of MTs and their assemblies with associated proteins is that to date there have been no efficient isotopic labeling protocols established for tubulin. This precludes in-depth structural characterization of MTs and limits the approaches for determination of intermolecular interfaces formed by MTs and their binding partners. To overcome this challenge, the application of dREDOR filters was explored to characterize the intermolecular interfaces that dynactin’s CAP-Gly forms with MTs and EB1 (Yan et al. 2015a). In these experiments, $^1$H–$^{13}$C and $^1$H–$^{15}$N REDOR filters were simultaneously applied to dephase all $^1$H magnetization arising from U-$^{13}$C, $^{15}$N CAP-Gly. Subsequently, polarization was transferred from $^1$H of natural abundance MTs or EB1 to their binding interface on the surface of CAP-Gly. A $^{13}$C–$^{13}$C CORD dimension was included to enable site-specific assignment of the binding interface. Figure 13 shows dREDOR–CORD and dREDOR–HETCOR spectra of CAP-Gly in complex with MTs (b), EB1 (c), and the intermolecular interfaces as determined by dREDOR (a left, green) and chemical shift perturbations (a right, orange/purple). The results confirmed that loop $\beta_3/\beta_4$ including the GKNDD motif comprises the primary binding interface with MTs. CAP-Gly interacts with its binding partners on the flat side of the protein, where most of the surface-exposed hydrophobic residues are located. dREDOR experiments of the CAP-Gly/EB1 complex were consistent with the known binding interface for this complex, which has been determined previously (Hayashi et al. 2005; Honnappa et al. 2006; Yan et al. 2013a), validating the approach for characterizing the CAP-Gly/MT interface.

### 3.1.3 Dynamics of CAP-Gly

MT-associated motors and their activators possess conformational plasticity, which is essential for their ability to bind to and slide along the MTs (Howard, 2001; Vale, 2003). Conformational plasticity is directly related to internal flexibility, and
therefore, knowledge of dynamics is essential for understanding the biological function of MAP assemblies. The dynamics of CAP-Gly free, in complex with EB1, and bound to polymeric MTs have been characterized using MAS NMR, over a range of functionally relevant timescales from nanoseconds to milliseconds (Yan et al. 2015b). Global dynamics were probed through a temperature series of 1D- and 2D13C spectra (Fig. 14a–c). Site-specific dynamics were characterized using 1H–15N and 1H–13C DOPs (Fig. 14d) (Hou et al. 2011b). As indicated by the temperature series in Fig. 14a–c, free and MT-bound CAP-Gly are dynamic across the entire range of timescales under investigation, and the motions are strongly temperature dependent. In contrast, the CAP-Gly in complex with EB1 is largely rigid on these timescales and its spectra are not temperature dependent. From the measurement of DOPs, it was found that the loops of CAP-Gly are mobile in the free protein as well as in complex with MTs. However, consistent with 1D and 2D spectra, the dynamics of CAP-Gly are notably attenuated when in complex with EB1. Remarkably, the loops of CAP-Gly show an increase in fast timescale backbone fluctuations (nanosecond-to-microsecond), but a decrease in slower dynamics (microsecond-to-millisecond) upon binding to MTs (Fig. 14d). It was proposed that these observed changes in dynamics have a critical function in CAP-Gly/MT interactions. The combined structural and dynamics studies of CAP-Gly highlight the structural plasticity of this protein and the essential role this flexibility plays in CAP-Gly’s ability to adopt different conformations and interact with different binding partners.

3.2 Bactofilins

Bactofilins are a recently discovered class of bacterial cytoskeletal proteins. Similar to eukaryotic cytoskeletal proteins, these proteins have diverse functions, such as cellular mobility, cell shape, and attachment. Bactofilins assemble rapidly and spontaneously, making them not amenable for characterization by many biophysical techniques (Kuhn et al. 2010). Bactofilins contain a conserved central rigid DUF583 domain, and terminal regions that are more dynamic. Lange and co-workers have recently reported resonance assignments and structure for BacA from Caulobacter crescentus (Shi et al. 2015; Vasa et al. 2015). Only the core DUF583 domain (residues 39–137) was observed in dipolar based 13C–13C and 15N–13C spectra, supporting the hypothesis that while the core is rigid, the termini are dynamic and believed to have a role in protein–protein interactions.
or protein–membrane interactions. Secondary structure analysis revealed at least ten distinct β-sheet segments (Fig. 15). To observe dynamic residues, through-bond INEPT (Morris & Freeman, 1979) $^{13}$C–$^{13}$C correlation spectra were acquired. Chemical shifts for residues in the INEPT spectra indicated random coil secondary structure. Interestingly, fewer resonances were observed in the INEPT spectrum than would be expected to arise from the N- and C-termini, indicating that not all residues in these regions are dynamic. While the secondary structure and dynamic behavior of BacA have similar features to amyloids (Daebel et al. 2012; Heise et al. 2005; Helmus et al. 2010), BacA has a distinct β-helical tertiary structure, as indicated by mass-per-length measurements by scanning transmission EM.

The atomic resolution structure determined by Shi et al. revealed that BacA is a right-handed β-helix with a triangular hydrophobic core and six windings (Shi et al. 2015). With sparsely labeled BacA samples ($^{1,3}$-13C and $^{2}$-13C glycerol), medium- and long-range distance restraints were obtained, as well as torsion angle restraints from TALOS+ and β-sheet hydrogen bond restraints. Additional $^{1}$H–$^{1}$H distance restraints from a 4D HN(H)(H)NH spectrum (acquired with sine weighted Poisson-gap non-uniform sampling (Hyberts et al. 2010, 2012)) and a 2D NHHC spectrum were essential for determination of the handedness of the β-helical structure. The presence of a right-handed β-helix had not been previously reported for any cytoskeletal protein. The hydrophobic core is triangular with highly conserved glycines at many of the corners and three parallel β-sheets per winding (Fig. 16). It is believed that hydrophobic interactions mediate polymerization/folding of bactofilins (Kuhn et al. 2010). Hydrogen bonds between adjacent β-strands help to stabilize the overall structure. Windings 1 and 6 were not as well restrained due to a lack of intermolecular distance restraints, which is attributed to increased dynamics in these regions of the protein. Mutations of hydrophobic residues in winding 6 affect BacA assembly in vivo (Kuhn et al. 2010). It is likely that dynamics in this region of the protein has a role in BacA assembly.

4. MAS NMR of viral assemblies and intact viral particles

Viruses are small pathogens that can impact all forms of life from bacteria, to plants and animals (Brussaard et al. 2004; Nelson & Citovsky, 2005; Pearson et al. 2009; Prangishvili, 2013; Smith & Helenius, 2004). The general virion structure is a single- or double-stranded DNA or RNA encapsulated by a protein coat or capsid. Some viruses include a lipid envelope surrounding the capsid as well. Viruses invade target cells and seize the host cell’s machinery to reproduce, while evading cellular defense mechanisms (Kaminskyy & Zhivotovsky, 2010). These key aspects of viral replication are of great interest as targets for the treatment of viral infections, but can also be exploited for nanotechnology and drug development. MAS NMR is an excellent tool to probe structure and dynamics of viral protein assemblies, and particularly promising is an integrated approach, where this technique is combined with other experimental (e.g., cryo-EM, X-ray crystallography, and solution NMR) and/or computational (e.g., MD simulations) methods to yield atomic-level understanding of structure and
dynamics of viral assemblies. Below we discuss two important classes of viruses, HIV-1 and bacteriophages, which to date have been the most thoroughly characterized by MAS NMR (Fig. 17).

4.1 HIV-1 capsid and maturation intermediates

Acquired immunodeficiency syndrome (AIDS), caused by HIV, is a global pandemic and affects approximately 37 million people globally (World Health Organization, 2015). A key step in the HIV lifecycle is maturation, where an immature viral particle is converted into a mature, infectious virion through a cascade of cleavage steps of the Gag polyprotein (Fig. 18a, Fig. 22, panel 1). The final step of maturation is the cleavage of the SP1 tail from CA (capsid protein) and the reorganization of CA into the mature conical capsid, which encapsulates the retroviral RNA (Briggs & Krausslich, 2011; Engelman & Cherepanov, 2012). Viral maturation is of poignant interest as the target of a novel class of therapeutics termed maturation inhibitors (Adamson et al. 2009; Salzwedel et al. 2007). Maturation inhibitors such as Bevirimat (BVM) inhibit HIV-1 maturation by binding to the CA-SP1 cleavage site and preventing cleavage. Tubular assemblies of CA mimic the native capsid lattice (Byeon et al. 2009; Zhao et al. 2013). Obtaining stable, morphologically homogeneous samples for MAS NMR typically requires high ionic strength (∼1–2 M NaCl). With the use of low-E and Efree probes designed to minimize heating due to
radiofrequency irradiation, outstanding quality MAS NMR spectra can be obtained (Byeon et al. 2012; Bayro et al. 2014; Chen & Tycko, 2010; Han et al. 2010, 2013; Lu et al. 2015a).

### 4.1.1 Structural characterization of HIV-1 capsid assemblies

CA is a 231 residue, predominantly α-helical protein with independently folding N-terminal (NTD, residues 1–143) and C-terminal (CTD, residues 148–231) domains, connected by an inter-domain linker (Deshmukh et al. 2013; Gres et al. 2015). Early studies (Chen & Tycko, 2010; Han et al. 2010) revealed that most residues in both NTD and CTD are relatively rigid, and CA assemblies retain the secondary and tertiary structure determined from solution NMR and x-ray crystallography studies. Upon maturation, the viral capsid morphology changes from spherical to conical. Mature conical HIV-1 capsids are built from hexameric and pentameric CA assemblies, whose stoichiometry in the final capsid is variable, and so is the capsid’s cone shape. Four kinds of intermolecular contacts are critical for capsid morphology and stability both in vivo and in vitro: intra-hexameric NTD–NTD and NTD–CTD, and inter-hexameric CTD–CTD interfaces around pseudo twofold and pseudo threefold axes (Fig. 18c–e) (Byeon et al. 2009; Pornillos et al. 2009). The capsid’s structural polymorphism is also observed in vitro, where the morphology can be controlled by assembly conditions. CA can assemble into cones, tubes, and spheres (Barklis et al. 1998; Ehrlich et al. 2001; Ganser-Pornillos et al. 2007; Gross et al. 1997; Han et al. 2010, 2013). MAS NMR studies of spheres, cones, and tubes indicated no major differences in secondary or tertiary structure among the three morphologies (Bayro et al. 2014; Chen & Tycko, 2010; Han et al. 2010, 2013), but no 3D structure is available so atomic-level details of structural organization of these in vitro assemblies remain poorly understood.

HIV-1 exhibits significant sequence variability and sequence-dependent viral infectivity (Price et al. 2012). To date, two HIV-1 strains have been studied by MAS NMR: HXB2 and NL4-3 (Byeon et al. 2012; Bayro et al. 2014; Chen & Tycko, 2010; Han et al. 2010, 2013; Lu et al. 2015a). The wild-type sequences of these proteins differ by only four amino acids, but this variation causes considerable conformational changes across the CA sequence (Han et al. 2013). These structural perturbations may play a role in the differing viral infectivity of the two strains. Viral infectivity is also regulated through the interactions with host factors. Cyclophilin A (CypA) is one such host factor protein that modulates HIV-1 infectivity and uncoating through direct interactions with the HIV-1 capsid’s CypA-binding loop located in the NTD. The mechanism of CypA is complex and poorly understood (Colgan et al. 1996; Luban et al. 1993). A recent study yielded the structure of CypA in complex with the assembled HIV-1 capsid at 8 Å resolution (Liu et al. 2016). It was discovered, using cryo-EM-guided all-atom MD simulations and MAS NMR spectroscopy that CypA simultaneously interacts with two CA subunits in different hexamers. This binding mechanism established through the integrated cryo-EM/MAS NMR/MD

---

**Fig. 16.** Top view (a) and side view (b) of BacA structure determined by MAS NMR. (c) Schematic representation of the 6 windings. Colors are as follows: white – hydrophobic residues, red – acidic residues, blue – basic residues, and green – other residues. Mutations of asterisked residues in winding 6 affect in vivo assembly. Reprinted with permission from Shi et al. (2015). Copyright 2015 American Association for the Advancement of Science.
approach provided insights into the mechanism of HIV-1 capsid stabilization by and recruitment of CypA to promote the viral infectivity.

Additional mutations in the various regions of CA sequence have been shown to affect the assembly morphology and efficiency, capsid stability, and interactions with host factors (Bocanegra et al. 2012; Forshey et al. 2002; Gres et al. 2015; Jiang et al. 2011; Manocheewa et al. 2013; McCarthy et al. 2013; Qi et al. 2008; Yang & Aiken, 2007). For example, mutations in the linker region of CA render the capsid assembly inefficient or abolish it completely (Jiang et al. 2011). On the contrary, the E45A mutation makes a ‘hyperstable’ capsid, which does not disassemble, thus dramatically reducing the viral infectivity (Forshey et al. 2002; Hulme et al. 2015; Yang et al. 2012). Mutations in the CypA-binding loop, such as A92E and G94D exhibit drastically reduced viral infectivity in the presence of the normally required CypA. Full infectivity can be restored by inhibition of CypA by cyclosporin. MAS NMR analysis of these mutants revealed that the mutations induce conformational changes, as well as changes in the dynamics, vide infra (Lu et al. 2015a).

To get insights into the structural changes occurring upon formation of the hexagonal CA lattice (predominant symmetry in the cones), Bayro et al. used MAS NMR-derived structural restraints in tubular CA assemblies for refinement against solution NMR and x-ray crystallography structures of monomeric CA (Bayro et al. 2014). 15N–13C backbone distances, determined by 15N–13C dipolar recoupling (15N-BARE), have been used by Tycko and co-workers to derive backbone torsion angles and secondary structure (Hu et al. 2012). In α-helices and tight turns, 15N–13C distances are short, leading to rapid signal decay, while the decay is slower in loop regions (Fig. 19a). Deviations of experimental 15N-BARE curves of CA tubular assemblies from simulated curves derived from crystal or solution NMR structures indicated differences in secondary structure. Using the 15N–13C distances and TALOS torsion angle predictions as structural restraints, structure refinement was carried out against CA coordinates derived from solution NMR (monomer, 2LF4 (Shin et al. 2011)) and crystal structures (hexamer, 3MGE (Pornillos et al. 2010)). Figure 19b–d shows regions of the CA structure that adopt a more extended conformation in tubular assemblies than in their respective initial structures. Among these regions that undergo changes are the 310 helix near the start of the CTD, as well as the loops between helices 3 and 4, and helices 10 and 11. These perturbations were attributed to conformational changes arising from higher-order oligomerization and/or lattice curvature.

Very recently Bayro and Tycko characterized the structure of tubular assemblies of the capsid pseudo-twowfold inter-hexameric interface (Bayro & Tycko, 2016) using mixed and selective labeling schemes and dipolar-based distance measurements, with additional restraints derived from the cryo-EM structure determined by Zhang and co-workers (Zhao et al. 2013). In this study, quantitative W184 to M185 intermolecular distances across the dimerization interface were obtained with 15N–13C REDOR dephasing experiments. A series of experiments were performed to characterize intra-residue distances of W184 and M185 in order to constrain the side-chain conformations, including 15N–13C REDOR distances, 13C–13C BroBaRR (BroadBand Rotational Resonance) distances for M185 (Chan & Tycko, 2004), and tensor correlation experiments (Dabbagh et al. 1994) of W184 to determine the angle between the backbone Ca–N bond vector and sidechain αC–Ne1 bond vector. The distances and angles determined differed somewhat from prior solution NMR (Byeon et al. 2009) and X-ray crystallography studies (Gres et al. 2015; Worthylake et al. 1999), suggesting the structure of the dimerization interface in tubular assemblies differs from that in crystals or CA in solution. Further intra-CA monomer distance constraints were obtained from 13C–13C correlations with long mixing (700 ms). With an integrated structure calculation approach described above, they were able to determine the structure of the inter-hexameric dimer interface in CA tubular assemblies, and further demonstrated that this interface is well ordered in this morphology.
Protein dynamics have indispensable functions in many stages of the HIV-1 lifecycle, including assembly, release, and maturation (Freed, 2015). While CA is relatively rigid in the individual α-helices, residues in loops and the flexible linker, which are key functional regions of the protein, are dynamic, giving rise to conformational plasticity, which is directly connected to capsid morphology, interactions with host factors, and viral infectivity.

Lu et al. have characterized the site-specific dynamics of several capsid constructs, including the HXB2 and NL4-3 strains discussed above, as well as CypA-bound CA and the A92E and G94D ‘escape mutants’ (Lu et al. 2015a). These mutants have approximately 10% the infectivity of the wild-type virus; however, infectivity can be restored upon CypA inhibition (Yilnen et al. 2009). The CypA loop of CA (residues 84–100) undergoes conformational changes upon binding of CypA as well as in the A92E and G94D ‘escape’ mutants. 

4.1.2 Conformational dynamics of HIV-1 capsid assemblies by MAS NMR

Protein dynamics have indispensable functions in many stages of the HIV-1 lifecycle, including assembly, release, and maturation (Freed, 2015). While CA is relatively rigid in the individual α-helices, residues in loops and the flexible linker, which are key functional regions of the protein, are dynamic, giving rise to conformational plasticity, which is directly connected to capsid morphology, interactions with host factors, and viral infectivity.

Lu et al. have characterized the site-specific dynamics of several capsid constructs, including the HXB2 and NL4-3 strains discussed above, as well as CypA-bound CA and the A92E and G94D ‘escape mutants’ (Lu et al. 2015a). These mutants have approximately 10% the infectivity of the wild-type virus; however, infectivity can be restored upon CypA inhibition (Yilnen et al. 2009). The CypA loop of CA (residues 84–100) undergoes conformational changes upon binding of CypA as well as in the A92E and G94D ‘escape’ mutants. 

1H–15N and 1H–13C dipolar tensors and resonance intensities reported

Fig. 18. (a) Late stages of HIV-1 viral life cycle from assembly at the host cell membrane to budding and virion maturation. (Freed, 2015) (b) All-atom MD derived model of the HIV-1 capsid based on cryo-ET and solution NMR, with a hexamer of hexamers subunit shown in the expansion (Lu et al. 2015a). (c) Hexmer of hexamers of HIV-1 capsid assembly. The interhexameric trimer interface is circled. (d, e) Helix 10 trimer interface and helix 9 dimer interface respectively. Blue: hexamer of hexamers, orange: pentamer of hexamers (Zhao et al. 2013). (f) CA monomer. Residues for which chemical shift perturbations are observed upon binding of CypA are highlighted orange. (g) 15N–13C and 13C–13C correlation spectra of free CA tubular assemblies (black) and CA tubular assemblies in complex with CypA (Lu et al. 2015a) (a) Adapted with permission from Freed (2015). Copyright 2015 Macmillan Publishers. (b, f, g) Adapted with permission from Lu et al. (2015a). Copyright 2015 National Academy of Sciences. (c–e) Adapted with permission from Zhao et al. (2013). Copyright 2013 Nature Publishing Group.
Fig. 19. (a) $^{15}$N-$^{15}$N BARE curves for selected CA residues. Circles indicate experimental curves for CA tubular assemblies. Simulated curves correspond to the following structures: (- -) 2LF4, (−) 3MGE, (−−) 2KOD. (b) Initial (red) and final (blue) structure refinement against 2LF4, indicating the change in the conformation of the $\alpha_3$ helix. (c, d) Initial (orange) and final (blue) structure refinement against 3MGE, indicating the change in conformation of loop 3/4 and loop 10/11, respectively. Adapted with permission from Bayro et al. (2014). Copyright 2014 Elsevier.
on CA dynamics on the nano- through millisecond timescales. The CypA loop was observed to exhibit an unprecedented degree of mobility over the entire range of timescales under study. As expected, the motions are attenuated upon CypA binding. A completely unexpected finding was that the dynamic profiles of the escape mutants A92E and G94D closely resembled that of CA in complex with CypA: mobility of CypA loop is dramatically attenuated due to the mutations. Furthermore, NMR parameters derived from all-atom MD trajectories were in remarkable agreement with experimental results, corroborating that the dynamics are modulated by the mutations in the CypA loop. Figure 20 demonstrates (a) DOPs and (b) $^1$H–$^{15}$N dipolar lineshapes of residues in the CypA loop with (c) peak intensities and (d) DOPs mapped to the structure of CA. Taken together, these results indicate that the capsid escapes the CypA dependence by a dynamic allostery mechanism and highlight the key role of conformational dynamics in the HIV-1 CA function.

Conformational plasticity of HIV-1 CA is essential for its assembly into the viral capsid core. As discussed above, the hinge that links the NTD and CTD has been shown to play a key role in HIV-1 core stability (Byeon et al. 2009; Jiang et al. 2011). The dynamic origin and mechanism of CA’s pleiomorphic assembly was investigated by an integrated MAS and solution NMR approach, using conical assemblies of U–$^{13}$C, $^{15}$N–Tyr-labeled samples of full length CA, of individual CTD constructs, and solutions of U–$^{13}$C, $^{15}$N-CA and CTD (Byeon et al. 2012). Tyrosine is an ideal probe of linker dynamics as functionally important Y145 is located in the linker; CA has only three additional Tyr residues located in rigid regions of the protein, which provide internal controls in the NMR characterization. To observe dynamics on the millisecond timescale, $^{13}$C–$^{15}$N REDOR dephasing experiments were performed. The dephasing profiles for the three Tyr residues located in the helical regions of CA are consistent with them being rigid. On the contrary, the Y145 signals could not be dephased due to the complete averaging of the $^{13}$C–$^{15}$N dipolar couplings, which is a clear manifestation of dynamics occurring on the timescale of the order of $\sim$10 ms (Gullion & Schaefer, 1989). To probe the dynamics of the interdomain linker on the microsecond to nanosecond timescale, $^1$H–$^{15}$N dipolar and $^{15}$N CSA lineshapes were acquired. All Tyr residues, including Y145 showed close to rigid limit dipolar and CSA parameters, establishing that Y145 is rigid on the corresponding timescales. On the basis of these results it was concluded that millisecond timescale conformational dynamics of the interdomain linker is essential for the CA function.
assembly. Solution NMR experiments revealed that, while these millisecond timescale motions of the linker are essential for opening up the conformational space, the number of accessible conformers is finite, and their populations are controlled through electrostatic intermolecular interactions between side chains of W184 and E175. This integrated approach established a molecular switch mechanism by which dynamics and electrostatic interactions work together to permit the formation of varied capsid morphologies in the mature HIV (Byeon et al. 2012).

In another study, Tycko and co-workers observed that while CA is generally rigid, there are regions of crucial biological function that exhibit static or dynamic disorder (Bayro et al. 2014). Using scalar-based $^{13}$C–$^{13}$C TOBSY (Hardy et al. 2001) experiments, dynamically disordered regions of the protein were identified. Residues attributed to the interdomain linker, loop 8/9, and the C-terminal tails were observed in this data set, indicating that these residues are dynamically disordered. To distinguish residues that are mobile on the sub-millisecond timescale, $^1$H $^2$F filtered NCA data sets were acquired. In this experiment, a $^1$H $^2$F filter is incorporated and signals decay according to the local $^1$H–$^1$H dipolar coupling strength. Thus, peak intensities of rigid residues are weakened to 20–30% of their peak intensity, while dynamic residues retain more than 40% of their peak intensity. $^1$H $^2$F filtered experiments indicated that many dynamic residues are in loops and/or solvent exposed, with the most mobile residues being in the CypA loop (Fig. 21), in agreement with the DOP measurements by Lu et al. (2015a). The N-terminal segment was also observed to be mobile.

4.1.3 MAS NMR of HIV-1 maturation intermediates

Understanding the viral maturation process has been of interest both from the fundamental science standpoint and for the development of small molecule maturation inhibitors as a venue for HIV-1 treatment. Despite intense research into the maturation mechanism, many key questions remain open, including how the conical CA capsids form from the immature Gag lattice. Two hypotheses have emerged to explain how capsid reorganization takes place upon maturation: (i) during

Fig. 21. (a) NCA spectrum of CA tubular assemblies, with sufficiently resolved peaks labeled. (b) $^1$H $^2$F filtered NCA spectrum with 168 µs spin echo. Label colors correspond to peak intensity from $^1$H $^2$F filtered NCA experiment: 15–30%, dark blue; 31–40%, light blue; 41–55%, green; 56–85%, magenta. Reprinted with permission from Bayro et al. (2014). Copyright 2014 Elsevier.
Fig. 22. (Panel 1) Sequence of Gag polyprotein cleavage during maturation. (Panel 2) (a–d) Direct-DARR, (e–j) INADEQUATE, and (k) CP-DARR spectra of CA (orange) and CA-SP1 (black) NL4–3 strain assembled into tubes. Selected regions show the presence of SP1 peaks not observed in the CA spectra. Reprinted with permission from Han et al. (2013). Copyright 2013 American Chemical Society.
maturation, the capsid gradually changes shape from spherical to conical, and (ii) cleavage of SP1 triggers disassembly of the immature lattice and subsequent de novo reassembly of the mature capsid core. Recently, a small molecule BVM was discovered, which was shown to inhibit viral maturation by abolishing the final step, the cleavage of the SP1 peptide from the CA-SP1 maturation intermediate (Fig. 22, panel 1) (Adamson et al. 2010; de Marco et al. 2010; Fontana et al. 2016; Nguyen et al. 2011). Studies on the BVM interaction with the Gag and maturation intermediates by EM and biochemical methods suggest that the second hypothesis is correct, i.e. capsid formation takes place through de novo assembly rather than by the gradual lattice remodeling (Keller et al. 2013). These EM studies are limited by resolution, and to gain atomic-level insights on the final step of HIV-1 maturation, MAS NMR spectroscopy was pursued. In a recent investigation, the SP1 peptide was characterized in tubular assemblies of two strains of the CA-SP1 maturation intermediate: wild type HXB2 and NL4-3 A92E mutant (Han et al. 2013). Dipolar and scalar-based correlation experiments (Fig. 22, panel 2) revealed that the SP1 peptide in the assembled state is dynamically disordered and does not adopt a stable helical structure proposed on the basis of early solution NMR experiments on isolated SP1 peptide in organic solvents (Datta et al. 2011). The presence of the SP1 peptide affects the conformation and dynamics of CA protein: residues 226–231 of the CTD directly preceding the SP1 peptide, and which are mobile on microsecond timescale in CA assemblies, become more rigid, with the motions occurring on millisecond timescale in the CA-SP1 tubes. Most surprisingly, the dynamics of the CypA loop is modulated by the presence of the SP1 peptide: the loop becomes more flexible. Taken together, the results of this investigation support the de novo reassembly hypothesis (Han et al. 2013; Keller et al. 2013; Lu et al. 2015a).

To summarize, MAS NMR is a powerful emerging method for atomic-level characterization of HIV-1 viral assemblies. Recent work has demonstrated the significance of sequence-dependent conformational changes and dynamics of the HIV-1 capsid assemblies, as well as brought to light the importance of dynamic allosteria in HIV maturation, infectivity, and interactions with host factor proteins.

4.2 Bacteriophages

Bacteriophages are a diverse class of viruses that target and infect bacteria. Bacteriophages have applications in many areas, including molecular biology (Bax & Grishaev, 2005; Messing, 2001; Smith, 1985), nanotechnology (Huang et al. 2005; Nam et al. 2006), and drug development (Clark & March, 2006; Omidfar & Daneshpour, 2015). For review see (Henry & Debarbieux, 2012; Salmond & Fineran, 2015). Two classes of bacteriophages have been studied by MAS NMR: filamentous andicosahedral bacteriophages. Filamentous phages, which include Pf1, fd, and M13 are long rod-like structures with a single-stranded DNA (ssDNA) encapsulated in a capsid, composed of symmetric repeats of several thousand copies of an α-helical coat protein (Marvin, 1998). A second class is represented by tailed icosahedral viruses, such as T4 and T7, which are composed of nucleic acid encased in an icosahedral capsid, with a tail to attach to the target host cell (Ackermann, 1999). In addition to structural studies of bacteriophage capsids, protein–nucleic acid interactions in bacteriophages have been addressed by MAS NMR.

4.2.1 Structural characterization of bacteriophage capsid proteins

McDermott and co-workers first characterized the intact Pf1 bacteriophage by MAS NMR, where resonance assignments and secondary structure determination were performed and the feasibility of this approach for atomic-level analysis was established (Goldbourt et al. 2007b). Pf1 undergoes a cooperative phase transition at 10 °C (Thiriot et al. 2005), with the high- and low-temperature forms having slightly different helical symmetries. Goldbourt et al. mapped chemical shift perturbations observed in 13C–13C correlation spectra to the structure of Pf1 (Goldbourt et al. 2010). Most chemical shift perturbations appeared in the side chains of residues in the hydrophobic region of the protein, at the interface between subunits. They postulated that adjustments in these hydrophobic regions enabled the temperature transition of what is an overall rigid structure. Hydration waters play critical roles in these hydrophobic regions, including maintenance of structure and facilitating protein dynamics. McDermott and co-workers (Sergeyev et al. 2014), and Sinha and co-workers (Purusottam et al. 2013) used 1H–15N and 1H–13C HETCOR spectroscopy, including MEduim-to-LOnge Distance (MELODI) HETCOR (Yao & Hong, 2001) to probe the hydration state of Pf1. REDOR dephasing pulses (Gullion & Schaefer, 1989) were applied during HETCOR experiments to dephase magnetization arising from directly bonded protons, such that observed magnetization is known to come from direct contact with water. McDermott and co-workers also utilized water-selective Tz-filtering for spectral simplification in some HETCOR experiments and a 3D MELODI HETCOR experiment which incorporates a train of 13C and 15N REDOR dephasing pulses on both 13C and 15N with a DARR 13C–15N mixing component which allowed for the assignment of numerous protein residues in contact with water, as shown in Fig. 23. SD-HETCOR experiments (HETCOR with 1H spin diffusion (Kumashiro et al. 1998)) were also included to allow for the observation of hydration waters on the encapsulated DNA. Hydration sites include surface exposed residues at the N-terminus, as well as C-terminal residues, which contact
the interior cavity of the phage, including Arg 44 and Lys 45, which participate in protein–DNA interactions, as discussed below. These results indicate that hydration waters are essential in the protein–DNA interactions, which stabilize the capsid structure. The data also revealed hydrophilic groves in the coat protein, which allow for water exchange between the interior of the phage and the exterior surface.

Goldbourt and co-workers reported the structure of the capsid protein from intact M13 bacteriophage (Morag et al. 2015). This is the first structure of an intact filamentous virus capsid solved by MAS NMR. The intra- and inter-subunit distance restraints were established using selectively labeled samples (Fig. 24a) (Morag et al. 2011). The ‘fold-and-dock’ protocol of CS-ROSETTA (Das et al. 2009; Shen et al. 2008) containing 35 M13 subunits to capture all unique interactions was used (DiMaio et al. 2011). This protocol is designed to determine optimum structures of higher-order oligomers. M13 and other Ff class bacteriophages have fivefold subunit symmetry around the virion axis with largely α-helical secondary structure. Many inter-subunit interactions were observed in hydrophobic regions, highlighting the importance of the hydrophobic pockets for subunit packing. Each monomer was found to participate in four hydrophobic pockets, with participating residues distributed across the monomer (Fig. 24c). Many of these key aromatic/hydrophobic residues are conserved among filamentous bacteriophages. The structure also reveals hydrogen bonding and electrostatic interactions that further contribute to capsid stability.

4.2.2 Characterization of bacteriophage capsid dynamics

Lorieau et al. probed backbone dynamics in Pf1 by Lee–Goldburg cross-polarization experiments (Lorieau et al. 2008). 1H–13C DOPs were measured for several Cα, Cβ, Cγ, and Cδ sites. The backbone of Pf1 was found to be remarkably rigid.
As would be expected, many solvent-exposed side chains exhibited reduced order parameters. Significantly, Arg 44 and Lys 45 side chains, which are directed into the phage cavity, were also observed to be highly mobile. It was postulated that the dynamics of these two residues is essential for mediating DNA–protein interactions. Opella and co-workers also measured $^1$H–$^{15}$N dipolar lineshapes of this system with $^1$H-detected experiments at 50 kHz MAS (Park et al. 2013).

4.2.3 Characterization of protein–nucleic acid interactions in bacteriophages

Nucleic acids are an integral component of viruses (Speir & Johnson, 2012). Recently several labs have demonstrated that MAS NMR can characterize nucleic acids (Cherepanov et al. 2010; Huang et al. 2012) and protein–nucleic acid interactions (Asami et al. 2013; Asami & Reif, 2013; Marchanka et al. 2015, 2013) including studies of protein–DNA interactions in intact bacteriophages (Morag et al. 2014; Sergeyev et al. 2011; Yu & Schaefer, 2008).

In contrast to other bacteriophages, Pf1 has an unusually high nucleotide-to-capsid subunit ratio of 1:1 (Day et al. 1988) and a highly extended conformation. (Other inoviruses have ratios of between 2–2.5 and 1.) Using dynamic nuclear polarization (DNP, (Ni et al. 2013)) for sensitivity enhancement, Sergeyev et al. were able to detect nucleotide resonances in the intact Pf1 bacteriophage (Sergeyev et al. 2011). The reported chemical shifts for the deoxyribose moieties capture the unusual DNA conformation of Pf1. As compared with average B-form DNA chemical shifts, C2′, C3′, C4′, and C5′ shifts in Pf1 appeared downfield, indicating a C2′-endo/gauche conformation with an anti glycosidic bond orientation. Nucleotide base chemical shifts, such as those of TC7 and CC5 fall outside the averages reported in the BMRB (Ulrich et al. 2008) indicating that Pf1 DNA does not form typical base-pairing interactions.

Goldbourt, Morag, and Abramov used $^{13}$C–$^{13}$C correlation experiments (Hou et al. 2011a, 2013a), described above in Section 2.2.1, were found to be ideal for the observation of DNA resonances, particularly for quaternary carbons with no directly attached protons. Nucleotide-specific assignments were achieved using a nucleotide walk, as shown in Fig. 25. In the CORD spectra acquired with long mixing times, sugar-to-capsid and base-to-capsid DNA–protein interactions were also observed. PHHC proton-mediated $^{31}$P–$^{13}$C correlation spectra detected capsid-to-phosphate backbone interactions. Altogether, 56 capsid–ssDNA interactions could be

---

**Fig. 24.** (a) $^{13}$C–$^{13}$C CORD spectrum of [1,3–$^{13}$C]glycerol, U-$^{15}$N M13 bacteriophage with 500 ms mixing time. Intra-residue contacts are labeled in black. Inter-residue contacts within the same subunit are labeled in green. Inter-residue contacts between residues in different subunits are labeled in blue. Select inter-subunit correlations are shown on the structure. (b) Sideview of the NMR-ROSETTA model of M13 containing 35 subunits, the minimum number of subunits required to contain all unique interactions. (c) Hydrophobic pockets formed by several subunits. Adapted with permission from Morag et al. (2015). Copyright 2015 National Academy of Sciences.
assigned. These protein–DNA interactions occur primarily between positively charged Lys side chains near the C-terminus and the DNA phosphate backbone. Fd was found to have a similar sugar conformation to Pf1 (Sergeyev et al. 2011) but a greater tendency for base stacking and a different protein–DNA interface.

Due to challenges with respect to spectral overlap and sample preparation, MAS NMR of nucleic acids has been much more limited than protein NMR. Abramov and Goldbourt were able to characterize the 40 kbp double-stranded DNA packaged within bacteriophage T7, a 50 MDa icosahedral phage (Abramov & Goldbourt, 2014). Using $^{13}$C–$^{13}$C DARR and $^{15}$N–$^{13}$C TEDOR correlations, all carbons and most nitrogens of the nucleic acid could be detected. The DNA packaged inside the phage was determined to be B-form with expected Watson–Crick base pairing. This study, along with the studies of nucleic acids in fd and Pf1, adds to the growing database of nucleic acid chemical shifts (particularly for large, native systems) and the relationship of chemical shifts to structure.

Yu and Schaefer used $^{15}$N–$^{31}$P REDOR experiments to characterize DNA packaging in bacteriophage T4 with U-$^{15}$N and [e-$^{15}$N]Lys-labeled T4 (Yu & Schaefer, 2008). The dependence of REDOR dephasing on the $^{15}$N–$^{31}$P distance enabled the determination that the DNA packaged inside T4 is B-form by determining the distance between the phosphate backbone and nitrogen atoms of the nucleic acid bases. T4 is packed into the phage capsid with over 1000 molecules of three lysine-rich proteins (Karam, 1994). The very close phosphate–lysine distances (as close as 3.5 Å) indicate that side-chain amine groups have an important role in charge balance in the phage. Further charge balance of DNA inside the phage comes from polyamines and ammonium cations.

Bacteriophages including Pf1, fd, and M13 have also been studied in membrane-associated states, as well as by static SSNMR techniques by Opella and others. These topics are beyond the scope of the current review. We direct readers to the following reviews and key publications: (Bechinger, 1997; Cross et al. 1983; Glaubitz et al. 2000; Marassi & Opella, 2003; Opella et al. 2008; Park et al. 2010; Shon et al. 1991; Tan et al. 1999; Thiriot et al. 2004, 2005; Zeri et al. 2003).

### 4.3 Other viral assemblies

Pintacuda and co-workers reported a study of the MeV nucleocapsid (Barbet-Massin et al. 2014a). Prior to this work, no atomic-level information was available for assembled nucleocapsid. The MeV nucleocapsid is comprised a relatively rigid NCORE domain and a more dynamic NTAIL domain. Proton detection was applied to both intact and cleaved assemblies.
with MAS at 60 kHz. Dipolar \(^{15}\text{N}-^{1}\text{H}\) correlations revealed rigid residues of \(N_{\text{CORE}}\), while scalar-based correlations detected mobile residues of \(N_{\text{TAIL}}\). Capsid hydration levels were also evaluated by observing \(T_{2p}\), relaxation of \(^{15}\text{N}-^{1}\text{H}\) peak intensities. The reduced intensities of the cleaved nucleocapsid relative to intact capsid indicate that the presence of the \(N_{\text{TAIL}}\) domain leads to a less ordered and hence more hydrated conformation.

5. Conclusions and future perspectives

MAS NMR has developed into a compelling technique for the characterization of biological assemblies at the atomic level. Integrated approaches to protein structure determination utilizing MAS NMR structural restraints in conjunction with information from other methods, such as cryo-EM and MD, are very promising for structural and dynamics analysis of high molecular weight assemblies. Continued advancements in sensitivity, resolution, and methodology will enable the detailed characterization of increasingly complex systems, such as whole cells, where substantial progress has already been made (Curtis-Fisk et al. 2008; Janssen et al. 2010; Pius et al. 2012; Reichhardt & Cegelski, 2014).

Acknowledgements

This work is supported by the National Institutes of Health NIGMS grants P50GM082251 and F32GM113452 and is a contribution from the Pittsburgh Center for HIV Protein Interactions. C.M.Q. acknowledges the support of the National Institutes of Health Postdoctoral Fellowship grant F32GM113452.

References


Agarwal, V., Diehl, A., Skrynnikov, N. & Reif, B. (2006). High resolution \(^1\text{H}\) detected \(^{1}\text{H},^{13}\text{C}\) correlation spectra in MAS solid-state NMR using deuterated proteins with selective \(^1\text{H},^{2}\text{H}\), isotopic labeling of methyl groups. Journal of the American Chemical Society 128, 12620–12621.


KAMINSKYJ, V. & ZHIHOVYTSKY, B. (2010). To kill or be killed: how viruses interact with the cell death machinery. Journal of Internal Medicine 267, 473–482.


WATT, E. D. & RIENSTRA, C. M. (2014). Recent advances in solid-state nuclear magnetic resonance techniques to quantify biomolecular dynamics. Analytical Chemistry 86, 58–64.


