

^{13}C Direct Detected NMR for Challenging Systems

Isabella C. Felli* and Roberta Pierattelli*

Cite This: *Chem. Rev.* 2022, 122, 9468–9496

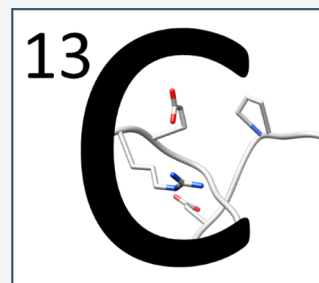
Read Online

ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: Thanks to recent improvements in NMR spectrometer hardware and pulse sequence design, modern ^{13}C NMR has become a useful tool for biomolecular applications. The complete assignment of a protein can be accomplished by using ^{13}C detected multinuclear experiments and it can provide unique information relevant for the study of a variety of different biomolecules including paramagnetic proteins and intrinsically disordered proteins. A wide range of NMR observables can be measured, concurring to the structural and dynamic characterization of a protein in isolation, as part of a larger complex, or even inside a living cell. We present the different properties of ^{13}C with respect to ^1H , which provide the rationale for the experiments developed and their application, the technical aspects that need to be faced, and the many experimental variants designed to address different cases. Application areas where these experiments successfully complement proton NMR are also described.



CONTENTS

1. Introduction	9468
1.1. Properties of Heteronuclear Spins	9469
1.2. Building Blocks of NMR Experiments: What's New for ^{13}C Direct Detection?	9470
1.2.1. Homonuclear Decoupling	9470
1.2.2. Starting Polarization Source	9471
2. Suite of ^{13}C Direct Detection NMR Experiments	9472
2.1. Sequence-Specific Assignment	9472
2.2. Multidimensional Experiments ($n\text{D}$, $n > 3$)	9475
2.3. CON and CACO Fingerprints	9476
2.4. Useful NMR Observables	9477
2.4.1. Chemical Shifts and Heteronuclear Relaxation Rates	9477
2.4.2. Chemical Exchange	9478
2.4.3. Residual Dipolar Couplings	9479
2.4.4. What about Homonuclear Nuclear Overhauser Effects (NOE)?	9479
3. Biomolecular Applications	9479
3.1. Focus on Amino Acid Side-Chains	9479
3.1.1. Monitoring Side-Chains through C' -Detected Experiments	9480
3.1.2. Positively Charged Amino Acids	9480
3.1.3. Hydrophobic Amino Acids	9482
3.1.4. The Special Case of Proline	9483
3.2. Overcoming Limitations Due to ^1H Transverse Relaxation: Paramagnetic and Large Proteins	9484
3.3. Intrinsically Disordered Proteins	9485
3.3.1. Mimicking Physiological Events	9486
3.3.2. Liquid–Liquid Phase Separation	9486
3.3.3. Low Complexity Regions	9487
4. Conclusions and Outlook	9488

Author Information	9489
Corresponding Authors	9489
Notes	9489
Biographies	9489
Acknowledgments	9489
Abbreviations	9490
References	9490

1. INTRODUCTION

NMR spectroscopy is an indispensable tool for investigations of biological molecules and their interactions. The power of NMR to link structural, dynamic, kinetic, and thermodynamic information makes it an essential component of cutting-edge research in structural biology. Provided NMR spectra can be acquired with high resolution and sensitivity, a virtually unlimited amount of atomic-resolution information can be achieved starting from chemical shift values, nuclear spin relaxation rates, scalar couplings, exchange effects, diffusion coefficients, and other highly sophisticated observables. This places NMR in a unique position with respect to the many spectroscopic methods that provide global information or to different high-resolution methods, such as X-ray crystallography and cryo-electron microscopy, that however fail to provide information about macromolecules in solution for highly dynamic or heterogeneous ones in particular.

Special Issue: Biomolecular NMR Spectroscopy**Received:** October 13, 2021**Published:** January 13, 2022

Needless to say that continuous development of the experimental approach is necessary to exploit at best this powerful spectroscopic technique to extend the complexity of the systems under investigation, as required by the many challenges in biomedical research. These would benefit from the availability of high resolution structural and dynamic details on biologically relevant molecular components. This in turn triggers instrumental technological improvements that enable expansion of the applications to problems of increasing complexity.

We would like to focus here on the example of carbon-13 direct detection NMR in solution, one of the widely used tools to characterize biological macromolecules. We are going to start by introducing the key properties of heteronuclear spins that make heteronuclear direct detection interesting and discuss the many strategies developed to overcome potential critical points, such as the problem posed by homonuclear decoupling in the direct acquisition dimension. The issue of the starting polarization source as well as coherence transfer methods is described since they constitute common basic ingredients of more complex NMR experiments. The flow of the review then proceeds to illustrate the many experiments developed, focusing on the applications where they reveal unique additional information with respect to more conventional approaches. Current challenging research areas where these methods provide useful data are also presented.

1.1. Properties of Heteronuclear Spins

The term “heteronuclear” refers to nuclear spins other than protons, the first ones to be considered in general in NMR spectroscopy (the magnetic field of an instrument, B_0 , is indeed often indicated through the ^1H Larmor frequency). The most widely used nuclear spins for biomolecular NMR investigations are ^1H , ^{13}C , ^{15}N , and ^{31}P (the latter of interest for the investigation of nucleic acids); we will focus here on the contributions of ^{13}C and ^{15}N to the study of proteins even if similar arguments can of course be extended to nucleic acids (including ^{31}P) and carbohydrates.

While sharing the same spin angular momentum ($S = 1/2$), heteronuclei have smaller gyromagnetic ratio (in absolute value) with respect to proton and thus smaller magnetic moment. The latter, which interact with the external magnetic field B_0 determining the Larmor frequency, is responsible for the intrinsic sensitivity that scales down moving from ^1H to ^{13}C to ^{15}N .¹ The magnetic moment associated with nuclear spins is also responsible for the dipole–dipole (DD) interactions with other nuclear spins in the surroundings, interactions that in isotropic solution are largely averaged by fast molecular tumbling and do not have an impact on line positions or do not cause line splitting like the scalar coupling. They are however the major interactions that promote nuclear relaxation. Therefore, in principle heteronuclear spins sense lower dipolar contributions to relaxation as a result of their lower magnetic moments. Considering the interaction between two dipoles at a specified distance, heteronuclei sense a lower contribution to relaxation with respect to protons by a factor of approximately $(\gamma_X/\gamma_H)^2$ (neglecting contributions from the different spectral densities). Paramagnetic contributions to nuclear relaxation provide a clear example in this respect. Once a particular paramagnetic center is defined, dipolar contributions to relaxation of nuclear spins in the surroundings depend on the distance ($1/r^6$) and on the square of the gyromagnetic ratio (γ_X^2) of the nucleus itself (as well as on spectral densities). Therefore, at a fixed distance from

the paramagnetic center, heteronuclear spins sense a smaller dipolar interaction and thus a smaller contribution to relaxation with respect to protons; in other words, the so-called paramagnetic relaxation enhancements at a specified distance from a paramagnetic center are smaller for heteronuclei than for protons.

Similar arguments hold for diamagnetic contributions to relaxation: dipolar interactions involving heteronuclear spins are lower than those involving protons, not easy to generalize in this case because the final result depends on many contributions which in turn depend on the local chemical structure. On the other hand, contributions to dipolar relaxation deriving from proton spins are generally dominant, both for proton itself and for heteronuclei due to the large γ_H , a situation that pushed the development of isotopic enrichment in deuterium to reduce the bath of proton dipoles in which heteronuclei are immersed and thus reduce contributions to nuclear relaxation.^{2–4} Another example in this respect is provided by ^{15}N relaxation often measured to investigate local dynamics in proteins. In this case, the dominant dipolar contribution to ^{15}N relaxation is provided by the directly bound proton.

The other striking difference when moving from protons to heteronuclei consists of the electronic structure/chemical environment that influences signals chemical shift. Largely averaged in solution due to fast molecular tumbling, the isotropic part of the chemical shift tensor determines peak positions. The major contributions to chemical shifts derive thus from the local molecular topology that translates into the different chemical shifts expected for the different functional groups, a property often exploited in the study of small molecules. In proteins, the chemical structure of the different amino acids as well as their link through the peptide bond is the major determinant of the observed signals chemical shift. The local 3D structure provides an additional contribution to it, which clearly shows a general trend toward larger chemical shift dispersion for heteronuclei with respect to protons, as can be appreciated by inspecting the Biological Magnetic Resonance Data Bank (BMRB, <https://bmr.io/>), the database in which chemical shifts of assigned proteins are deposited.

The anisotropic part of the chemical shift tensor instead contributes to nuclear relaxation. Therefore, large contributions to relaxation are expected when significant chemical shift anisotropy (CSA) is present such as for example for nuclear spins involved in the peptide bond or aromatic rings, certainly an important aspect to consider when exploiting heteronuclear spins. Several solutions have indeed been proposed to use constructively interference between CSA and DD to mitigate the contributions to transverse relaxation that may broaden lines beyond detection for globular proteins of increasing molecular mass due to the increasing rotational correlation time. On the other hand, these contributions to nuclear relaxation do not have a detrimental impact when focusing on highly flexible proteins, in which the fast motions reduce contributions to transverse relaxation.

Scalar couplings have a strong impact on the spectra, in particular in the direct acquisition dimension, and their magnitude and topology are also influenced by the type of nuclei under investigation. These are mediated by electronic effects and depend on many factors including the gyromagnetic ratio of the nuclear spins involved in the coupling as well as the electronic structure and the local geometry. However, the most striking difference moving from ^1H to ^{13}C relates to the fact that large one-bond homonuclear scalar couplings are present when

focusing on ^{13}C nuclear spins in uniformly labeled samples, while analogous ones are not observed for protons. This property, widely exploited in many multidimensional experiments to achieve coherence transfer, also causes very large signals splitting and the relative complex multiplet structures complicate direct detection of ^{13}C . This constitutes an important aspect to be considered in order to convert ^{13}C direct detection into a useful tool for the study of complex macromolecules. On the other hand, when moving to multiple bond effects, these are smaller than the analogous ones involving ^1H (for example, ranges for $^3J_{\text{HH}}$ are often larger than those for $^3J_{\text{CC}}$ in aliphatic chains).

Heteronuclei constitute the molecular backbone; protons are at the edge of chains of chemical bonds and in many cases form the exposed surfaces of macromolecules. This property is often used in the study of interactions in which changes in ^1H chemical shifts are investigated to identify interaction surfaces. Also heteronuclei are of course affected by changes in the nucleus surrounding but are more sensitive to changes in the local structure, such as changes in dihedral angles, and these complementary features can result useful.

Proteins are studied in water, the solvent of life. Interactions of proteins with water are key for many aspects: they prompt polypeptide chains to fold into stable globules or to remain flexible and solvent-exposed as well as to create membrane-less organelles through liquid–liquid phase separations. Therefore, the interactions with the solvent have a very relevant role in protein function. Proton NMR can be used to detect changes with pH or protonation state of a protein, but it is also influenced by exchange processes that, depending on their magnitude, can broaden lines beyond detection. Heteronuclei in this context can act as “spies” of nonprotonated states providing information also in cases in which protons are not present or when fast exchange between the free and bound forms broadens ^1H NMR lines beyond detection.

1.2. Building Blocks of NMR Experiments: What’s New for ^{13}C Direct Detection?

Heteronuclear spins have interesting properties as also exploited in the indirect dimensions of many multidimensional NMR experiments based on ^1H direct detection (“inverse detection” of heteronuclei). We would like to discuss here the most important aspects to consider when moving to ^{13}C detection.

Starting from simple 1D ^{13}C NMR spectra, often used also for the study of small molecules to identify specific functional groups or coupling topologies, the large chemical shifts dispersion is accompanied, when studying isotopically labeled macromolecules, by the onset of complicated multiplets determined by the large one bond homonuclear ^{13}C – ^{13}C scalar couplings. These can range from simple doublets for ^{13}C nuclei that only have one ^{13}C bound, or to more complicated multiplet structures observed for the different side-chains with ^{13}C nuclei directly bound to two or three other ^{13}C nuclei. The onset of complicated multiplets, that on one hand can provide valuable information regarding the type of spin system, is detrimental when seeking high-resolution information needed to study macromolecules. Indeed, the complex multiplet structure of ^{13}C signals deriving from homonuclear scalar couplings drastically reduces both sensitivity and resolution of the NMR spectra, two key features for the study of complex macromolecules.

Several strategies have become widely used to suppress these homonuclear couplings in the indirect dimensions of NMR spectra such as the inclusion of constant time evolution periods,

the use of band-selective inversion pulses to refocus scalar coupling evolution, etc.⁵ In many cases, the resolution that can be achieved in the indirect dimension is however limited by the number of points that can be acquired in the time that can be dedicated to a specific experiment and the reduction in resolution brought about by signals’ splitting is seldom a limiting factor. On the other hand, when nuclear spins are investigated in the direct acquisition dimension, the FID can in principle be acquired as long as desired (just limited by the transverse relaxation properties of the system under investigation rather than by the time needed to accumulate increments to construct an indirect dimension). This constitutes a contribution to spectral resolution, provided the complex multiplet structures are simplified by the implementation of homonuclear decoupling. The problem of ^{13}C homonuclear decoupling is more demanding than heteronuclear decoupling because the two nuclear spins involved in the coupling are close in frequency and thus radio frequency irradiation on one of them can be sensed by the other. However, several elegant solutions have been proposed and constitute a relevant aspect for the design of NMR experiments for the study of complex biomolecules.

1.2.1. Homonuclear Decoupling. Let us start discussing the “simple” case of a ^{13}C nuclear spin that only has one large bond homonuclear coupling such as for example carbonyl nuclear spins ($^{13}\text{C}'$) in protein backbones, which share a large one bond scalar coupling with $^{13}\text{C}^\alpha$ nuclear spins. One nice feature of the one bond scalar couplings ($J_{\text{C}'\text{C}^\alpha}$) consists of the fact that they are not as variable as, for example, three bond couplings, which largely depend on the local conformation of the molecule. In the case of carbonyl moieties of protein backbone, the one bond scalar coupling with C^α is fairly uniform throughout the primary sequence and relatively independent of the type of amino acid and on the local conformation,⁶ a property that renders the problem of homonuclear decoupling generally amenable and more easy to address with respect to, for example, ^1H homonuclear decoupling. The most straightforward approach thus consists of deconvolution of the spectra, exploiting a defined value of the coupling. Initially proposed for indirect dimensions of triple resonance experiments⁷ and implemented for ^{13}C detected experiments,^{8,9} this approach has been recently revived by the incorporation of AI methods.¹⁰

The other possibility, provided the chemical shifts of the two nuclear spins involved are sufficiently different to allow for their selective irradiation, consists of band-selective homonuclear decoupling in which the acquisition time is shared between acquisition and decoupling mode in alternating time intervals (Figure 1).^{11–13} This approach requires an additional radio frequency channel; decoupling sidebands are observed depending on the frequency of acquisition and decoupling periods. An elegant alternative approach to homonuclear decoupling in which the acquisition time is shared between acquisition and 180° refocusing pulses is the BASHD.¹⁴ Introduced for solid-state NMR experiments,¹⁵ it was adapted for direct detection of carbonyl carbon nuclei in solution measurement.

The most widespread approach used nowadays consists of using virtual homodecoupling, that is exploiting spin-state-selective methods in which scalar couplings are preserved and different components of the signal are collected: these constitute the basis to achieve virtual decoupling through the appropriate linear combinations of the acquired signals to separate the different multiplet components, followed by a shift to the center of the original splitted signal.¹⁶

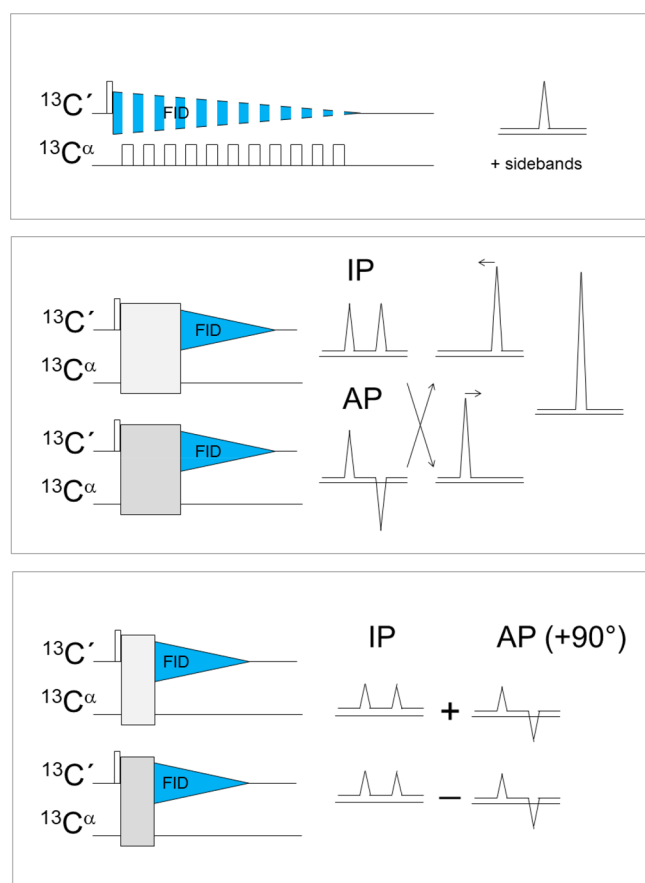


Figure 1. Homonuclear decoupling strategies initially implemented for ^{13}C direct detection: band-selective homonuclear decoupling (top panel), virtual decoupling achieved through the IPAP (middle panel), and S 3 E (bottom panel) approaches. A scheme illustrating the linear combinations performed to achieve virtual decoupling is also reported on the right side of the middle panel. The contribution of the IP and AP components to the two independent FIDs acquired through the S 3 E approach is also schematically indicated on the right side of the bottom panel.

Several experimental variants based on this idea have been developed and may result useful for different applications. The most straightforward implementation of this idea consists of the IPAP approach¹⁷ in which the in-phase (IP) and antiphase (AP) components of the carbonyl carbon signals (with respect to C^α) are acquired and separately stored (Figure 1).^{18–20} The postacquisition treatment of the acquired data, which can also be performed directly in the time domain, enables the removal of the large one bond splitting from the spectra. Interestingly, this approach preserves the coupling and can in principle be applied to mutually decouple and still observe the two nuclear spins involved in the coupling, as demonstrated in solid-state applications.²¹ An experimental variant in which AP and IP signals for virtual decoupling are acquired sequentially in a single scan has been recently proposed.²²

When fast transverse relaxation becomes a limiting factor, shorter experimental variants can be implemented.^{18,20,23} Indeed the IPAP approach relies on complete interconversion between the IP and AP components of the signal and requires a time that is inversely proportional to the scalar coupling itself ($1/2J_{\text{C}^\alpha\text{C}^\beta}$). However, partial interconversion between in-phase and antiphase already provides the two components in half the time; in this case, changing the sign of one of the two

components in alternate scans allows storage of two FIDs that can be used to separate the two multiplet components needed to perform homonuclear decoupling through appropriate manipulations of the acquired FIDs (Figure 1).²⁰

Direct acquisition of the AP signal component was also proposed for systems in which transverse relaxation is a key limitation such as for very large proteins²⁴ or for paramagnetic ones.²⁵ Finally, different variants exploiting sensitivity enhancement strategies were proposed for homonuclear decoupling in COCA (COCAINE)²⁶ and CON²⁷ experiments. They were also shown to be useful alternatives for heteronuclear decoupling.²⁸

These principles implemented for backbone carbonyl homonuclear decoupling can of course be extended to analogous cases such as ^{13}C nuclear spins that only have one large scalar coupling with a second ^{13}C nuclear spin that also has a different chemical shift, sufficiently different to allow band-selective inversion of the two spins independently. Indeed this strategy also performs well for amino acid side-chains that have a carbonyl/carboxylate moiety such as aspartate, glutamate, asparagine and glutamine residues. It has also been implemented to investigate terminal nuclear spins of aliphatic side-chains and to decouple them from their next neighbors.^{29–32}

The situation becomes more complex for spins that are coupled to more than one additional ^{13}C spin through large one-bond scalar couplings such as $^{13}\text{C}^\alpha$ spins (for all amino acids except glycine) as well as for the vast majority of spins of amino acid side-chains. The ideas described can in principle be extended also to this case; selected applications were so far implemented. As an example, several approaches were proposed to perform homonuclear decoupling of $^{13}\text{C}^\alpha$ signals by clever combinations of spin-state selective approaches (DIPAP, DS 3 E).^{20,30} The capacity to selectively invert $^{13}\text{C}^\alpha$ from $^{13}\text{C}^\beta$ (in addition to $^{13}\text{C}'$) constitutes a key feature, and some compromises may be necessary for signals with similar $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ shifts that fall close to the transition regions of the inversion profiles of the band-selective pulses employed. The acquisition of the different components also has a cost in terms of overall sensitivity. Despite these additional complications, these methods were successfully used to investigate very large proteins in which extensive isotopic enrichment was mandatory and for which carbonyl carbon direct detection provided too broad lines.³⁰ Similar approaches were also proposed for the investigation of nucleic acids through ^{13}C direct detection^{33–37} as well as to focus on aromatic residues in proteins.³⁸

1.2.2. Starting Polarization Source. The starting polarization source used in NMR experiments also constitutes an important element that can be used to modulate experimental sensitivity. On the other hand, the latter also depends on the kind of information desired. Direct detection of heteronuclei for biomolecular applications in solution, after pioneering work in 1988³⁹ was abandoned in favor of inverse detection methods.^{40–42} It was proposed at the beginning of 2000 as a strategy to recover information on paramagnetic proteins in regions in which proton resonances are broadened beyond detection due to paramagnetic relaxation enhancements.^{38,43–47} Dipolar contributions to relaxation sensed by nuclear spins when a paramagnetic center is present in a molecule depend on the properties of the paramagnetic center itself, on the effective correlation time modulating the interaction, on the electron–nuclear distance ($1/r^6$), and on the square of the gyromagnetic ratio of the nuclear spin under investigation (γ_X^2).⁴⁸ Therefore, shifting our attention from ^1H to ^{13}C and ^{15}N ensures reduction

of the paramagnetic enhancement (considering the same distance); from a different point of view, similar enhancements are sensed at shorter distances from the paramagnetic center and thus shifting the focus from ^1H to ^{13}C (and to ^{15}N) enables researchers to observe resonances of nuclear spins closer to the paramagnetic center since the additional contributions to relaxation are scaled by $(\gamma_{\text{X}}/\gamma_{\text{H}})^2$.^{49,50}

When considering experiments that are more complicated than 1D experiments, these should not actively perturb ^1H nuclear spins in any of the coherence transfer steps because this would reintroduce the dependence on ^1H transverse relaxation, much more sensitive to paramagnetic relaxation enhancements. For this reason initial variants of ^{13}C detected experiments for biomolecular applications in solution were based on ^{13}C as a starting polarization source and never exploited protons in any of the coherence transfer steps; these experiments were called “protonless” NMR experiments.²⁰ These experiments also resulted useful for the study of very large proteins in which high levels of deuteration were necessary,⁵¹ reducing the amount of information that could be achieved through proton direct detection.^{9,19,52}

While focusing on this topic, it became evident that heteronuclear direct detection could result useful also for different applications in which ^1H fast relaxation does not constitute a limiting factor.^{23,53,54} For these applications, the use of ^1H as a starting polarization source brings a significant increase in experimental sensitivity,^{55–58} still exploiting heteronuclear chemical shifts in all the detected dimensions of multidimensional NMR experiments to benefit from their contribution to resolution, a key feature when focusing, for example, on proteins devoid of a stable tertiary structure. These experiments were thus generally referred to as “exclusively heteronuclear” to indicate that only heteronuclei were frequency labeled in all dimensions, regardless of the starting polarization source exploited.^{24,28,53,59}

Considering for the moment backbone nuclear spins, the proton polarization source can be provided by amide protons ($^1\text{H}^{\text{N}}$) as well as by aliphatic protons ($^1\text{H}^{\alpha}$) with advantages and disadvantages that depend on the properties of these two nuclei (Figure 2). Amide protons, which can be easily correlated to the directly bound nitrogen and then to carbonyl carbon nuclei through the $^1J_{\text{HN}}$ and $^1J_{\text{CN}}$, respectively, are influenced by

solvent exchange that can become so pronounced that coherence transfer becomes inefficient. In addition proline residues do not have an amide proton, a feature that reduces the information content of CON spectra that exploit $^1\text{H}^{\text{N}}$ polarization as a starting source, in particular for proline-rich proteins.

Backbone aliphatic protons (H^{α}) are instead present for all amino acids and are nonexchangeable nuclei. Therefore, experiments based on $^1\text{H}^{\alpha}$ as a starting polarization source can in principle provide complete information. Also in this case coherence transfer pathways exploiting large scalar couplings ($^1J_{\text{H}^{\alpha}\text{C}^{\alpha}}$, $^1J_{\text{C}^{\alpha}\text{C}^{\beta}}$, and $^1J_{\text{C}^{\alpha}\text{N}}$) enable the transfer of ^1H polarization to backbone nuclear spins in an efficient way without major losses; care should be taken to consider the role of $^{13}\text{C}^{\alpha}$ – $^{13}\text{C}^{\beta}$ couplings in the coherence transfer pathway to ensure that no information is lost for amino acids with similar $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ chemical shifts.

Longitudinal relaxation enhancement (LRE) strategies can be used to reduce the interscan delay and reduce experimental time (or increase the S/N per unit time).⁶⁰ Borrowing ideas proposed for amide proton detected NMR experiments, in which amide protons are selectively perturbed to enhance the recovery to equilibrium and reduce interscan delays (SOFAST, BEST),^{61,62} different variants of $^1\text{H}^{\text{N}}$ -start experiments were proposed exploiting $^1\text{H}^{\text{N}}$ band-selective pulses ($\text{H}^{\text{NBESTCON}}$).⁶³ It is interesting to note that in these experiments ^1H polarization is only used as a starting polarization source so that the longitudinal recovery starts well before acquisition of the FID, a feature that enabled the acquisition of the $\text{H}^{\text{NBESTCON}}$ without introducing any longitudinal recovery delay after the acquisition of the FID.⁶³

Similar longitudinal relaxation enhancement approaches would be useful also for $^1\text{H}^{\alpha}$ protons, used as a starting polarization source in several variants of exclusively heteronuclear NMR experiments. However, $^1\text{H}^{\alpha}$ spins are more difficult to manipulate through band-selective pulses with respect to the case of $^1\text{H}^{\text{N}}$ protons because they fall in a more crowded spectral region. To this end, a variant to selectively manipulate a subset of nuclear spins while leaving others unaffected was proposed that exploits the scalar coupling with the attached heteronuclear spins (H^{flip}).⁵⁹ In this way, longitudinal relaxation enhancement can be implemented in a general way for protons used as a starting polarization source of exclusively heteronuclear NMR experiments, both for $^1\text{H}^{\text{N}}$ -start and $^1\text{H}^{\alpha}$ -start variants (Figure 2).

The use of proton–nitrogen cross-polarization, in place of the more widespread INEPT block, was also proposed since this excitation mechanism uses the large water-magnetization reservoir to continuously replenish the amide-proton; combined with the CON reading scheme, this represents another useful approach to study systems in which solvent exchange is very pronounced.^{64,65}

2. SUITE OF ^{13}C DIRECT DETECTION NMR EXPERIMENTS

2.1. Sequence-Specific Assignment

Similarly to ^1H detected experiments, ^{13}C detected experiments can be differentiated based on the active interactions exploited in the coherence transfer steps (scalar or dipolar coupling), the starting polarization source (^{13}C or ^1H), and the kind of detected nuclei ($^{13}\text{C}'$ or $^{13}\text{C}^{\alpha}$ or others).

The experiments initially proposed for assignment purposes were based on ^{13}C -start, ^{13}C -detection, and completely avoided protons in any of the magnetization transfer steps (protonless

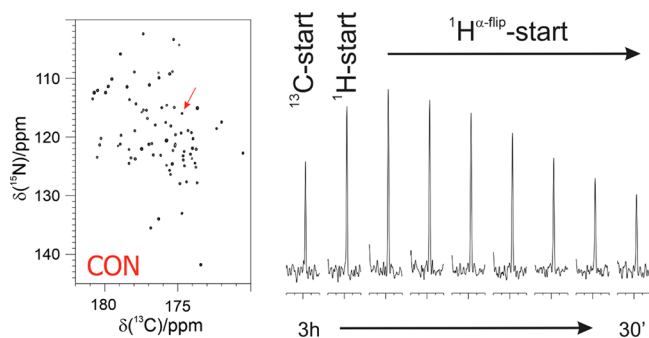


Figure 2. Intensity of one of the cross peaks obtained in 2D CON experiments (the correlation for C'59–N60 of ubiquitin indicated by an arrow in the full spectrum reported on the left) recorded with different pulse schemes and different interscan delays (d1) is compared by showing its trace. From left to right: ^{13}C -start (relaxation delay d1 = 2.5 s), ^1H -start (d1 = 1.5 s), and $^1\text{H}^{\alpha}$ -flip with different d1 (1.5, 1.2, 0.9, 0.7, 0.5, 0.35, 0.2 s). Adapted from ref 59. Copyright 2009 American Chemical Society.

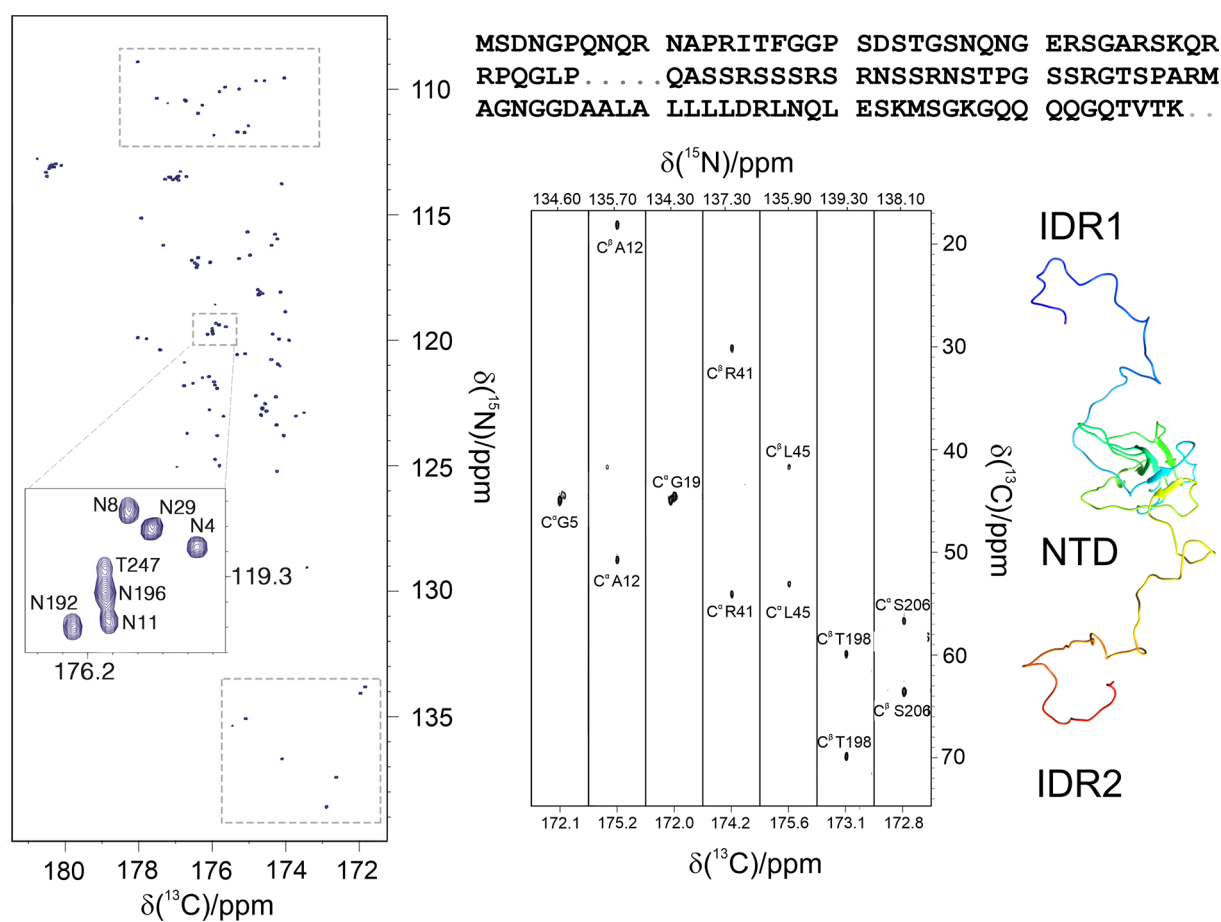


Figure 3. Left: The 2D CON of a 248 amino acids long construct of the SARS-CoV-2 nucleocapsid protein comprising the NTD folded domain and the two intrinsically disordered regions flanking it (IDR1-NTD-IDR2). The high resolution provided by this experiment enables to easily resolve resonances in the usually very crowded Gly-region (upper squared region) and to directly observe connectivities involving proline residues (lower squared region). In the expansion shown in the center of the map, the resolution of several repeating fragments comprising asparagine residues can be appreciated (the assignment reported is referred to the amide nitrogen of the mentioned amino acid). Right: Seven strips derived from the 3D (H)CBCACON experiment extracted at the ^{15}N chemical shift of proline residues, flanked by a cartoon of the IDR1-NTD-IDR2 construct. The C', C^α, and C^β frequencies belong to the preceding amino acid leading to the X-Pro assignment. The upper part of the panel reports the IDR1-NTD-IDR2 primary sequence (the sequence of the NTD domain is omitted for sake of clarity). Five proline residues are found in the IDR1 and two in IDR2 domain. Adapted from ref 73. Copyright 2021, The Authors under the terms of a Creative Commons CC BY license <http://creativecommons.org/licenses/by/4.0/>.

NMR experiments).²⁰ The set of protonless NMR experiments included various kinds of 2D experiments correlating carbon nuclei such as the ^{13}C – ^{13}C COSY, ^{13}C – ^{13}C TOCSY, and the CACO MQ correlation experiment.^{20,30,43,45,46,66,67} The ^{13}C – ^{13}C NOESY experiment proved also useful in cases in which fast transverse relaxation represents a major limitation.^{9,19} Despite the ^{13}C – ^{13}C cross-relaxation rates being expected to be much smaller than ^1H – ^1H ones, the ^{13}C – ^{13}C NOE effects between directly bound carbon nuclei are easily detectable and spin diffusion through their nuclei is a very efficient process able to provide very useful spectra for amino acid type identification. These initial results opened the way to the development of a wide variety of experiments for sequence-specific assignment.

A backbone ^{13}C nucleus that can be used to design a whole series of experiments is C^α. It is characterized by a large chemical shift dispersion, so it can provide highly dispersed spectra. However, the coupling to the C^β present in all amino acids other than glycine makes homonuclear decoupling less straightforward (see section 1.2.1). Nevertheless, the CAN experiment is useful to detect the two correlations between C^α and the intra- and inter-residue nitrogen.⁵⁰ Variants to highlight the sequential

correlation and thus discriminate it from the intraresidue have been proposed.⁶⁸ These gain also in resolution since the inter-residue correlations are generally more resolved than the intraresidue ones.⁶⁹

Thanks to the potential of experiments based on C^α direct detection for the study of higher molecular mass proteins, it was proposed to exploit the isotopic labeling strategy designed for solid-state applications that avoids neighboring ^{13}C spins (and thus also the need for homonuclear decoupling in the direct ^{13}C dimension).^{51,70} Several experiments based on C^α direct detection exploiting this isotopic labeling scheme were proposed including the CANCA and the CACA TOCSY experiments.^{68,71}

The most successful applications of ^{13}C direct detected NMR rely on carbonyl carbon detection, which presents a single and uniform splitting that can be easily removed (see section 1.2.1). Carbonyl carbon nuclei can be directly correlated to the neighboring C^α nuclei through a ^{13}C – ^{13}C transfer step ($^1J_{\text{C}^\alpha\text{C}^\beta}$), which can be further correlated to the C^β and to all the aliphatic side chain to obtain the 2D CBCACO and 2D CCCO experiments.²⁰ By introducing a further transfer step, it is also possible to include an additional dimension in which the

attached nitrogen is frequency labeled.^{20,72} Therefore, the 3D experiments CACON, CBCACON, and CCCON were designed.⁷² These enable the identification of all the spin systems in a protein including those involving proline residues. Actually the correlations involving the proline nitrogen with the carbon frequencies of the previous amino acid ($^{13}\text{C}'$, $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, and other aliphatic ^{13}C spins) often become the starting point for sequence-specific assignments because, already at this initial stage, they provide inter-residue information regarding X-Pro pairs, as illustrated for two intrinsically disordered regions of the nucleocapsid protein from SARS-CoV-2 (Figure 3).⁷³

Several additional experiments were then proposed in which different couplings were exploited to detect the complementary correlations to link amino acid spins systems in a sequence-specific manner. The CANCO experiment,^{20,74} and its variant also including the information about the C^β chemical shift (CBCANCO),⁷⁵ provides the complementary information for sequence-specific assignment. These experiments exploit C^α and C^β chemical shifts to match different spins-systems identified in CBCACON/CCCON spectra and enable researchers to link them to achieve their sequence-specific assignment. However, additional information is useful when focusing on complex systems, in particular when contributions to chemical shifts are drastically reduced due to high flexibility and disorder of the polypeptide. Other experiments were thus designed that provide correlations involving carbonyl carbon nuclei as a further tool to reduce potential ambiguities.^{23,67,76} This is the case of the COCON experiment, which correlates the backbone nitrogen with the attached carbonyl carbon and with the previous and following carbonyl carbon nuclei in the sequence exploiting the $^3J_{\text{C}'\text{C}}$ scalar coupling.^{23,76}

As previously described, in most cases of practical interest it is possible to exploit ^1H as a starting polarization source to boost sensitivity. This has been accomplished for the multinuclear experiments described above to obtain the (H)CBCACON, (H)CCCON, (H)CBCANCO, and (HCA)COCON variants.^{28,77} Moreover, additional experiments were designed that needed a leap in S/N to be feasible due to their intrinsic low sensitivity such as the case of the 3D NMR experiments correlating several residues in a row ((H)NCANCO, (H)N-(COCA)NCO, and (HN)CO(CA)CON).^{28,78}

As an example of the quality of the spectra, a few slices extracted from the 3D spectra acquired on quail osteopontin, a disordered protein, illustrate the process of sequence-specific assignment through this strategy based on 3D C' -detected NMR experiments (Figure 4).

In the case of particularly crowded spectra, these experiments can also be tailored to select resonances of specific amino acids to simplify the spectra. There are many ways to select the signals of a specific amino acid or a group of amino acids with common characteristics. Implementation of multiple quantum filters to select XH_n groups, exploitation of band selective pulses for specific nuclei excitation, tuning of specific delays to select coherence transfer pathways, or matching of increasing numbers of coherence transfer steps to a particular spin topology are all selection blocks that can be included in a pulse sequence. With combinations of these approaches, it is possible to simplify the spectra and identify resonances belonging to virtually all different amino acids.

These experiments are particularly well suited to simplify crowded spectra of highly flexible proteins, which generally have favorable relaxation properties that limit losses resulting from the multiple magnetization transfer steps used for filtering, and

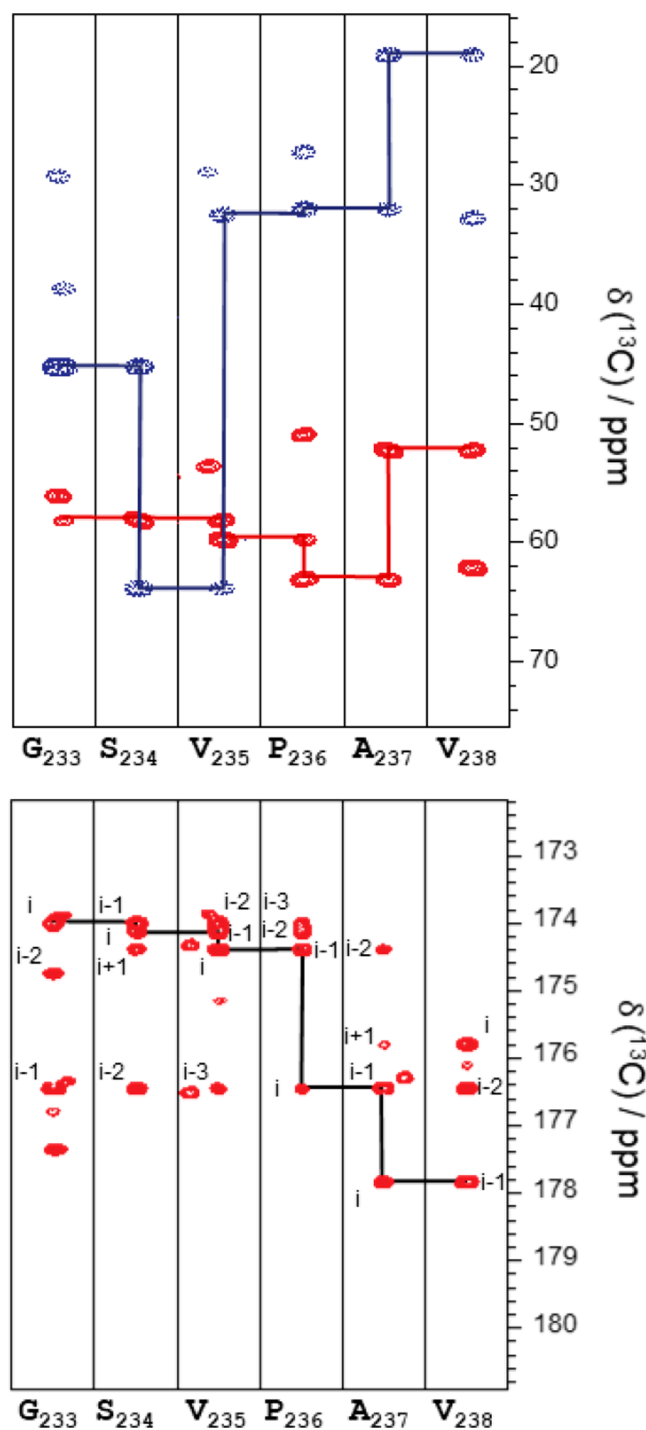


Figure 4. Assignment strips for quail osteopontin. Upper panel: Strip plots of the region 233–238 of the 3D (H)CBCANCO allowing the connectivity of dipeptides by C^α and C^β carbon nuclei. Lower panel: 3D (HCA)COCON strip plots of region 233–238 showing the connectivity between at least three consecutive C' signals. Reprinted in part from ref 77. Copyright 2019 Elsevier Ltd. All rights reserved.

for which band-selective pulses are highly effective due to the reduced chemical shifts dispersion.

The first set of experiments dedicated to this purpose was a ^{13}C -adaptation of the MUSIC approach developed by the Oschkinat group.^{79–81} These experiments, based on the CACON and CANCO sequences, are simple in their implementation and can provide information to aid sequential

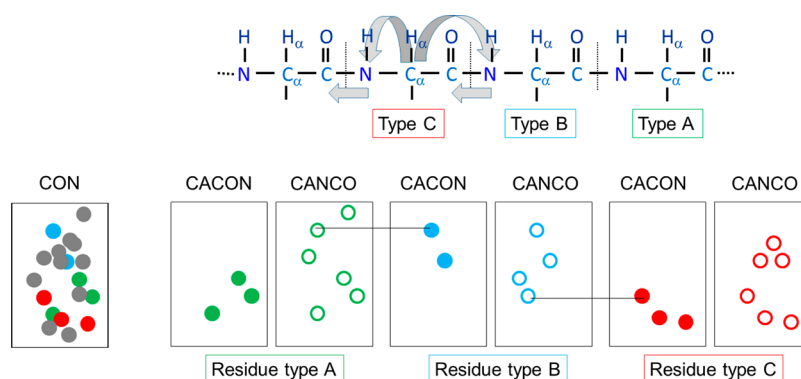


Figure 5. Schematic diagram showing the sequence specific assignment strategy using the amino acid-selection approach for three generic amino acids A, B, and C. The correlations obtained by (CA)CON-based experiments are reported as filled circles, while the additional ones obtained by the (CA)NCO-based experiments are reported as open circles.

assignment⁸² (Figure 5). On the other hand, many experiments should be acquired to discriminate as much as possible the different amino acids and in some cases the sensitivity is limited by the many transfer steps and pulses included for optimal selection.

A second approach, based on the CBCACO sequence, proposed the use of a selection method based on the $^{13}\text{C}^\beta$ topology to distinguish the different amino acid types.⁸³ From a single experiment recorded with a sequence that includes several selective pulses cleverly chosen, it is possible to generate subspectra in which the residue signals can be grouped in six classes according to their topology. The classification is coarse but useful for specific purposes.

Although these amino acid selective experiments produce spectra similar to those obtainable by amino acid selective labeling, they offer the clear advantage that only a single sample is required for all experiments and can be exploited in case a specific subset of residues should be monitored, for example, for simplified chemical shift mapping.

2.2. Multidimensional Experiments (nD, $n > 3$)

The boost in S/N ratio achieved with pulse sequence design and hardware innovations opened the way to the implementation of experiments with many coherence transfer steps, such as multidimensional experiments capable to correlate many diverse heteronuclei, and, when useful, for assignment purposes, reintroduce also the proton dimension.⁸⁴ These experiments, as for the ^1H -detected counterparts, require a very long experimental time and thus it becomes crucial to implement approaches to obtain highly resolved spectra in a reasonable time frame.

The acquisition of a multidimensional NMR experiment is generally performed sampling time-points equally spaced on a Cartesian grid (with time-point, we refer to each repetition of the experiment with different delays for chemical shift evolution). This is dictated by the Nyquist theorem, which states that the interval between the time-points sampled cannot exceed the inversion of the spectral width within which all the expected peaks appear. On the other hand, the resolution of the NMR signals is inversely proportional to the length of the acquisition time. This results in an enormous increase of experimental time as the dimensionality of the experiment is increased, especially when a large number of increments is required to achieve optimal resolution in the indirect dimensions.

In the last decades, alternative sampling approaches known with the collective name of nonuniform sampling (NUS) have been introduced in NMR to reduce experimental time, or achieve higher resolution in the same amount of time or be able to acquire spectra of high dimensionality with appropriate resolution.⁸⁵ NUS eludes the limitation of the conventional acquisition scheme by sampling only a subset of the time points of the Cartesian grid, according to some predetermined sampling scheme. The price to pay is that the data cannot be processed with the usual fast Fourier transform (FFT) but need different strategies for proper treatment of the acquired data and for reconstructing the final data-matrix including some postprocessing to remove spectral artifacts in the final spectrum arising from the so-called “sampling noise” intrinsically related to the method.⁸⁶ Several algorithms for data reconstruction or processing have been proposed and optimized such as maximum entropy (ME),⁷ multidimensional decomposition (MDD),⁸⁷ compressed sensing (CS),⁸⁸ and multidimensional Fourier transform (MFT)⁸⁹ to mention just a few. Some of these reconstruction methods have been implemented in the acquisition programs of the spectrometers to be routinely used, not only for ^{13}C direct detected experiments of course.

The use of NUS in ^{13}C direct detected spectra was tested first with the CANCO experiment on ubiquitin, a small globular protein of 73 amino acids.⁵⁹ Even without the inclusion of ^1H as starting polarization source, it was possible to reduce at about 40% the number of acquired points, reconstructing the final data matrix with the MDD algorithm.⁵⁹ Different approaches were also tested for the 3D CBCACON experiment.⁹⁰

In the following years, a number of additional experiments with higher dimensionality were designed, all tailored for spin system identification and backbone resonance assignment. The most successful example is provided by the suite of experiments developed by Koźmiński and co-workers, which takes advantage of the sparse multidimensional Fourier transform (SMFT) method to handle the spectra.⁹¹ One of the advantages of this method is the possibility to process only a subspace of the full multidimensional spectrum at arbitrary frequency coordinates and to simplify the analysis of multidimensional spectra by displaying only 2D cross-sections computed at some predefined frequencies collected in a lower dimensionality “basis spectrum” (2D/3D), which shares the same dimensions with the higher dimensionality experiment (4D/5D). In this way, the inspection of the 4D/5D is reduced to the analysis of a collection of lower dimensionality maps. The complete set of experiments, particularly useful for the investigation of intrinsically disordered

proteins, has the CON and the CACON experiments as reference spectra to process 4D and 5D data sets, respectively. The choice of these two is based on the fact that they provide the best result in terms of reliability of the sequential backbone assignment taking advantage of the excellent chemical shift dispersion of the resulting spectra.^{92,93}

The multidimensional ^{13}C detected NMR experiments that have been proposed share as common features direct detection of carbonyl carbon nuclei and exploit coherence transfer pathways mediated by the one-bond and two-bonds scalar couplings between backbone heteronuclear spins ($^1J_{\text{C}'\text{N}}$, $^1J_{\text{C}'\text{C}^\alpha}$, $^2J_{\text{C}'\text{C}^\alpha}$, $^1J_{\text{C}'\text{C}^\beta}$). Since these involve coherence transfer between several nuclear spins, the extension to higher dimensions can be easily achieved through frequency labeling of different spins exploited for coherence transfer. The most straightforward approach thus consists of extending the dimensionality of 3D experiments by frequency labeling additional nuclear spins exploited in coherence transfer pathways, as proposed for 4D HCBCACON, 4D HCCCON, 4D HCBCANCO, 4/5D HNCACON and 4/5D HNCANCO, and 3/4D HCANCAO experiments.^{94,95} Additional variants were also proposed that exploit the most informative and well-resolved heteronuclear spins in the indirect dimensions, such as the 5D CACONCAO to focus on backbone assignment, followed by the 5D HC(CC-TOCSY)CACON^{96,97} to discriminate between different amino acid types.

An approach to accelerating the long NMR experiments consists of the implementation of ^1H -start and longitudinal relaxation enhancement that significantly shortens the effective longitudinal recovery time and allows for shorter delays between the consecutive acquisitions (section 1.2.2). This was exploited to propose improved variants of the 4D and 5D NMR experiments based on selective excitation of $^1\text{H}^\text{N}$ nuclei (not touching the aliphatic ones as well as the water resonance),⁹⁶ or on the $^1\text{H}^\text{flip}$ approach that can be exploited also when $^1\text{H}^\alpha$ are used as a starting polarization source.⁹² Several variants of the same core NMR experiment can be used ($^1\text{H}^\text{NBEST}$, $^1\text{H}^\text{N-flip}$, $^1\text{H}^\alpha\text{-flip}$) allowing researchers to perform a sequence specific walk through the backbone by “hopping” from one CON correlation to the neighboring one (CON-CON strategy). These experiments can be performed as 4D experiments (by using a 2D CON as a reference) as well as in the 5D version in which the reference spectrum is the 3D CACON. The proficiency of this CON-CON strategy then led to the design of the analogous experiments based on ^1H direct detection,⁹⁸ providing a complete set to meet different experimental needs.^{99–102} Importantly, this set of spectra is well suited for the implementation of automated protocols for assignment.^{103,104}

Another strategy that can be used to speed up the acquisition of a multidimensional spectrum is to exploit projection spectroscopy, in which a series of projections are acquired rather than the full spectrum.¹⁰⁵ In this way, the analysis of the spectrum is facilitated as it consists of a collection of a series of 2D spectra. Automation of such analysis, called automated projection spectroscopy (APSY),¹⁰⁶ yields a peak list of the full dimensionality spectrum without reconstructing it. The methods for ^{13}C assignment purposes have been demonstrated with α -synuclein¹⁰⁷ and successfully applied also for large intrinsically disordered proteins.¹⁰²

2.3. CON and CACO Fingerprints

Acquisition of 2D NMR spectra that provide well-resolved cross-peaks constitutes a useful tool to achieve a fingerprint of a

protein. 2D HN spectra are often the first ones used for this purpose and sometimes the analogous 2D HC spectra are collected to obtain (and follow) ^{13}C signals. However, information about carbonyl carbon nuclei is not available through ^1H detected 2D spectra. It is thus worth to record experiments that correlate carbonyl carbon resonances with the directly bound heteronuclear spins, C^α and N, exploiting the one bond couplings ($^1J_{\text{C}'\text{C}^\alpha}$ and $^1J_{\text{C}'\text{N}}$), to obtain the CACO and CON spectra. The various solutions for homonuclear decoupling described in section 1.2.1 ensure the removal of the large one-bond coupling in the direct acquisition dimension. The inclusion of an additional building block allows designing the 2D CBCACO and CCCO NMR experiments that collectively provide a suite of 2D NMR experiments that significantly enrich the information content that can be obtained through simple 2D experiments (Figure 6).¹⁰⁸

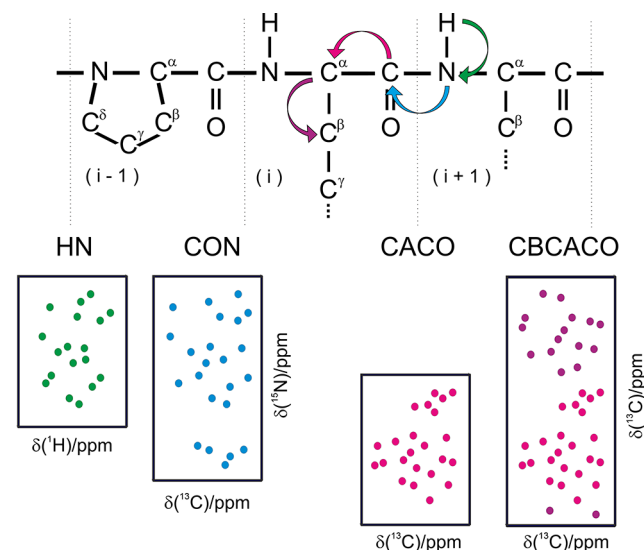


Figure 6. Schematic representation of the 2D spectra based on ^{13}C direct detection that can be used, in conjunction with the 2D HN (^1H – ^{15}N HSQC), to obtain a fingerprint of a protein. These include the 2D CON ($^{13}\text{C}'$ – ^{15}N correlation), 2D CACO ($^{13}\text{C}'$ – $^{13}\text{C}^\alpha$ correlation), 2D CBCACO ($^{13}\text{C}'$ – $^{13}\text{C}^\alpha$ and $^{13}\text{C}'$ – $^{13}\text{C}^\beta$ correlations) spectra. The scalar couplings exploited to detect the correlations in the various spectra are schematically indicated in the top panel ($^1J_{\text{NH}'}$, $^1J_{\text{C}'\text{N}}$, $^1J_{\text{C}'\text{C}^\alpha}$, $^2J_{\text{C}'\text{C}^\alpha}$). In addition to the correlations involving backbone nuclei shown in the figure, correlations involving specific side-chains can be observed (for Asn and Gln in HN and CON 2D spectra and for Asn/Asp and Gln/Glu in CACO and CBCACO 2D spectra). Adapted from ref 112. Copyright 2014, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.

The additional information available for a protein, if compared to that available in a 2D HN, is evident. The large dispersion of heteronuclear chemical shifts and additional nuclear spins that can be monitored (^{13}C) provide a useful tool to investigate proteins in solution through simple 2D experiments. The correlation of nuclear spins belonging to two different amino acids involved in the peptide bond provides an important contribution to resolution in 2D spectra, particularly for intrinsically disordered proteins in which resonances tend to cluster in regions typical for each amino acids type.⁶⁹ Interestingly, the protonless variants are amenable to combination with ^1H detected NMR experiments using multiple

receivers in order to collect two experiments simultaneously as shown through the CON//HN implementation.^{109–111}

^{13}C NMR is not particularly sensitive to solution conditions and is not affected by the detrimental effect of the presence of high salt concentration, pH, or temperature, which cause extensive line broadening of the H^{N} resonances in the NMR spectra.⁸ This holds in particular for the spectra of highly mobile proteins or protein regions, which have peptidic protons largely exposed to the solvent and not engaged in stabilizing interactions like H-bonds. In these cases, upon increasing experimental ionic strength, pH, or temperature, the quality of 2D HN correlation spectra may get worse due to the more efficient hydrogen-exchange mechanism, resulting in increasing number of peaks that are extensively line broadened, while the quality of ^{13}C -detected spectra is maintained or improved. In Figure 7, the comparison of the 2D HN and 2D CON correlation spectra acquired for α -synuclein at pH 7.4 clearly shows that ^{13}C NMR allows recovering the missing information in 2D HN spectra, such as the correlations of Gly, Ser and Thr

residues, that are the first ones to disappear when increasing temperature.

As described in Section 1.2.2, amide protons can also provide a starting polarization source and different experimental variants of CON spectra have been proposed including $^1\text{H}^{\text{N}}$ -start,²⁷ $^1\text{H}^{\text{N}}$ -flip,⁵⁹ $^1\text{H}^{\text{N}}$ BEST,⁶³ and $^1\text{H}^{\text{N}}$ -CP.⁶⁴ These experiments lack correlations involving proline residues and the dependence on solvent exchange processes of amide protons is reintroduced that has an impact on the outcome of the spectra, eventually leading to loss of information when chemical exchange is very efficient. However, if properly tuned, this can even result in a sensitivity increase with respect to the $^1\text{H}^{\text{N}}$ -start variant by exploiting one of the variants with LRE ($^1\text{H}^{\text{N}}$ -flip, $^1\text{H}^{\text{N}}$ BEST) or the $^1\text{H}^{\text{N}}$ -CP one.

2.4. Useful NMR Observables

2.4.1. Chemical Shifts and Heteronuclear Relaxation Rates. The process of sequence-specific resonance assignment provides a list of chemical shifts associated with different nuclei in the protein that represents the key to access atomic resolution information on complex macromolecules. These allow us to associate specific nuclei to each of the cross-peaks detected in multidimensional NMR spectra and follow spectral changes upon changes in the experimental conditions. For example, 2D spectra are often used to acquire protein fingerprints, as described in the previous section, and then follow spectral changes upon addition of potential partners, ligands, metal ions, etc. The additional contribution provided by ^{13}C detected NMR experiments consists of the possibility to access information on solvent-exposed regions of proteins, such as external loops of globular proteins, intrinsically disordered proteins or protein regions, to follow proline residues and to recover information on paramagnetic proteins or highly deuterated large proteins.

The chemical shifts, in particular heteronuclear ones, are also very informative about the secondary structural elements present in a protein.^{114,113} Indeed, on top of contributions deriving from the chemical structure and the presence of specific functional groups, a significant contribution also derives from the local conformation, which in turn is linked to the secondary structural element.^{115,116} Therefore, chemical shifts provide the first source of structural and dynamic information just by comparison with chemical shifts expected for a hypothetical “random-coil” state, representative of the contribution from the local chemical topology. The extraction of this information is by no means a trivial task. Computational tools have been developed to predict reference random-coil chemical shifts to enable the interpretation of experimentally determined chemical shifts in terms of structural and dynamic properties of a protein.^{117–120} As an example, Figure 8 shows the case of one of the “flexible linkers” of CBP, a 207 amino acids long intrinsically disordered region of the multidomain CBP protein (CBP-ID4).^{99,121} Even if largely unstructured, two partially populated helical elements separated by proline rich regions can clearly be identified from chemical shift analysis.

Variants of basic 2D experiments were developed to determine heteronuclear relaxation rates often used to access information about local flexibility.¹²² CON-based experimental variants are available to determine ^{15}N relaxation rates (longitudinal and transverse).^{123,124} The most useful experiments however are probably those focusing on proline ^{15}N spins as they provide information that is not accessible through ^1H detected analogues. The possibility to focus on the ^{15}N region of prolines selectively also allows researchers to focus on a narrow

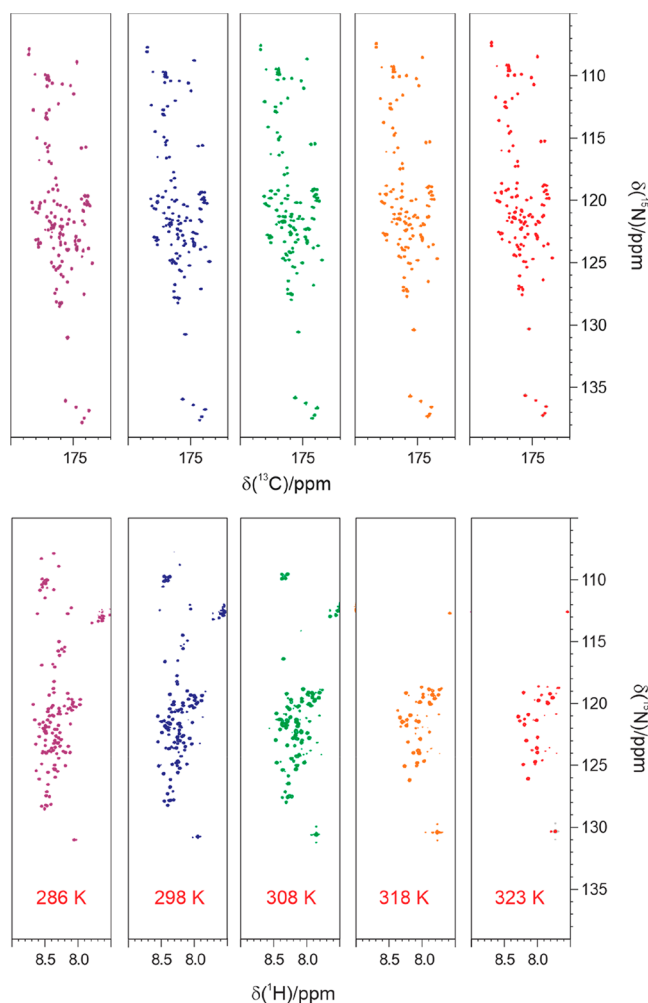


Figure 7. 2D spectra correlating the backbone amide nitrogen either with the directly bound amide proton (^1H – ^{15}N HSQC spectra, lower panels) or with the directly bound carbonyl (^{13}C – ^{15}N CON spectra, upper panels) acquired on α -synuclein at pH 7.4, are shown as a function of increasing temperature, from 286 to 323 K (from left to right). Each spectrum was acquired at 16.4 T with one scan per increment on a 1 mM protein sample in 20 mM sodium phosphate buffer and 200 mM NaCl.

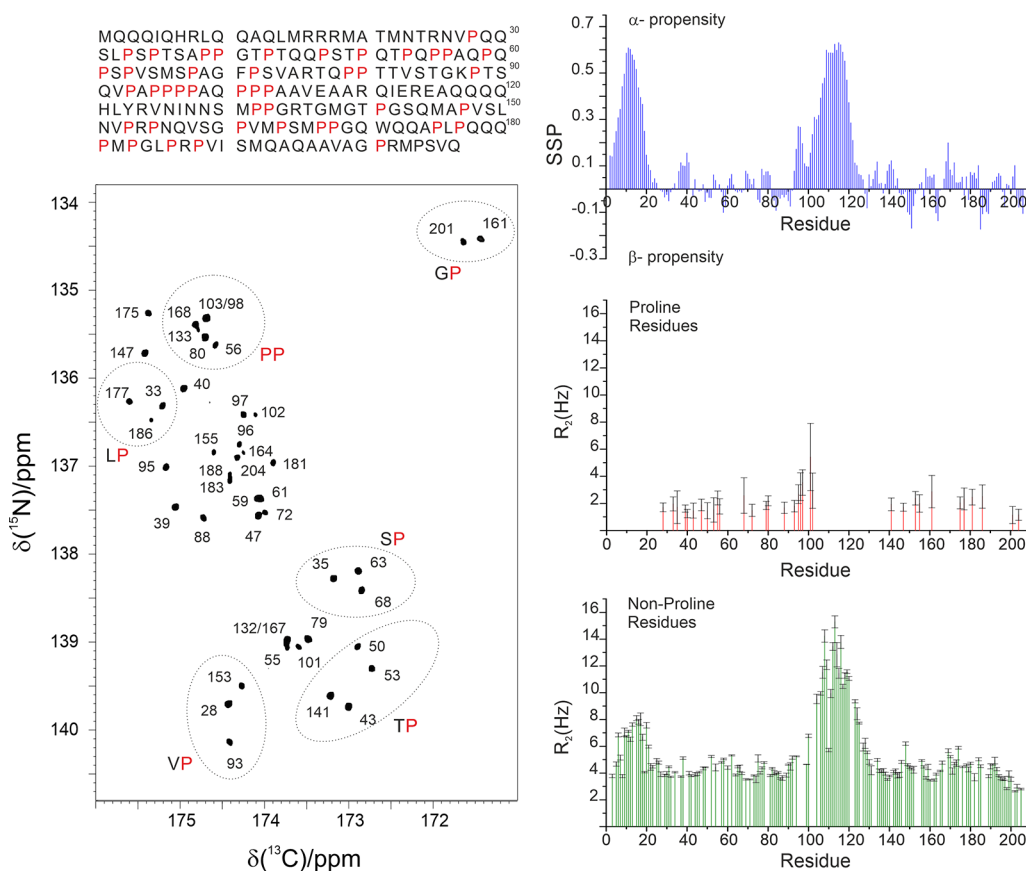


Figure 8. Left: 2D proline-fingerprint spectrum of CBP-ID4. In the 2D spectrum, the signals are numbered according to the proline position in the sequence; C'–N correlations involving identical residue pairs are circled. In the upper part, the primary sequence of the linker is reported, with the proline residues colored in red. Right: On the upper part, the secondary structure propensity (SSP) plot indicates that the protein is largely disordered, with two regions that have a measurable α -helical propensity. In the other two panels, ^{15}N R_2 data measured for proline nitrogen nuclei and ^{15}N R_2 data measured for the nonproline residues recorded at 16.4 T are reported. Adapted from ref 121. Copyright 2018 Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim.

spectral region and achieve an excellent resolution with very few increments, if compared to the whole amide region in proteins, an attractive feature for the investigation of complex proline-rich proteins. It is interesting to note that the absence of the directly bound proton contributes to a reduction of the ^{15}N nuclear relaxation rates, reflected in sharp NMR lines. As an example, the data obtained for CBP-ID4 are shown in Figure 8.

It is worth noting that the CON spectrum in Figure 8 reveals a peculiarity of this experiment when recorded for unstructured proteins, that is the clustering of the signals in groups depending on the previous amino acid type, whose nature determines to a great extent the chemical shift of the $^{13}\text{C}'$ nucleus, being all other contributions to the chemical shift typical of a folded protein largely averaged out.

With increasing pH and temperature, the efficient solvent exchange renders the interpretation of ^{15}N relaxation rates for nonproline amino acids quite complicated as also exchange may contribute to the observed results, reducing possible advantages of CON-based approaches. In these cases, other heteronuclear relaxation rates involving nonexchangeable nuclei such as ^{13}C may result useful and provide complementary information to that accessible through ^{15}N relaxation. The picture is more complex because of the dense network of directly bound carbon nuclei, which implies ^{13}C – ^{13}C interactions in uniformly labeled samples. However, cases of interest in which these effects are mitigated and valuable information can be obtained include the

determination of carbonyl carbon longitudinal and transverse relaxation rates^{123,125} as well as heteronuclear ^1H – ^{13}C NOEs.¹²⁶ CON and CACO variants were used for the purpose and tested on model proteins;^{123,127} these may result useful for the investigation of intrinsically disordered proteins, for example, in all the cases in which chemical exchange is so pronounced to interfere with the determination of ^{15}N relaxation rates.¹²⁸

2.4.2. Chemical Exchange. CON experiments also have interesting properties for the experimental determination of exchange processes with the solvent, an observable used since the early days of NMR to discriminate between amide protons easily accessible to the solvent from those protected into globular cores or involved in tight hydrogen bonds.¹²⁹

In fact, the experiment correlates two heteronuclear spins not directly involved in chemical exchange with the solvent and it constitutes a useful tool to recover information also in conditions in which amide protons are broadened beyond detection, as described above (Figure 7).^{63,72} However, simple modifications of the basic pulse sequences can be designed to reintroduce a dependence on exchange processes with the solvent although in a more indirect way, allowing to monitor the process in a less perturbative way.

The ^1H CON variant^{126,130} of course constitutes the first obvious one to reintroduce effects deriving from solvent exchange through the starting polarization source used in the experimental variant.¹¹² Exchange indeed influences the

observability of the signals, even if to lower extent with respect to ^1H detected experiments. Exchange is also responsible for a pronounced enhancement of the recovery of amide proton polarization to equilibrium provided amide protons are selectively perturbed with respect to those of water.^{60,131,132} Therefore, in a certain range of exchange rates, this effect can become a favorable feature for the detection of CON spectra as shown through the example of the $^1\text{H}^{\text{NBEST}}$ CON acquired on α -synuclein in less than 1 min⁶³ (without introducing any delay between the end of one transient and the start of the next one thanks to the fast recovery of amide proton magnetization in the experimental conditions used). This is actually a qualitative observation of exchange which however could be quantified by exploiting similar approaches to those proposed by Schanda and Brutscher.¹³¹ Selective manipulation of the water resonance to highlight only residues whose amide proton senses this perturbation^{133,134} constitutes another widely used strategy that was implemented prior to $^1\text{H}^{\text{N}}$ -start CON variants to take advantage of the nice chemical shift dispersion typical of the CON reading scheme.¹³⁵

However, probably the most interesting approaches are those that exploit the most simple CON variant, the ^{13}C -start one, to indirectly detect effects of exchange by monitoring its effect through ^{15}N spin coherences or spin orders. In an elegant paper, Atreya and co-workers proposed to monitor the effect of isotopic shift induced by deuterium in solutions constituted by 50% H_2O and 50% D_2O , and then measure exchange effects between the two resonances.¹³⁶ Effects of chemical exchange can thus be determined by observing the direct exchange of polarization between the two isotopomers as well as by the effect on the scalar coupling of ^{15}N with ^1H , coupling that can be quenched by fast exchange processes. A similar approach was also proposed to investigate exchange rates of arginine side-chains.¹³⁷

Finally, the CON reading scheme is particularly well suited to implement variants in which exchange is monitored through subtle effects on coherences or spin orders involving ^{15}N . For example, the determination of ^{15}N relaxation under CPMG, in the presence of efficient exchange processes of amide protons with the solvent, is influenced by whether ^1H RF decoupling is applied or not. This difference, properly modeled, can be used to access information on exchange rates also in cases in which direct observation of the ^1H resonance is precluded. This approach, initially designed to access information on exchange rates of selected amino acid side-chains^{138–140} was then implemented also for the study of amide protons in proteins in ^1H and ^{13}C detected experimental variants.^{141–143}

Along similar lines, a three spin order operator can be created without perturbing the water resonance and then allowed to evolve in a free evolution period.¹⁰⁸ In the presence of efficient solvent exchange processes, this becomes the major determinant of the disappearance of the three spin order by “decorrelation”, as initially described by Skrynnikov and Ernst.¹⁴⁴ The CON variant allows implementation of this idea in a very clean way since the water resonance is not perturbed at all in the experiment and the 2D CON reading scheme allows to profit by the excellent resolution also in intrinsically disordered proteins.¹⁰⁸

Finally, solvent accessible protein backbones can be “illuminated” by hyperpolarized HDO produced using dissolution-dynamic nuclear polarization (D-DNP).^{145–149} The hyperpolarized solvent provides a polarization reservoir, enhancing the ^1H signals of sites undergoing chemical-exchange with the hyperpolarized solvent itself; implementation of a 2D

CON reading scheme provides a nice resolution and reduced hurdles in dealing with the resonance of hyperpolarized HDO when performing ^1H detection.¹⁴⁹

2.4.3. Residual Dipolar Couplings. Other observables deriving from dipole–dipole interactions include residual dipolar couplings (RDCs) resulting, as the term itself tells, from noncomplete averaging of dipolar interactions in solution.^{150–152} This can be originated by the natural magnetic anisotropy of the molecule under investigation that induces a partial degree of alignment when immersed in high magnetic fields or by external agents. Residual dipolar couplings are generally determined measuring changes in signals splitting resulting from a well resolved scalar coupling upon induction of a partial degree of alignment of the molecule in solution. The resolved $^{13}\text{C}'$ – $^{13}\text{C}^\alpha$ scalar couplings are thus an obvious candidate,¹⁵³ easily accessible from carbonyl detected experiments, to determine RDCs. Experiments to measure a variety of different RDCs, including several that involve proton nuclei,^{123,154,155} were developed to enable the determination of RDCs also in cases in which methods based on ^1H direct detection fails to provide information.^{123,156}

2.4.4. What about Homonuclear Nuclear Overhauser Effects (NOE)? By analogy to observables that can be detected through ^1H detected experiments, an important observable that comes in mind is the homonuclear NOE, one of the most robust sources of internuclear distances useful for solution structure determination. However, when moving from ^1H to ^{13}C , homonuclear NOEs are drastically reduced due to the lower gyromagnetic ratio of ^{13}C .^{157,158} Ideally, long-range ^{13}C – ^{13}C NOEs would be very useful observables that would contribute to the investigation of very large systems. However, since they scale with $1/r^6$, they become tiny effects that add to the much stronger interactions between directly bound carbon atoms (from 1 to 3 Å the NOE decreases by 3 orders of magnitude), so small that it is really difficult to measure them experimentally in uniformly labeled protein samples with the current hardware possibilities. On the other hand, ^{13}C – ^{13}C NOEs can be easily detected between directly bound carbon nuclei and spin diffusion becomes a very efficient effect within these networks of directly bound carbon atoms.^{9,19,30,52,158} This information is not so useful to recover unknown structural information but contains nevertheless information that can be used for assignment purposes.

3. BIOMOLECULAR APPLICATIONS

3.1. Focus on Amino Acid Side-Chains

The possibility to focus the NMR experiments on specific amino acids is a useful approach for proteins' investigation. This can be used to aid the sequence specific assignment of macromolecules, as described in section 2.1, but it can also be exploited to address specific biological questions.

Amino acids' side-chains are generally assigned using ^1H -detected experiments, but there are many cases in which the resonances due to the nuclei of a long chain are missing or the signals vanish upon interaction due to conformational exchange. In all these cases, ^{13}C -detected experiments can be of use. In addition, with simple 2D experiments is possible to focus on the key correlations relevant to monitor the behavior of a side chain upon changes in its environment, either due to changes in the solution conditions or to the interaction with a partner or to binding of a small molecule/metal ion. Since different kinds of nuclei are affected to a different extent by the various

interactions, the analysis of carbon nuclei chemical shifts can provide additional information on conformational changes, variations in solvent exposure or in hydrogen bond pairing, etc. providing useful information to describe the processes of interest.

3.1.1. Monitoring Side-Chains through C'-Detected Experiments. The set composed of CACO, CBCACO, and CCCO experiments provides an excellent tool to monitor resonances of virtually every ^{13}C nuclear spin of aliphatic side-chains by correlating them to the backbone carbonyl of each amino acid. Moreover, it also provides information on the terminal carbon spins of side-chains of residues that contain the carboxylate (aspartate and glutamate) or the carbonyl (asparagine and glutamine) moiety since their correlations can be detected in a clean region of the spectra. The one-bond correlation between the terminal carbon nuclei and the neighboring one can be collected in 2D CACO spectra; the cross-peaks observed for the side-chains of these four amino acid types fall in specific spectral regions, as illustrated in Figure 9.

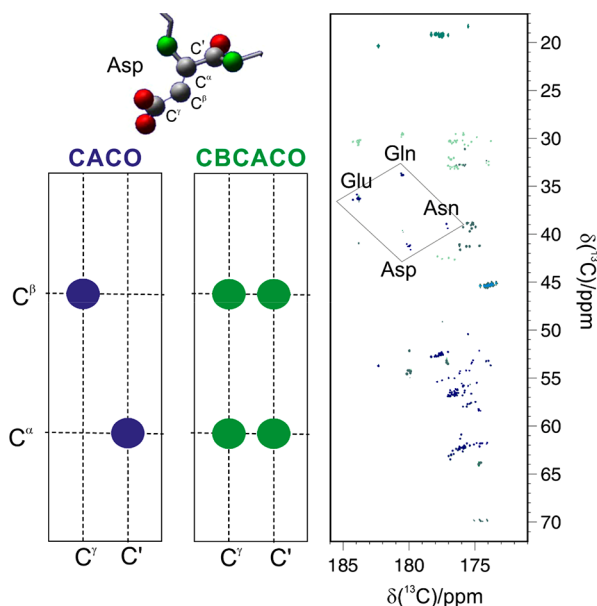


Figure 9. Superposition of the CACO (blue contours) and the CBCACO (green contours) spectra recorded on a 600 μM sample of α -synuclein in 20 mM Tris-Cl buffer, pH 7.4, and 310 K at 16.4 T. On the left, the correlations observable in the two spectra for an Asp residue, sketched above, are schematically reported.

Comparison with a 2D CBCACO allows detection of the neighboring aliphatic carbon spin, providing the information to assign in a sequence-specific manner the ^{13}C terminal resonance of aspartate and asparagine residues. The 2D CCCO is needed to complete the assignment and also identify in a sequence-specific manner the ^{13}C resonances of glutamate and glutamine residues.

The CON experiment provides instead the intraresidue $\text{N}^{\delta}-\text{C}'$ for asparagine and $\text{N}^{\epsilon}-\text{C}^{\delta}$ for glutamine residues. A combination of CON and CACO/CBCACO has been proposed to remove from the latter spectra the correlations due to the carbonyl carbon nuclei bound to nitrogen, producing maps where only the cross-peaks of aspartate and glutamate nuclei are present.¹⁵⁹ Since the coupling between the backbone carbon and the carboxylate carbon of the aspartate side chain has a strong conformational dependence, in some favorable cases it

is even possible to observe a splitting of the signals, whose magnitude can provide information on the chain's conformation.

Carbonyl and carboxylate functional groups of amino acid side-chains are seldom assigned despite their key role in many biological processes and their investigation can be very instructive. A recent example in this context is provided by the investigation of α -synuclein when subject to concentration jumps of calcium metal ions, a process associated with the transmission of nervous signals. Negatively charged side-chains of aspartate and glutamate residues are expected to be among the first candidates to interact with positively charged ions such as Ca^{2+} , and it is interesting to access direct information on these side-chain functional groups to identify whether specific regions of the polypeptide are affected to different extents by the interaction. The set of 2D ^{13}C detected NMR experiments confirmed that the final part of the protein, rich in negatively charged amino acids, is the one sensing Ca^{2+} concentration increase, as also previously described through ^1H detected experiments.^{160–162} These experiments, by monitoring also carboxylate groups of glutamate and aspartate residues, revealed that even within a disordered protein in which all of them are expected to be exposed to the solvent, only a subset of them are initially perturbed by the addition of Ca^{2+} (Figure 10A,C). Interestingly, among the most affected residues there are aspartate and glutamate residues separated by a proline, in a peculiar motif (DPD and EPE motif). The presence of a proline residue in between two residues featuring carboxylate groups could be a strategy to adapt the local conformation for Ca^{2+} interaction. It is noteworthy that these two proline residues were the most perturbed upon the addition of Ca^{2+} , an observation that could be easily achieved through the 2D CON (Figure 10B). The 2D CACO and CON spectra thus provide a useful tool to focus on the metal ion coordination sphere and more generally to investigate interactions of negatively charged residues with complementary charged molecules in the cell.

3.1.2. Positively Charged Amino Acids. Lysine and arginine side-chains are crucial for driving protein–protein interaction and for forming intramolecular and intermolecular salt bridges. Interactions and salt bridges often involve acidic protein side-chains such as those of aspartate or glutamate. Positively charged residues are relevant also to establish interactions with charged phosphodiester moieties of nucleic acids.

The first example of exploitation of CON and CBCACO type experiments for investigating protein–protein interaction dates back to 2006, in the study of the metal-mediated complex formed in the presence of Cu(I) between two copper chaperons, Atx1 and Ccc2a.²⁹ These two proteins have a very specific role in copper trafficking and the solution structure of the complex obtained in the presence of Cu(I) was solved by NMR.¹⁶³ However, the H^{N} signals of some of the residues considered important to characterize the metal-binding region were missing due to exchange with the solvent protons, preventing clear-cut characterization of their role. The correlations of these residues were instead present in the CON experiment and allowed the characterization of the interacting interfaces of the two oppositely charged proteins. In addition, through the comparative analysis of the CON and the CBCACO, it was possible to establish the direct involvement of selected carboxylate moieties in electrostatic interactions as well as in H-bonds necessary to stabilize the conformation of an otherwise very mobile region of the polypeptide.

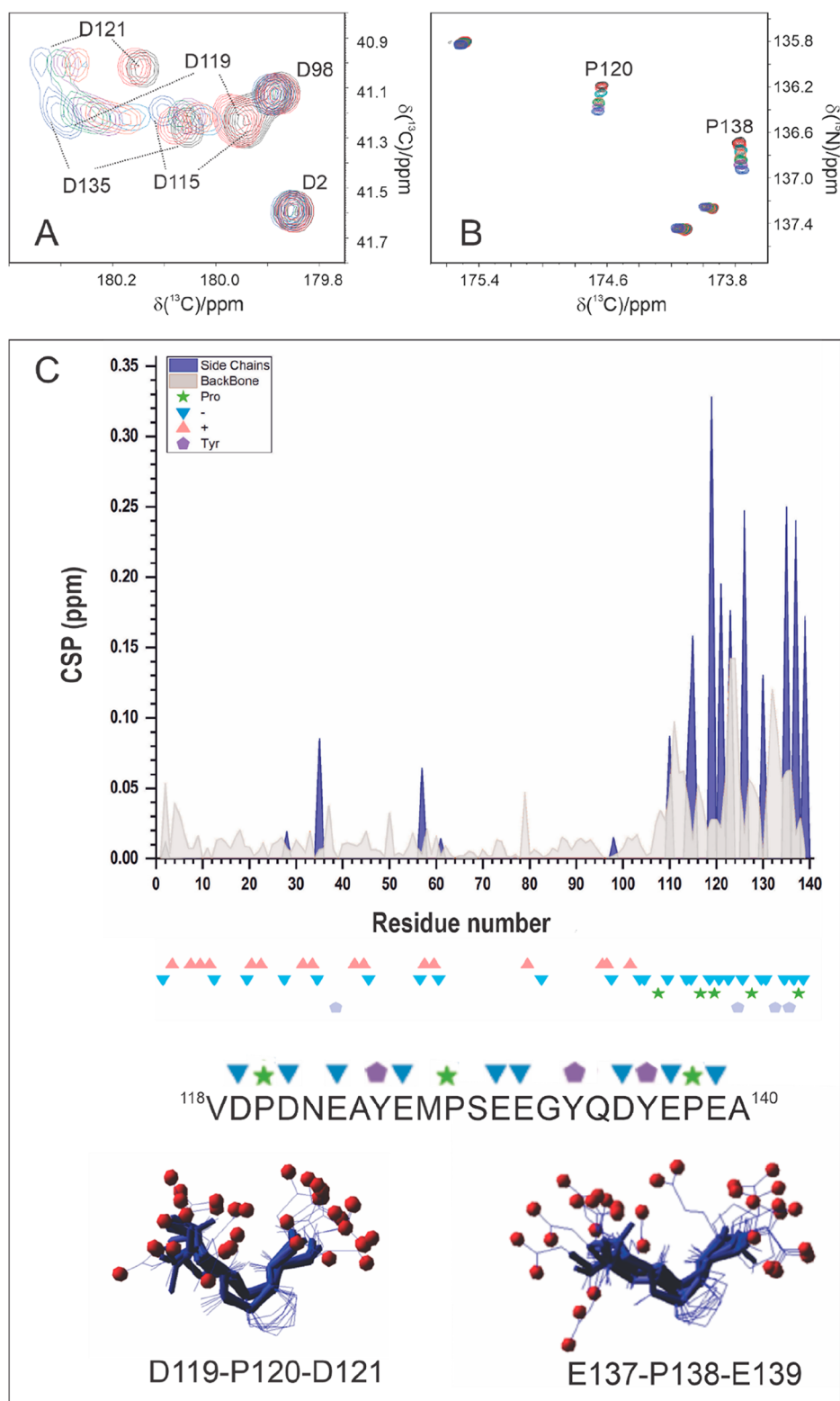


Figure 10. (A) Expansion of CBCACO spectra recorded on α -synuclein at increasing concentration of Ca^{2+} showing the chemical shift perturbation for the most affected Asp residues during the titration. (B) Expansion of CON spectra recorded on α -synuclein at increasing concentration of Ca^{2+} showing the chemical shift perturbation for two proline residues (P120 and P138). (C) Comparison of chemical shift perturbations (CSP) of side chain carboxylate/carbonyl carbon chemical shifts (blue) with backbone carbonyl carbon chemical shifts (red) determined through 2D-CACO and 2D-CON spectra ($\text{CSP} = |\Delta(\delta(^{13}\text{C}))|$). Backbone CSP values are smaller in magnitude with respect to those of side-chains and not necessarily maximal for Asp/Glu/Asn/Gln amino acids, reflecting a more indirect effect experienced by backbone nuclear spins upon interaction with calcium ions. On the bottom the portion of primary sequence affected by metal ion binding is reported, with two sketches of the three-amino acid motifs most affected. Adapted from ref 108. Copyright 2020 Wiley-VCH GmbH.

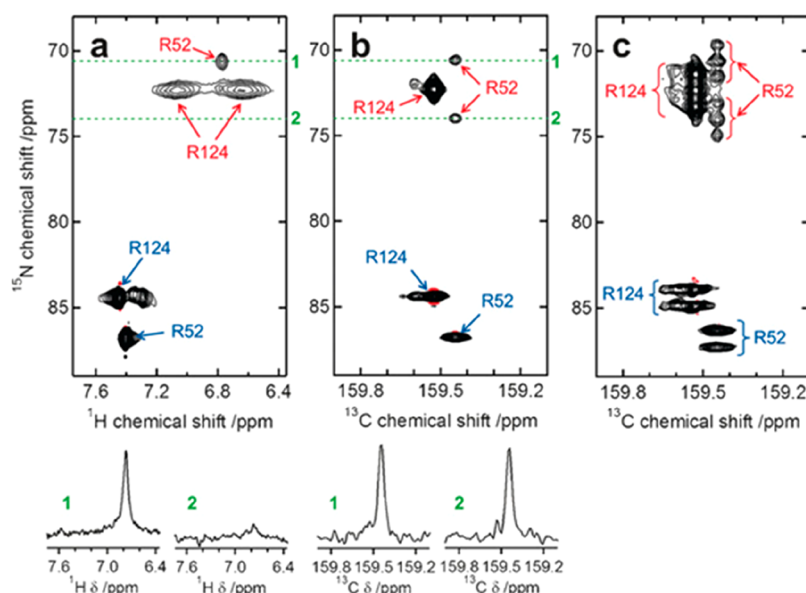


Figure 11. (a) ^1H - ^{15}N HSQC spectrum of 1 mM PYP in 5 mM potassium phosphate at pH 6.2. The $^{15}\text{N}^{\eta}$ - $^1\text{H}^{\eta}$ and $^{15}\text{N}^{\epsilon}$ - $^1\text{H}^{\epsilon}$ correlations are indicated in red and blue, respectively. (b) ^1H -decoupled and (c) ^1H -coupled $^{15}\text{N}^{\eta/\epsilon}$ - $^{13}\text{C}^{\zeta}$ correlation spectra of 3 mM PYP in 5 mM potassium phosphate at pH 6.5. The $^{15}\text{N}^{\eta}$ - $^{13}\text{C}^{\zeta}$ and $^{15}\text{N}^{\epsilon}$ - $^{13}\text{C}^{\zeta}$ correlations are indicated in red and blue, respectively. To avoid additional signals due to the presence of ^2H isotopomers, D_2O for field lock was added externally to a separate compartment of the NMR tube. The number of scans is as follows: (a) 16, (b) 128, and (c) 192. In (a) and (b), 1D traces from the 2D spectra, indicated by the broken lines, are shown. Reproduced with permission from ref 31. Copyright 2017 Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim.

Since Atx1 has several lysine residues considered important for driving binding, an experiment to detect the C^{δ} - N^{ϵ} was designed, which allowed the identification of small but meaningful chemical shift variations upon complex formation allowing to delineate the interaction surface. One of the lysine residues (K65) experienced in the apo-form a very peculiar downfield shift of 6 ppm with respect to the usual value, suggesting its involvement in an intramolecular H-bond.²⁹

Novel experiments were also designed for monitoring arginine side-chain. The guanidinium group of this amino acid is often involved in salt-bridges, H-bonds and in cation- π interactions with aromatic side-chains and with nucleic acid molecules. It is thus very informative to be able to determine its ionization state and its dynamics; however, ^1H -detected experiments are often not effective due to line broadening and signal crowding and ^{13}C detection has been exploited by several groups. For example, Hansen and co-workers proposed an experiment to obtain a $^{13}\text{C}^{\zeta}$ - $^{15}\text{N}^{\epsilon}$ co-relation spectra avoiding transfer from $^{13}\text{C}^{\zeta}$ to $^{15}\text{N}^{\eta}$. Such experiment proved useful for ^{15}N relaxation measurements and quantification of the squared order parameter, S^2 , that reports on the motions of the arginine side-chain on some model proteins and on a 42 kDa enzyme, the human histone deacetylase 8 (HDAC8).¹⁶⁴ With a similar approach it is possible to probe the rotational dynamics around the C^{ζ} - N^{ϵ} bond.¹⁶⁵ This experimental scheme found an interesting application on a mutant of T4 lysozyme, a 19 kDa protein containing 13 arginine residues with different exchange regimes: some are in fast exchange; five of them have ^{15}N resolved resonances, suggesting some restriction in the rotation of the side-chain. A variant to include in the ^{13}C detected scheme the so-called divided-evolution approach¹⁶⁶ was designed and successfully demonstrated that two arginine residues have an exchange regime consistent with their involvement in a hydrogen-bonding network. An additional variant of the experiment including selective pulses proved useful to establish

the exchange regime for two slower rotating guanidinium groups, which belong to residues shown to be involved in an interaction network together with a tryptophan side-chain in one case and in strong ionic bidentate hydrogen bonds in the other. Exploiting double-quantum experiments it is also possible to obtain $^{13}\text{C}^{\zeta}$ - $^{15}\text{N}^{\epsilon}$ correlation spectra with reduced exchange broadening, obtaining sharper lines. These can report about chemical shift perturbations occurring due to the interaction of the side-chain such as the measurement of very small isotopic effects when protons are substituted with deuterium to highlight the involvement of a specific group in stabilizing H-bonds or salt bridges.¹⁶⁷

Another approach, proposed by Mulder and co-workers,³¹ provides the correlation of the $^{13}\text{C}^{\zeta}$ with both $^{15}\text{N}^{\epsilon}$ and $^{15}\text{N}^{\eta}$, and enables the determination of the protonation pattern of the nitrogen nuclei. They tested the experiment, which based its efficiency on cross-polarization transfers, on three proteins differing in number of arginine residues present and in their protonation state. In the case of the photoactive yellow protein (PYP), which presents a particularly relevant arginine for protein function (R52), they demonstrated the possibility to detect all the protons bound to nitrogen nuclei, one H^{ϵ} and four H^{η} , contrary to previous crystallographic studies (Figure 11).³¹

3.1.3. Hydrophobic Amino Acids. The use of ^{13}C direct detected NMR experiments has been proposed also to map methyl groups, often exploited for protein structure and dynamics, particularly in large proteins. In ^1H detected experiments often the signals of the CH_3 groups fall in regions where other signals resonate, complicating the extraction of selective information. A solution is to recur to selective labeling of the protein;⁵¹ otherwise, one can exploit a spectroscopic filter to remove all the undesired cross-peaks, as in the ^{13}C -Methyl COSY experiment proposed by Atreya and co-workers.¹⁶⁸ In this experiment they exploited the fact that the $^{13}\text{CH}_3$ signal of the methyl group located at the terminal end of an amino acid chain has only a single J-coupling to its directly attached ^{13}C

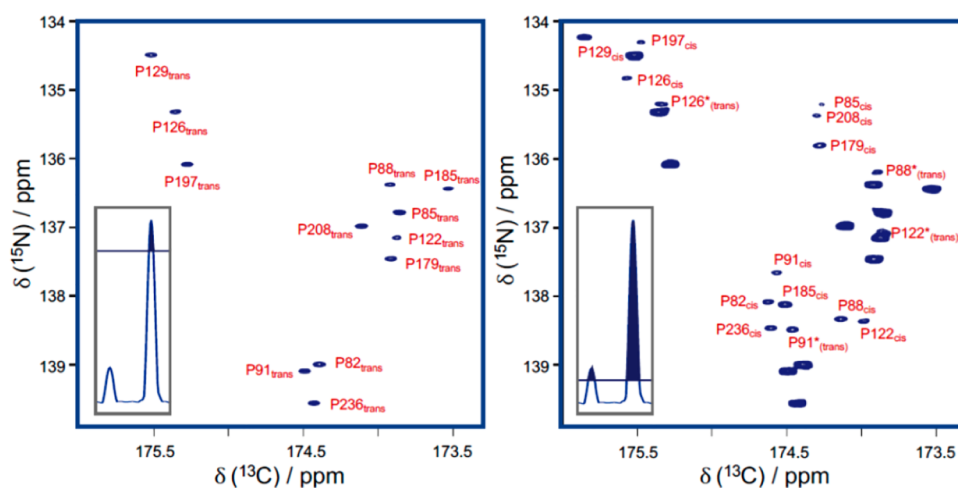


Figure 12. Proline region in the 2D CON spectrum. The main forms of prolines are observed (left) and minor forms appear (right) when the contour levels are lowered. In the inset, the level of the 2D slice and the relative intensity of major and minor forms are reported. Sequence specific assignment of cross-peaks is reported in the spectra; tentative assignments are indicated with an asterisk. Reprinted in part from ref 77. Copyright 2019 Elsevier Ltd. All rights reserved.

nucleus. Using this selection method they demonstrated that the $^{13}\text{CH}_3$ correlation peaks for the various amino acids can be clearly separated into distinct spectral regions. Such a filtering building block can be incorporated in a 3D experiment wherein the additional dimensions can be utilized to provide intra-amino acid or long-range correlations. This approach is particularly well suited to study highly flexible proteins, for which the overlap is dramatic, but can be of use also for chemical shift mapping of large proteins or to obtain long-range information accurately quantifying paramagnetic relaxation enhancement.

Aromatic residues are often of great interest but not trivial to be investigated in detail, in particular in very large globular proteins or in intrinsically disordered ones. They often form the core of protein folds and recently they were shown to play a key role in the formation of membrane-less organelles through the process of liquid–liquid phase separation. The quaternary carbon (C^γ) linking aromatic rings to C^β is not straightforward to be detected through ^1H detected experiments while it gives rise to well resolved cross peaks with C^β that fall in a very clean spectral region of simple 2D ^{13}C – ^{13}C COSY or TOCSY or NOESY spectra. These 2D spectra were used to identify metal ion ligands in paramagnetic proteins by comparison with the spectra obtained with a diamagnetic analogue.^{44,45} A special sequence was also designed to remove the two large one bond scalar couplings influencing the C^β ($^1J_{C^\beta C^\gamma}$ and $^1J_{C^\beta C^\alpha}$) to improve the resolution of 2D spectra that report the C^β – C^γ correlation.³⁸

3.1.4. The Special Case of Proline. The cyclic nature of the proline side-chain and the lack of the backbone peptidic H^N atom sets it apart from all other natural amino acids. Well studied in globular proteins, proline residues are often very abundant in regions of proteins that would not lead to diffraction in X-ray crystallographic studies and on this basis were classified among “disorder promoting” amino acids in polypeptides. However, due to the constrained side-chain of proline residues, they confer local rigidity. The steric hindrance given by the cyclic structure of proline lowers the energy barrier present between the *cis* and the *trans* conformation of the peptide bond involving the nitrogen of this residue. When proline residues are part of globular folds the local structure can favor the *cis* conformation, still allowing interconversion to the *trans* conformer through minor changes, inducing a bend in the polypeptide and therefore

protein folding.¹⁶⁹ In highly flexible protein regions, while the occurrence of *cis* conformation is usually around 0.5% for all the residues in a polypeptide, for proline this can be much higher even in the absence of constraints from a stable 3D structure.^{170,171}

The possibility to directly monitor correlations involving proline residues in clean regions of CON spectra stimulated the design of a series of experiments to investigate the properties of these residues in globular as well as disordered proteins. One of the first experiments proposed was the (H)CCCdNCO experiment,⁸² which correlates the intrareidue ^{15}N and the $^{13}\text{C}'$ of the preceding residue with all the ^{13}C resonances of a proline ring, providing information to identify the isomerization of the peptide bond. An alternative strategy to select in a clean way resonances involving proline nitrogen resonances consists of exploiting the evolution of the $^1J_{\text{NH}}$ scalar coupling to suppress resonances of all amino acids except proline.⁸² Finally, when the ^{15}N chemical shifts of prolines are sufficiently isolated from those of all other amino acids, such as for highly flexible disordered proteins (^{15}N chemical shift range centered at 137 ppm), the use of a band selective pulse on ^{15}N spins of proline residues allows to detect proline fingerprint with high resolution with a reduced spectral width and to design Pro-selective experiments to focus the assignment on proline residues. This experimental scheme allows tailoring all the experiments to measure observables such as ^{15}N transverse and longitudinal relaxation rates.¹²¹ Lately, several other experiments were designed following a similar approach, which provides the sequence-specific assignment of proline residues to complement any assignment obtained either with ^{13}C or ^1H NMR (or, more likely, a combination of the two).¹⁷²

These experiments provided a wealth of information that is beginning to reveal possible functional roles of proline in highly flexible proteins. The investigation of flexible linkers of CBP, which constitute more than half of the primary sequence of this complex multidomain coactivator of transcription, revealed that often proline residues are clustered in specific regions of the primary sequence, indicating that these residues can be exploited to maintain elongation of the backbone, such as in the case of CBP-ID3,¹⁰¹ or act as helix-breakers, separating regions with high helical propensities, like in the case of CBP-ID4.⁹⁹ Peculiar

motifs with several proline residues in a row (PP, PPP, PPPP) could be detected and characterized with the novel experimental tools developed.¹⁷² Other motifs in which prolines are flanked by specific amino acids are also emerging, as described above for the interaction of α -synuclein with calcium ions through the DPD or EPE motifs. Similar motifs were also identified in different proteins, such as viral proteins^{102,173–175} as well as in WASP-interacting protein (WIP), an intrinsically disordered polypeptide with a key role in actin polymerization in activated T cells.^{176,177} Another interesting example in this context is provided by aromatic-proline pairs as recently monitored in different proteins^{77,178,179} and illustrated below through the example of quail osteopontin.⁷⁷ The ^{13}C detected experiments used to complete resonance assignment revealed in addition to the major form, also a subset of peaks with lower intensity that could be clearly identified in the clean region of CON spectra reporting the correlation of proline nitrogen (Figure 12). This set of peaks could be assigned in a sequence specific manner and the ^{13}C chemical shifts of proline rings confirmed the presence of *cis/trans* isomers and provided their relative population. All major peaks have a secondary form and for some of them, which appear clustered in specific regions of the primary sequence, even additional forms due to various combinations of *cis* and *trans* conformation could be identified. The population of *cis* isomers results more abundant for proline residues near aromatic residues. In this protein the rigidity of the closed proline ring is exploited to establish π – π interactions with the aromatic side-chain of neighboring amino acids that imparts a compact state to the otherwise very flexible tertiary structure.⁷⁸

An equilibrium between the *cis* and *trans* isomers of proline peptide bonds was shown to be relevant in case of RNA polymerase II C-terminal domain, an intrinsically disordered region of the protein.¹⁷⁸ Another interesting observation was reported for the BMAL1 transactivation domain that is involved in circadian clock modulation,¹⁷⁹ showing a direct link with function mediated by an aromatic-proline pair within an intrinsically disordered region. The ^{13}C detected experiments are thus revealing information about novel sequence motifs that are not yet well described.

3.2. Overcoming Limitations Due to ^1H Transverse Relaxation: Paramagnetic and Large Proteins

Paramagnetic centers are known to provide a wealth of additional contributions to nuclear spins in their surroundings, including contributions to nuclear relaxation, to chemical shifts as well as to partial orientation in high magnetic fields.¹⁸⁰ The protonless NMR experiments provide, by default, additional unique information with respect to ^1H detected experiments. They enable monitoring signals also at distances from the paramagnetic center in which the effects sensed by protons are so large that resonances are broadened beyond detection, recovering information that would be lost when using as reading schemes ^1H detected experiments (such as ^1H – ^{15}N or ^1H – ^{13}C HSQCs).^{43–47,181–185} A variety of observables that contain structural information can thus be determined. Several approaches were proposed to quantify paramagnetic contributions to relaxation (either longitudinal or transverse, or both, depending on the type of paramagnetic center) and to exploit them to access distance information between nuclei and paramagnetic centers themselves for structure determination (Figure 13).¹³

Paramagnetic contributions to chemical shifts also contain a wealth of structural information and protonless NMR experi-

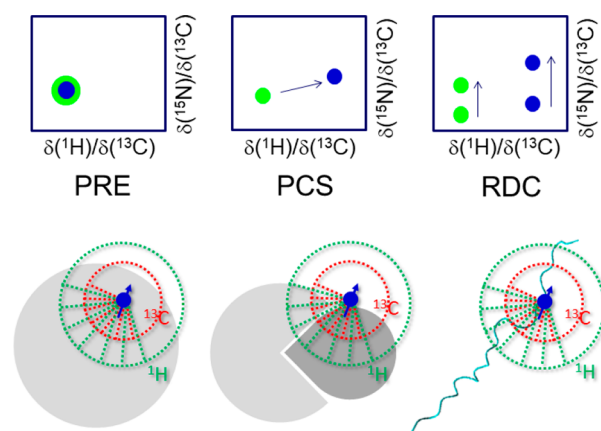


Figure 13. Presence of a paramagnetic center in a molecule affects its NMR spectrum (blue dots) with respect to the one of a diamagnetic analogue (green dots). The change in relaxation rates, in chemical shifts and signals splitting provide restraints called paramagnetic relaxation enhancement (PRE), pseudocontact shift (PCS), or contact shift (CS) and paramagnetism-induced residual dipolar coupling (RDC), respectively. Paramagnetic effects on nuclear relaxation and chemical shifts are generally smaller for ^{13}C with respect to ^1H , and protonless NMR allows to measure observables closer to the paramagnetic center with respect to proton NMR. The approach can be applied for structural purposes to establish the reciprocal orientation of two interacting partners or to establish long-range crosstalk within a protein.

ments enable detection of these effects also in regions of macromolecules in which protons are broadened beyond detection.^{20,25,47,182,183} The same holds for residual dipolar couplings that can be induced when a significant anisotropy derives from the paramagnetic metal ion itself.^{123,156} Therefore, the set of protonless experiments can provide excellent information for the structure determination of paramagnetic proteins.

After initial pioneering work showing the potential of the approach, the methods were applied to the study of type II Cu(II) containing proteins such as CopC⁴⁵ as well as to oxidized Cu(II)/Zn(II) superoxide dismutase¹³ to demonstrate that a 3D structure can be obtained also for proteins containing paramagnetic centers that broaden proton resonances beyond detection in large spheres around the metal ion itself. ^{13}C detected experiments were also used to investigate binding of Cu(II) to αB -Crystallin¹⁸⁶ as well as to follow the uptake of iron ions by ferritin.¹⁸⁷ A similar strategy was used to investigate iron-containing proteins^{182,188,189} and calcium binding proteins where the metal ion was replaced with lanthanide ions.⁴⁷ In the latter case, the lanthanide ion can be selected to modulate the magnitude and type of paramagnetic effects.¹⁹⁰

Paramagnetic tags are also often engineered on purpose to study intermolecular interactions as well as possible residual structure in disordered proteins.^{191–197} The additional interactions deriving from the presence of a paramagnetic center are generally larger in magnitude with respect to internuclear ones on the grounds of the larger electronic magnetic moment. For this reason, they are well suited to reveal long-range information as well as information about short-range distances even if only partially populated in the ensemble of conformers describing disordered systems. In this context protonless NMR experiments provide valuable complementary information to that available through ^1H detection.^{198–200} Actually, different variants of CON experiments were used to modulate the extent of paramagnetic effects that contribute to cross-peaks intensity

by exploiting ^1H -start variants ($^1\text{H}^\alpha$, $^1\text{H}^\text{N}$) and comparing them to effects observed through ^{13}C -start variants, still exploiting the CON reading scheme to take advantage of its excellent resolution.¹⁹⁹ Indeed while ^{13}C -start variants only sense paramagnetic relaxation enhancements through ^{13}C nuclei, ^1H -start variants reintroduce a dependence also on effects sensed by protons, offering an additional way to modulate the extent to which paramagnetic effects are sensed.

This approach was recently demonstrated for a highly flexible disordered system such as osteopontin.¹⁹⁹ Upon comparing C' and $\text{H}^\alpha\text{CON}$ -derived long-range PREs in the protein grafted with MSTL, it appears that both experiments highlight the same regions of the protein (indicated by arrows in Figure 14), with

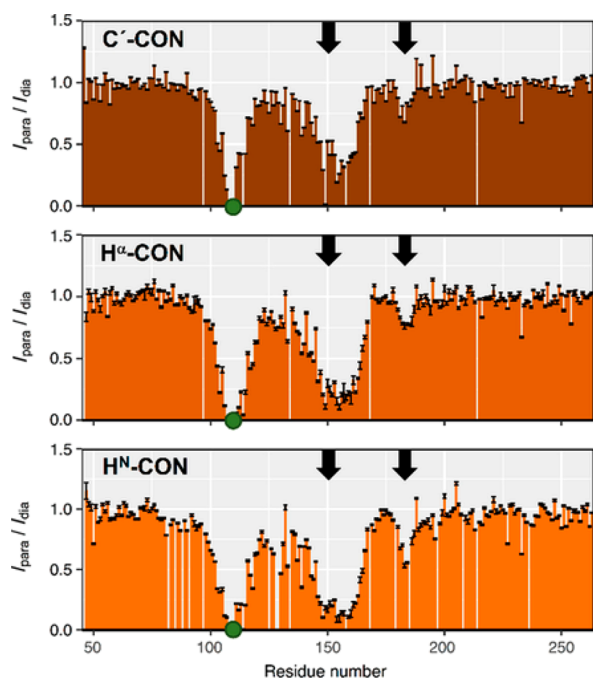


Figure 14. PRE intensity profiles determined through the CON-based approach exploiting different starting magnetization sources: C' -CON (top), H^α -CON (middle), H^N -CON (bottom). The green circle indicates the position of the spin label (S108C). Black arrows indicate the observed long-range contacts with the spin label. Reproduced with permission from ref 199. Copyright 2019 Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim.

the former showing less pronounced effects, in agreement with ^{13}C nuclear spins sensing similar paramagnetic effects at shorter distances from the paramagnetic center with respect to protons. The comparison of the PREs observed through the H^α and $\text{H}^\text{N}\text{CON}$, which both exploit ^1H as a starting polarization source, highlighted the presence of proline residues, which are clustered in the vicinity of the so-called compact state of the polypeptide. In addition, the $\text{H}^\text{N}\text{CON}$ profiles are less uniform with respect to the $\text{H}^\alpha\text{CON}$ profiles; a feature that could be due to additional solvent-mediated relaxation enhancement effects that are picked up by $\text{H}^\text{N}\text{CON}$ and do not influence $\text{H}^\alpha\text{CON}$ data.

Paramagnetic tags can also provide useful long-range information about intermolecular interactions as well as about the relative orientation of two domains in multidomain proteins, such as for the case of tandem RNA Recognition Motif (RRM) domains (RRM1-RRM2) of the splicing factor U2AF65 bound to a polyuridine (U9) RNA oligonucleotide. A paramagnetic tag was introduced at position 155 in the RRM1 domain; the

complementary information achieved through ^{13}C protonless experiments in conjunction with ^1H detected ones was very effective in defining the relative orientation of the two RRM domains connected by a flexible linker.¹⁹⁸

Paramagnetic relaxation enhancements sensed by proteins upon addition of highly paramagnetic complexes to the solution (solvent PREs) are also modulated by the experimental tool used to measure them and protonless experiments resulted useful to access complementary information to that achieved through ^1H - ^{15}N HSQC experiments.²⁰¹ Alternatively, solvent PREs have been proposed as a tool to shorten longitudinal recovery times.^{202,203}

Another area where heteronuclear direct detection can provide useful information on the grounds of the reduced contributions to relaxation sensed by heteronuclear spins with respect to protons as in case of paramagnetic proteins, is that of multimeric proteins²⁰⁴ or large protein assemblies for which extensive enrichment in ^2H (in place of ^1H) is needed to reduce the dipolar contributions to nuclear relaxation and thus the amount of protons that can be investigated is limited to exchangeable amide protons. These can be investigated through transverse relaxation optimized spectroscopy (TROSY)²⁰⁵ that exploits the constructive interference of CSA and DD interactions involving the amide proton and the directly bound amide nitrogen. A similar approach is more difficult to implement in a general way for ^{13}C nuclear spins because of the higher variability in chemical topology and magnitude of the effects. Methyl TROSY,²⁰⁶ which exploits the interference between ^1H - ^1H and ^1H - ^{13}C DD interactions instead, represents the most successful application but only allows the study of methyl groups; other approaches have been proposed for aromatic rings as well as for methylene groups. However, on highly deuterated samples, ^{13}C direct detection offers the most general strategy to recover information that is not easy to access otherwise. This approach, initially demonstrated on dimeric SOD (in the reduced, diamagnetic state)⁹ resulted useful to provide information on a multimeric protein of high molecular mass⁵² as well as for the Fc portion of immunoglobulin G (IgG) with a molecular mass of 56 kDa as model glycoprotein²⁰⁷ and for malate synthase G (MSG) protein (82-kDa).³² With increasing field strength, the utility of this approach is expected to increase thanks to the increased sensitivity and resolution provided by the increase in B_0 , in particular for aliphatic ^{13}C resonances that only experience a modest contribution to relaxation due to the chemical shift anisotropy tensor at increasing magnetic field.^{208,209}

3.3. Intrinsically Disordered Proteins

The field where ^{13}C direct detection finds most applications is that of intrinsically disordered proteins (IDPs) or regions (IDRs) of complex multidomain/heterogeneous proteins. Research in this area flourished almost in parallel to the development of the suite of NMR experiments based on ^{13}C direct detection, triggering the design of the most widely used experiments which in turn provided a wealth of valuable information at atomic resolution about their structural and dynamic properties. Since the early studies on unfolded systems, the importance of exploiting heteronuclei to increase the resolution of cross-peaks in multidimensional NMR spectra was recognized.²¹⁰ Uniform isotopic labeling with ^{13}C and ^{15}N indeed contributed to the increase in the complexity of the systems that could be studied in unfolded states. However, initial studies were largely based on ^1H direct detection, in particular

on detection of amide protons that can be easily correlated to the directly bound ^{15}N , the latter greatly contributing to excellent chemical shift dispersion, a key feature to overcome the main critical point arising from the high flexibility and lack of structure typical of IDPs. Amide protons however are influenced by exchange processes with the solvent (see section 2.4.2) and thus are prone to extensive line broadening in particular in conditions enhancing exchange processes. An interesting work¹¹⁸ analyzed the chemical shifts of disordered proteins deposited in the BMRB and revealed that most studies were based on ^1H detected triple resonance experiments and were performed in conditions deviating from “physiological” ones because a reduction of pH or temperature (or of both) was necessary to diminish the impact of exchange processes on amide proton line widths and to obtain informative 2D HN spectra. In this context the possibility to shift to direct detection of $^{13}\text{C}'$ and exploit the suite of multidimensional NMR experiments offered an excellent alternative to investigate IDPs. Carbonyl carbon nuclei are not influenced by exchange processes and can be easily correlated to the directly bound nitrogen nuclei which, as mentioned above, are the nuclear spins that retain a large chemical shift dispersion also in the absence of a 3D structure. In addition the $\text{C}'\text{--N}$ correlation exploits the scalar coupling deriving from the peptide bond, linking one amino acid with the neighboring one. This offers an important contribution to chemical shift dispersion deriving from the covalent structure of the protein which is particularly relevant for IDPs in which chemical shifts tend to average to those predicted for each amino acid type.⁶⁹ This set of experiments was used to perform the sequence-specific assignment of a large number of IDPs, starting from initial work on α -synuclein,²³ Src²¹¹ and securin.⁷⁵ As a result, the size and complexity of IDPs investigated significantly increased, and assignments deposited were more extended as also proline resonances could be monitored contributing in this way to more accurate characterization of IDPs.

3.3.1. Mimicking Physiological Events. The possibility to study IDPs in physiological conditions also represents an important motivation to choose ^{13}C detected NMR, an important aspect in general but particularly for IDPs that are highly solvent-exposed and thus sensitive to environmental conditions.

Exposed backbones are often sites of post-translational modifications (PTMs), a field in which NMR can provide valuable contributions by highlighting at atomic resolution which amino acids are affected by these chemical reactions in real-time. Phosphorylation of OH groups of serine, threonine and tyrosine residues is one of the most extensively studied post-translational modifications through NMR.²¹² The impact of this reaction on amide protons and nitrogen chemical shifts is very large and 2D HN spectra have been often used to follow these reactions. However, several kinases change their activity depending on pH and temperature and it is often important to investigate their effect in conditions near physiological ones. Therefore, the possibility to follow these reactions with an additional experiment, the CON, in particular if acquired simultaneously to the 2D HN,¹¹¹ offers an excellent tool of investigation for this challenging topic.^{63,213} Another strategy that has been proposed, in particular in cases in which sensitivity is an issue (either for low protein concentration or for instrumental setup) is the use of the most sensitive 2D experiment, the ^1H -start CACO. Several variants focusing on glycine or nonglycine residues were indeed proposed and used to follow phosphorylation reactions, in the presence of purified

kinases or in cell extracts, on a range of previously problematic targets, namely Mdm2, BRCA2, and Oct4.²⁰³

In-cell NMR spectroscopy provides a unique spectroscopic tool to investigate a protein in conditions that resemble physiological ones and the potential of ^{13}C detected experiments has been tested with several proteins characterized by different molecular mass and motional properties.²¹⁴ The results clearly indicate that for well-folded proteins the method of choice is ^1H NMR since line broadening typical for in-cell protein samples severely affects the detectability of the signals, particularly when CON-type experiments are used. The most valuable contribution provided by ^{13}C detected experiments in this context consists of picking up in a clean way signals deriving from highly flexible protein regions that also retain these properties of high flexibility within whole cells. Indeed, when dealing with IDPs, ^{13}C detected experiments offer a valuable solution to contrast the reduced chemical shift dispersion and high solvent exchange rates of the H^{N} signals, allowing to map entirely the protein backbone. The case of α -synuclein is paradigmatic.¹¹² The complementary features of the different variants of 2D CON spectra ($^1\text{H}^{\text{N}}$ -start, $^1\text{H}^{\alpha}$ -start, $^1\text{H}^{\alpha}$ -flip, etc.) as well as of CACO spectra ($^1\text{H}^{\alpha}$ -start, $^1\text{H}^{\alpha}$ -flip) offer many opportunities to pick up the desired correlations within the lifetime of in-cell NMR samples. As an example, the $\text{H}^{\text{N}}\text{BEST}$ CON and H^{α} CON can be acquired in less than 1 h on in-cell α -synuclein samples.⁶³ More recently, the CON//HN multiple receivers variant was implemented to monitor two experiments simultaneously,¹¹¹ an important feature for samples of limited stability, such as in-cell NMR samples, showing that also the ^{13}C -start variant of the CON can provide useful information in a very limited time (<1 h, Figure 15). Comparison of the two NMR spectra acquired simultaneously reveals that the increase in line width when passing to the in-cell samples is more pronounced for ^1H than for ^{13}C , a possible origin for this could be the local inhomogeneities that affect more nuclear spins with higher gyromagnetic ratios such as ^1H . The simultaneous observation of the two experiments of course provides a wealth of atomic resolved information on the state of the protein within cells.

3.3.2. Liquid–Liquid Phase Separation. Liquid–liquid phase separation (LLPS) is a phase transition process in which a homogeneous solution separates into two different coexisting liquid phases. The interface between the dense and the light phases allows the passage only of certain molecules making these droplets act as compartments which are called membrane-less organelles.²¹⁵ IDPs and protein containing IDRs are over-represented among proteins that undergo phase separation processes, especially in condensates containing RNA. This is likely due to their ability to form transient multivalent interactions. The LLPS behavior of IDPs is governed by various interactions, going from charge–charge interactions to π – π and cation– π interactions as well as hydrophobic contacts and charge pairing; in vitro, phase separation is often driven by changes in pH or temperature or ionic strength of the solution.²¹⁶

In living cells, condensates can act as organizational centers promoting reactions or sequestering proteins and nucleic acids to inhibit metabolic pathways and it is crucial to be able to understand the driving forces leading to their formation. Furthermore, phase transitions are also associated with a range of neurodegenerative diseases and it is pharmaceutically relevant to understand how to interfere with their formation.^{217,218} NMR spectroscopy is one of the main players on this grounds and ^{13}C

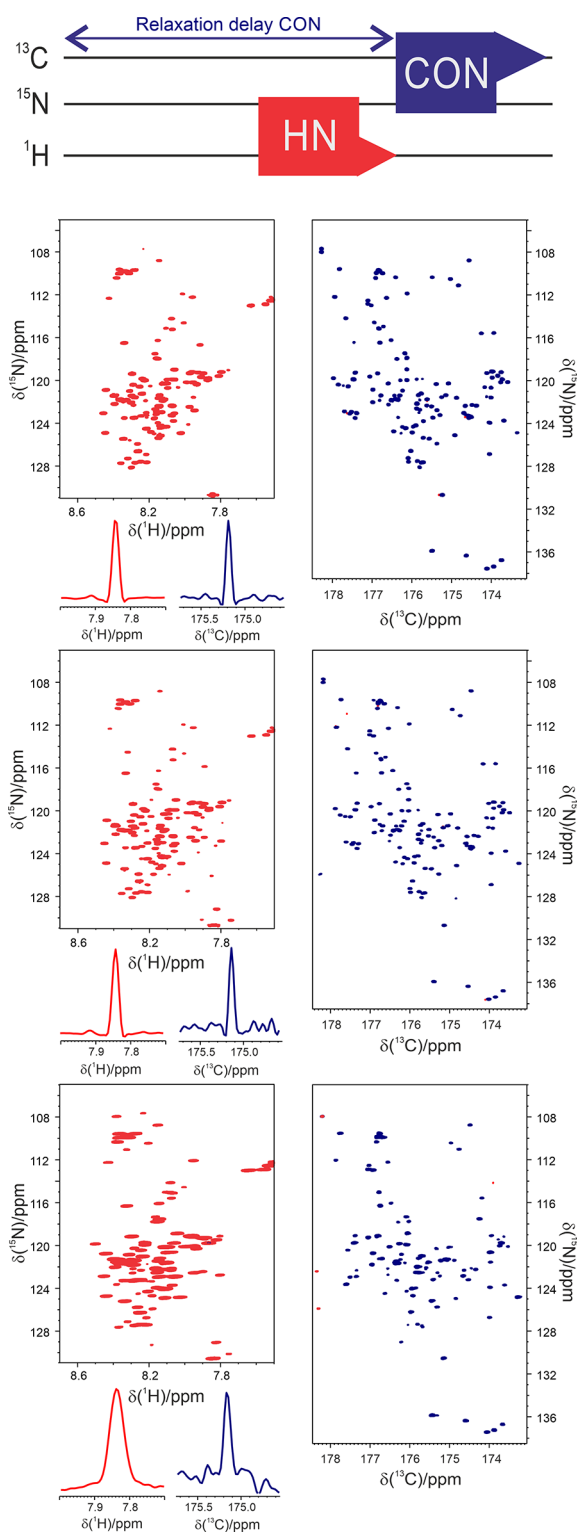


Figure 15. Scheme of the CON//HN experiment (top panel) and comparison of the 2D spectra (HN left, CON right) acquired through this experiment on ^{13}C , ^{15}N labeled α -synuclein at 310 K in different experimental conditions. From top to bottom: purified sample in 100 mM NaCl, 50 μM EDTA, 20 mM phosphate buffer at pH 7.4; in *Escherichia coli* cells lysate resuspended in the same buffer and in-cell. The traces of a representative signal extracted from the HN (red left) and the CON (blue right) spectra are also reported. Adapted with permission from ref 111. Copyright 2019 Biophysical Society.

NMR has been exploited to characterize at the molecular level some of these processes.

An interesting example is the phase transition of Tau, an IDP that functions in microtubule nucleation, assembly, and stabilization. This large protein contains a N-terminal projection domain, two proline-rich regions, and a C-terminal domain that includes three (3R) or four (4R) imperfect 31- or 32-residue repeats. The repeat domain constitutes the microtubule-binding domain and has the ability to bind microtubules and to promote their assembly.²¹⁹ Tau and its shorter variant containing the 4R repeats (termed K18) undergo liquid–liquid phase separation in vitro at pH 8.8 and 37 °C (Figure 16) and ^{13}C NMR was key to be able to map at the residue level changes occurring upon LLPS formation since most of the signals of $^1\text{H}^{\text{N}}$ -detected spectra, and those of Gly, Ser and Thr residues in particular, were broadened beyond detection in these conditions. The high quality of the spectra obtained, including CON and CBCACO, suggested the possibility to follow chemical shift perturbations for each residue and highlight changes in secondary structure propensities upon droplets formation.²²⁰

Another example was recently provided by the investigation of the C-terminal disordered regions of two interacting proteins, FMRP and CAPRIN1.²²¹ These two fragments form condensates individually and in interaction when FMRP is phosphorylated in specific positions. The comparison of the CON spectrum of CAPRIN1 in the presence of phase-separated FMRP and of CAPRIN1 cophase-separated with phosphorylated FMRP (pFMRP) highlighted the involvement of two arginine-rich regions in the interaction as well as the involvement of some tyrosine residues.

3.3.3. Low Complexity Regions. “Low complexity regions” is a term widely used to describe parts of polypeptide chains constituted by a simple composition in terms of amino acids, featuring a few or even only one amino acid-type. These are not expected to be parts of globular folds, at least considering what we know as of today, and are often found as parts of IDPs/IDRs.

Among them an interesting example is provided by the so-called poly-Q, that is parts of polypeptide chains constituted by many glutamine residues (Qs) in a row. These motifs attracted the attention of the scientific community because some proteins featuring elongated stretches of Qs are linked to the onset of the so-called poly-Q diseases, a set of neurodegenerative diseases caused by malfunction of proteins that only have in common a part of the polypeptide chain with a high number of glutamine residues.²²² The atomic resolution structural and dynamic investigation of these proteins is not easy because they do not crystallize, preventing the use of X-ray crystallography, and also from the NMR point of view their investigation represents a challenge due to the highly repetitive primary sequence. With ^{13}C detected NMR experiments it was possible to perform complete sequence specific assignment of the N-terminal fragment of the androgen receptor (AR) comprising up to 25 Q in a row (Figure 17), not just as an isolated peptide but embedded in the N-terminal fragment of AR (135 and 156 amino acids in the 4Q and 25Q variants, respectively). This study revealed a pronounced helical propensity of this region and highlighted the important role of four leucine residues preceding the poly-Q tract to induce helicity in the poly-Q region, as confirmed through a deletion mutant.^{100,223} This may provide a mechanism to protect the poly-Q tract from aggregation, a relevant process for the onset of a neurodegenerative disease (spinal bulbar muscular atrophy). Interest-

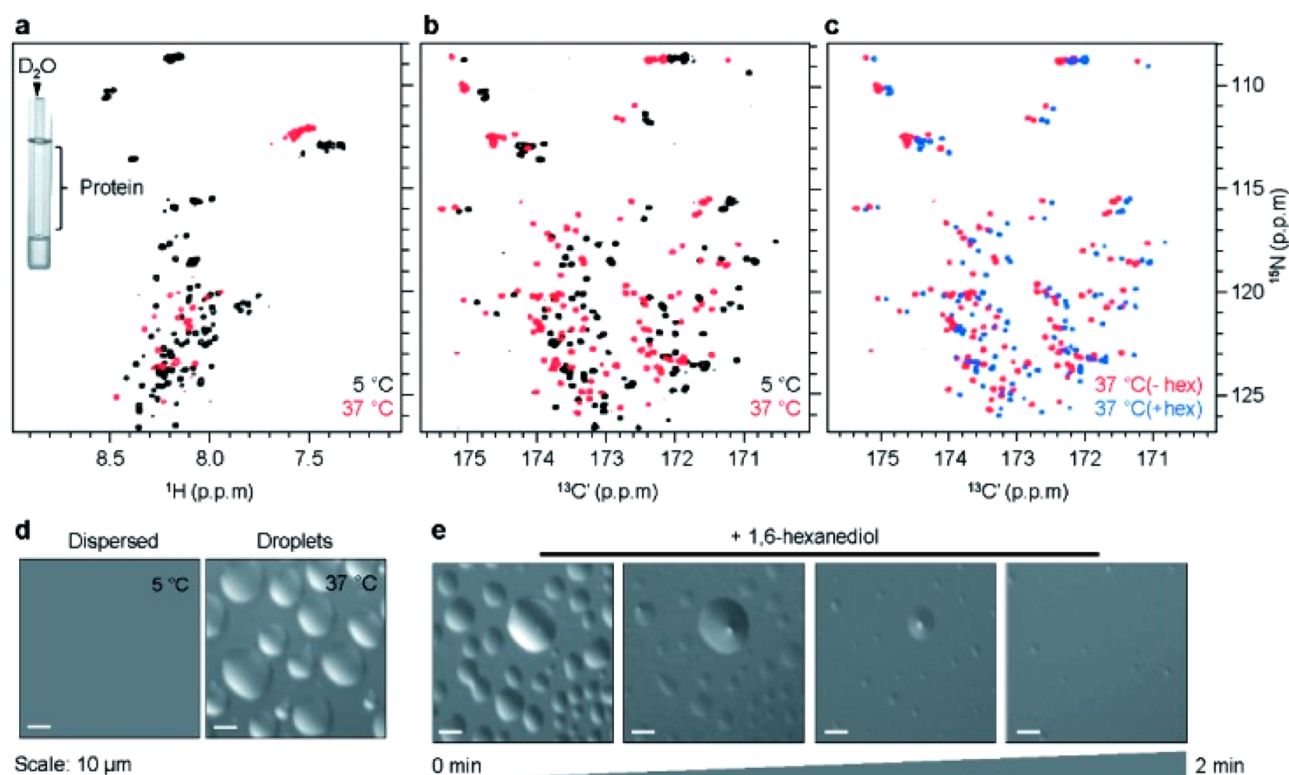


Figure 16. NMR spectroscopy of liquid–liquid phase separation of the repeat domain of tau. (a, b) Superposition of 2D ^1H - ^{15}N HSQC (a) and 2D CON (b) spectra of K18 in the monomeric dispersed state (5 °C, black) and the droplet phase (37 °C, red). To avoid contributions of solvent exchange to NMR signal broadening in ^1H - ^{15}N correlation spectra, D_2O was placed into a separate capillary tube (insert in a). (c) Superposition of CON spectra of K18 recorded at 37 °C in the absence (–hex, red) and presence (+hex, blue) of 3% 1,6-hexanediol, which rapidly dissolved K18 droplets (e). (d) DIC microscopy demonstrates K18 droplet formation at 37 °C. (e) Time-dependent dissolution of K18 droplets in the presence of 3% 1,6-hexanediol observed by DIC microscopy. Reproduced from ref 220 with permission from the Royal Society of Chemistry.

ingly, these poly-Q helices, induced by preceding leucine residues were also identified in functional proteins such as the IDS linker of CBP.²²⁴ The investigation of poly-Q containing proteins has now become a very active field of investigation.^{225–227}

4. CONCLUSIONS AND OUTLOOK

Largely neglected for many years in favor of approaches exploiting direct detection of proton on the grounds of the intrinsic higher sensitivity of the latter, from the early 2000s ^{13}C direct detected NMR in solution has been rediscovered thanks to hardware development and experimental improvements.

The first spectra we recorded at 16.4 T (700 MHz ^1H frequency) were obtained with a spectrometer equipped with a room-temperature inverse detection probehead, with a S/N ratio on the ASTM standard sample of about 200:1; today we are routinely using a cryogenically cooled probe for direct ^{13}C observation with a S/N ratio of 2800:1 on the same sample. This increase in sensitivity opened the way to routine use of ^{13}C direct detection by allowing a significant reduction of experimental times to access the same information (a factor of about 200). With the latest magnet and probeheads technology, the instrumental sensitivity jumps above 4400 for a 28.1 T (1.2 GHz) spectrometer.

In a stepwise fashion, solutions to overcome the potential limitations of the approach have been devised. Automation for removing the large splitting due to one bond scalar couplings in the ^{13}C direct dimension constituted an important contribution for making it handy. Different polarization sources were used as

the basis to design a wide variety of NMR experiments, ranging from simple 2D experiments to achieve a snapshot of a protein in solution, to 3D experiments for sequence-specific resonance assignment, all the way to complex multidimensional experiments (nD, with $n > 3$) exploiting the most advanced strategies to acquire and process the data. In addition, a wide variety of additional experiments to detect different NMR observables were developed and made available in the NMR spectrometers' library of pulse sequences, thus becoming of general use.

Browsing the literature, it can be realized that next to papers dedicated to the development of the experiments and exploring potential novel applications, there are many other papers including ^{13}C detected experiments side by side to ^1H detected ones for fingerprinting or to complete the characterization or the assignment of a protein. This is a clear indication that since the early 2000s when it was considered a sort of revival of a vintage technology, ^{13}C direct detection NMR in solution is now considered a useful complement to ^1H NMR that can provide unique information or simplify assignment and characterization of biomolecules in general.

One of the research areas where ^{13}C direct detection has proven very successful, as also witnessed but the many applications populating this account, is the characterization of intrinsically disordered proteins and heterogeneous proteins presenting highly mobile portions of their chain. Indeed, ^{13}C NMR provides larger chemical shift dispersion with respect to proton NMR and benefits from the favorable relaxation properties of disordered polypeptides. We expect that this area will be even more important in the future since these highly

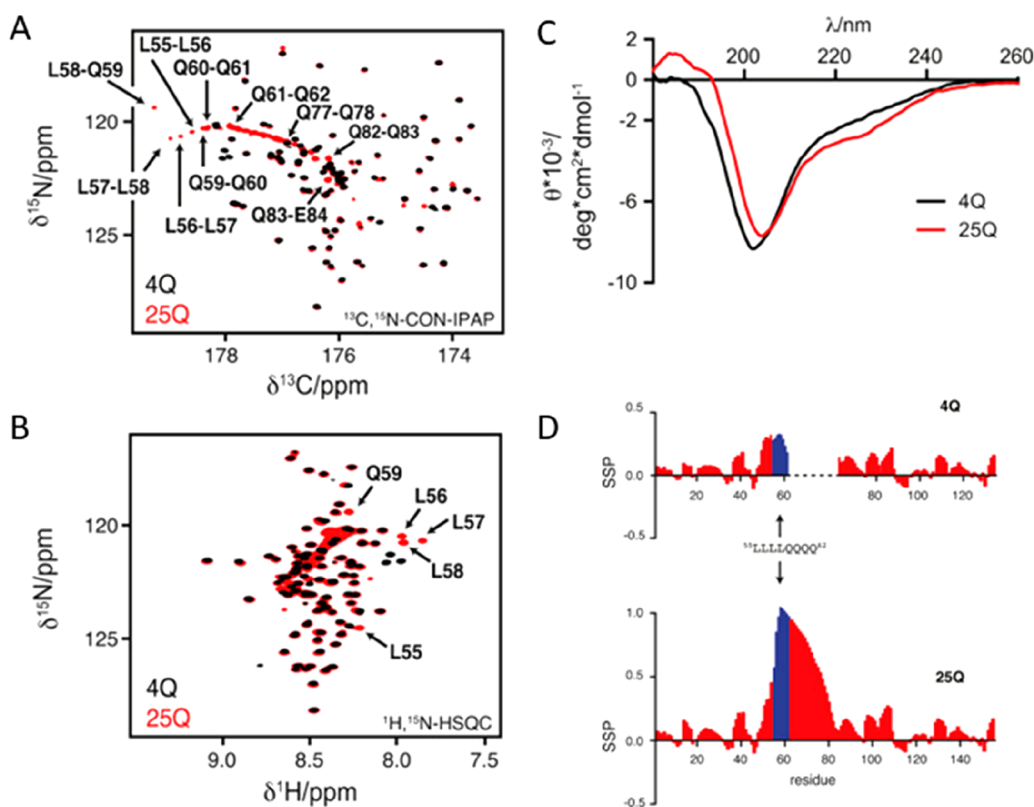


Figure 17. Comparison of the (A, B) NMR and (C) CD spectra of 4Q and 25Q; secondary structure propensity of 4Q and 25Q is reported in (D). The central region of (A) the ^{13}C , ^{15}N -CON-IPAP spectrum and (B) the ^1H - ^{15}N HSQC spectrum of 400 μM 4Q (black) and 25Q (red) at 278 K are shown, with the assignment of the resonances that experience the largest chemical shift variations upon increasing the length of the poly-Q tract. (C) CD spectrum of 120 μM 4Q and 130 μM 25Q at 310 K. In (D) Values for residues 55 to 62, corresponding to the $^{55}\text{LLLL}^{58}$ motif and the first 4 Gln of the poly-Q tract, are shown in blue to highlight the variation of the structural properties of the protein due to the different length of the poly-Q tract. To facilitate the comparison, the values for the residues of 4Q that follow the poly-Q tract are shifted to the right by 21 units. Adapted from ref 100. Copyright 2016 Biophysical Society.

flexible disordered proteins constitute a large share of the proteome and need to be investigated at atomic level.

This leads to the topic of the capability of highly flexible proteins to engage in intermolecular interactions still in a nonspecific way, interactions that lead to fuzzy complexes as well as to the process of liquid–liquid phase separation, still not well characterized. The investigation of these interactions at atomic resolution has great potential to reveal the driving forces of these processes and exclusively heteronuclear experiments can play a central role, monitoring both protein backbone as well as key side-chains involved in this process.

AUTHOR INFORMATION

Corresponding Authors

Isabella C. Felli – Department of Chemistry “Ugo Schiff” and Magnetic Resonance Center, University of Florence, 50019 Sesto Fiorentino, (Florence), Italy; orcid.org/0000-0002-6018-9090; Email: felli@cerm.unifi.it

Roberta Pierattelli – Department of Chemistry “Ugo Schiff” and Magnetic Resonance Center, University of Florence, 50019 Sesto Fiorentino, (Florence), Italy; orcid.org/0000-0001-7755-0885; Email: roberta.pierattelli@unifi.it

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.chemrev.1c00871>

Notes

The authors declare no competing financial interest.

Biographies

Isabella C. Felli graduated with a degree in chemistry in 1994 and obtained a PhD in 1998 at the University of Florence with Prof. Bertini. She then worked with Prof. Bodenhausen in Tallahassee, and as an Alexander von Humboldt fellow with Prof. Griesinger in Frankfurt. Since 2005, she is Associate Professor of Chemistry at the University of Florence. She is interested in the development of new methods for biomolecular NMR.

Roberta Pierattelli graduated with a degree in chemistry at the University of Florence and received a PhD in chemistry at the same university in 1995. After a postdoctoral year at the University of Southampton, she was appointed as a Researcher at the University of Florence, where she is now Full Professor of Chemistry. Her research interests are mainly related to applications of NMR spectroscopy in solution and in the solid state to study the structure and function of proteins and their interactions.

ACKNOWLEDGMENTS

Support of the CERM/CIRMMP center of Instruct-ERIC, of the Italian Ministry of University and Research and of Fondazione CR Firenze is gratefully acknowledged.

ABBREVIATIONS

AP	antiphase
APSY	automated projection spectroscopy
BMRB	biological magnetic resonance data bank
CPMG	Carr–Purcell–Meiboom–Gill
CS	compressed sensing
CSA	chemical shift anisotropy
DD	dipole–dipole
D-DNP	dissolution dynamic nuclear polarization
DIPAP	double IPAP
DS3E	double S3E
FFT	fast Fourier transform
FID	free induction decay
INEPT	insensitive nuclei enhanced by polarization transfer
IP	in-phase
LRE	longitudinal relaxation enhancement
MDD	multidimensional decomposition
ME	maximum entropy
MFT	multidimensional Fourier transform
NMR	nuclear magnetic resonance
NUS	nonuniform sampling
PCS	pseudocontact shift
PRE	paramagnetic relaxation enhancement
PTM	post-translational modification
RDC	residual dipole coupling
S3E	spin states selective
SMFT	sparse multidimensional Fourier transform

REFERENCES

- (1) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Oxford University Press, Ed.; Oxford, 1990; pp 1–640.
- (2) LeMaster, D. M.; Richards, F. M. NMR Sequential Assignment of *Escherichia Coli* Thioredoxin Utilizing Random Fractional Deuteration. *Biochemistry* **1988**, *27*, 142–150.
- (3) Markus, M. A.; Dayie, K. T.; Matsudaira, P.; Wagner, G. Effect of Deuteration on the Amide Proton Relaxation Rates in Proteins. Heteronuclear NMR Experiments on Villin 14T. *J. Magn. Reson., Series B* **1994**, *105*, 192–195.
- (4) Sattler, M.; Fesik, S. W. Use of Deuterium Labeling in NMR: Overcoming a Sizeable Problem. *Structure* **1996**, *4* (11), 1245–1249.
- (5) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G. I.; Rance, M.; Skelton, N. J. *Protein NMR Spectroscopy*; Elsevier, 2007; pp 1–915.
- (6) Schmidt, J. M.; Howard, M. J.; Maestre-Martínez, M.; Pérez, C. S.; Löhr, F. Variation in Protein C α -Related One-Bond J Couplings. *Magn. Reson. Chem.* **2009**, *47*, 16–30.
- (7) Hoch, J. C.; Maciejewski, M. W.; Mobli, M.; Schuyler, A. D.; Stern, A. S. Nonuniform Sampling and Maximum Entropy Reconstruction in Multidimensional NMR. *Acc. Chem. Res.* **2014**, *47*, 708–717.
- (8) Shimba, N.; Kovacs, H.; Stern, A. S.; Nomura, A. M.; Shimada, I.; Hoch, J. C.; Craik, C. S.; Dötsch, V. Optimization of ^{13}C Direct Detection NMR Methods. *J. Biomol. NMR* **2004**, *30*, 175–179.
- (9) Bertini, I.; Felli, I. C.; Kümmerle, R.; Moskau, D.; Pierattelli, R. ^{13}C – ^{13}C NOESY: An Attractive Alternative for Studying Large Macromolecules. *J. Am. Chem. Soc.* **2004**, *126*, 464–465.
- (10) Karunanithy, G.; Hansen, D. F. FID-Net: A Versatile Deep Neural Network Architecture for NMR Spectral Reconstruction and Virtual Decoupling. *J. Biomol. NMR* **2021**, *75*, 179–191.
- (11) Weigelt, J.; Hammarström, A.; Bermel, W.; Otting, G. Removal of Zero-Quantum Coherence in Protein NMR Spectra Using SESAM Decoupling and Suppression of Decoupling Sidebands. *J. Magn. Reson., Series B* **1996**, *110*, 219–224.
- (12) Vögeli, B.; Kovacs, H.; Pervushin, K. Simultaneous ^1H - or ^2H -, ^{15}N - and Multiple-Band-Selective ^{13}C -Decoupling during Acquisition in ^{13}C -Detected Experiments with Proteins and Oligonucleotides. *J. Biomol. NMR* **2005**, *31*, 1–9.
- (13) Bertini, I.; Felli, I. C.; Luchinat, C.; Parigi, G.; Pierattelli, R. Towards a Protocol for Solution Structure Determination of Copper(II) Proteins: The Case of $\text{Cu}^{12}\text{Zn}^{11}$ Superoxide Dismutase. *ChemBioChem* **2007**, *8*, 1422–1429.
- (14) Ying, J.; Li, F.; Lee, J. H.; Bax, A. ^{13}C Decoupling during Direct Observation of Carbonyl Resonances in Solution NMR of Isotopically Enriched Proteins. *J. Biomol. NMR* **2014**, *60*, 15–21.
- (15) Struppe, J. O.; Yang, C.; Wang, Y.; Hernandez, R. v.; Shamansky, L. M.; Mueller, L. J. Long-Observation-Window Band-Selective Homonuclear Decoupling: Increased Sensitivity and Resolution in Solid-State NMR Spectroscopy of Proteins. *J. Magn. Reson.* **2013**, *236*, 89–94.
- (16) Felli, I. C.; Pierattelli, R. Spin-State-Selective Methods in Solution and Solid-State Biomolecular ^{13}C NMR. *Prog. Nucl. Magn.* **2015**, *84*–85, 1–13.
- (17) Ottiger, M.; Delaglio, F.; Bax, A. Measurement of J and Dipolar Couplings from Simplified Two-Dimensional NMR Spectra. *J. Magn. Reson.* **1998**, *131*, 373–378.
- (18) Duma, L.; Hediger, S.; Lesage, A.; Emsley, L. Spin-State Selection in Solid-State NMR. *J. Magn. Reson.* **2003**, *164*, 187–195.
- (19) Bertini, I.; Felli, I. C.; Kümmerle, R.; Luchinat, C.; Pierattelli, R. ^{13}C – ^{13}C NOESY: A Constructive Use of ^{13}C – ^{13}C Spin-Diffusion. *J. Biomol. NMR* **2004**, *30*, 245–251.
- (20) Bermel, W.; Bertini, I.; Duma, L.; Felli, I. C.; Emsley, L.; Pierattelli, R.; Vasos, P. R. Complete Assignment of Heteronuclear Protein Resonances by Protonless NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2005**, *44*, 3089–3092.
- (21) Laage, S.; Lesage, A.; Emsley, L.; Bertini, I.; Felli, I. C.; Pierattelli, R.; Pintacuda, G. Transverse-Dephasing Optimized Homonuclear J-Decoupling in Solid-State NMR Spectroscopy of Uniformly ^{13}C -Labeled Proteins. *J. Am. Chem. Soc.* **2009**, *131*, 10816–10817.
- (22) Furuita, K.; Sugiki, T.; Takamuku, M.; Hattori, Y.; So, M.; Kawata, Y.; Ikegami, T.; Fujiwara, T.; Kojima, C. Sensitivity Enhancement by Sequential Data Acquisition for ^{13}C -Direct Detection NMR. *J. Magn. Reson.* **2021**, *322*, 106878.
- (23) Bermel, W.; Bertini, I.; Felli, I. C.; Lee, Y.-M.; Luchinat, C.; Pierattelli, R. Protonless NMR Experiments for Sequence-Specific Assignment of Backbone Nuclei in Unfolded Proteins. *J. Am. Chem. Soc.* **2006**, *128*, 3918–3919.
- (24) Hu, K.; Eletsky, A.; Pervushin, K. Backbone Resonance Assignment in Large Protonated Proteins Using a Combination of New 3D TROSY-HN(CA)HA, 4D TROSY-HACANH and ^{13}C -Detected HACACO Experiments. *J. Biomol. NMR* **2003**, *26*, 69–77.
- (25) Bertini, I.; Jiménez, B.; Piccioli, M. ^{13}C Direct Detected Experiments: Optimization for Paramagnetic Signals. *J. Magn. Reson.* **2005**, *174*, 125–132.
- (26) Lee, D.; Vögeli, B.; Pervushin, K. Detection of C', C α Correlations in Proteins Using a New Time- and Sensitivity-Optimal Experiment. *J. Biomol. NMR* **2005**, *31*, 273–278.
- (27) Hu, K.; Vögeli, B.; Clore, G. M. Spin-State Selective Carbon-Detected HNCO with TROSY Optimization in all Dimensions and Double Echo-Antiecho Sensitivity Enhancement in Both Indirect Dimensions. *J. Am. Chem. Soc.* **2007**, *129*, 5484–5491.
- (28) Bermel, W.; Bertini, I.; Csizmok, V.; Felli, I. C.; Pierattelli, R.; Tompa, P. H-Start for Exclusively Heteronuclear NMR Spectroscopy: The Case of Intrinsically Disordered Proteins. *J. Magn. Reson.* **2009**, *198*, 275–281.
- (29) Bertini, I.; Felli, I. C.; Gonnelli, L.; Pierattelli, R.; Spyridi, Z.; Spyroulias, G. A. Mapping Protein–Protein Interaction by ^{13}C '-Detected Heteronuclear NMR Spectroscopy. *J. Biomol. NMR* **2006**, *36*, 111–122.
- (30) Bermel, W.; Bertini, I.; Felli, I. C.; Matzapetakis, M.; Pierattelli, R.; Theil, E. C.; Turano, P. A Method for C α Direct-Detection in Protonless NMR. *J. Magn. Reson.* **2007**, *188*, 301–310.
- (31) Yoshimura, Y.; Oktaviani, N. A.; Yonezawa, K.; Kamikubo, H.; Mulder, F. A. A. Unambiguous Determination of Protein Arginine Ionization States in Solution by NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2017**, *56*, 239–242.

- (32) Pritchard, R. B.; Hansen, D. F. Characterising Side Chains in Large Proteins by Protonless ^{13}C -Detected NMR Spectroscopy. *Nat. Commun.* **2019**, *10*, 1747.
- (33) Fiala, R.; Sklenář, V. ^{13}C -Detected NMR Experiments for Measuring Chemical Shifts and Coupling Constants in Nucleic Acid Bases. *J. Biomol. NMR* **2007**, *39*, 153–163.
- (34) Fares, C.; Amata, I.; Carlomagno, T. ^{13}C -Detection in RNA Bases: Revealing Structure-Chemical Shift Relationships. *J. Am. Chem. Soc.* **2007**, *129*, 15814–15823.
- (35) Richter, C.; Kovacs, H.; Buck, J.; Wacker, A.; Fürtig, B.; Bermel, W.; Schwalbe, H. ^{13}C -Direct Detected NMR Experiments for the Sequential J-Based Resonance Assignment of RNA Oligonucleotides. *J. Biomol. NMR* **2010**, *47*, 259–269.
- (36) Fürtig, B.; Schnieders, R.; Richter, C.; Zetzsche, H.; Keyhani, S.; Helmling, C.; Kovacs, H.; Schwalbe, H. Direct ^{13}C -Detected NMR Experiments for Mapping and Characterization of Hydrogen Bonds in RNA. *J. Biomol. NMR* **2016**, *64*, 207–221.
- (37) Schnieders, R.; Wolter, A. C.; Richter, C.; Wöhnert, J.; Schwalbe, H.; Fürtig, B. Novel ^{13}C -Detected NMR Experiments for the Precise Detection of RNA Structure. *Angew. Chem., Int. Ed.* **2019**, *58*, 9140–9144.
- (38) Babini, E.; Felli, I. C.; Lelli, M.; Luchinat, C.; Pierattelli, R. Backbone and Side-Chains ^1H , ^{13}C and ^{15}N NMR Assignment of Human β -Parvalbumin. *J. Biomol. NMR* **2005**, *33*, 137–137.
- (39) Oh, B. H.; Westler, W. M.; Darba, P.; Markley, J. L. Protein Carbon-13 Spin Systems by a Single Two-Dimensional Nuclear Magnetic Resonance Experiment. *Science* **1988**, *240*, 908–911.
- (40) Westler, W. M.; Stockman, B. J.; Markley, J. L.; Hosoya, Y.; Miyake, Y.; Kainosho, M. Correlation of Carbon-13 and Nitrogen-15 Chemical Shifts in Selectively and Uniformly Labeled Proteins by Heteronuclear Two-Dimensional NMR Spectroscopy. *J. Am. Chem. Soc.* **1988**, *110*, 6256–6258.
- (41) Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A. Three-Dimensional Triple-Resonance NMR Spectroscopy of Isotopically Enriched Proteins. *J. Magn. Reson.* **1990**, *89*, 496–514.
- (42) Sattler, M.; Schleucher, J.; Griesinger, C. Heteronuclear Multidimensional NMR Experiments for the Structure Determination of Proteins in Solution. *Prog. Nucl. Magn.* **1999**, *34*, 93–158.
- (43) Machonkin, T. E.; Westler, W. M.; Markley, J. L. $^{13}\text{C}\{^{13}\text{C}\}$ 2D NMR: A Novel Strategy for the Study of Paramagnetic Proteins with Slow Electronic Relaxation Rates. *J. Am. Chem. Soc.* **2002**, *124*, 3204–3205.
- (44) Bermel, W.; Bertini, I.; Felli, I. C.; Kümmerle, R.; Pierattelli, R. ^{13}C Direct Detection Experiments on the Paramagnetic Oxidized Monomeric Copper, Zinc Superoxide Dismutase. *J. Am. Chem. Soc.* **2003**, *125*, 16423–16429.
- (45) Arnesano, F.; Banci, L.; Bertini, I.; Felli, I. C.; Luchinat, C.; Thompson, A. R. A Strategy for the NMR Characterization of Type II Copper(II) Proteins: The Case of the Copper Trafficking Protein CopC from *Pseudomonas Syringae*. *J. Am. Chem. Soc.* **2003**, *125*, 7200–7208.
- (46) Machonkin, T. E.; Westler, W. M.; Markley, J. L. Strategy for the Study of Paramagnetic Proteins with Slow Electronic Relaxation Rates by NMR Spectroscopy: Application to Oxidized Human [2Fe-2S] Ferredoxin. *J. Am. Chem. Soc.* **2004**, *126*, 5413–5426.
- (47) Babini, E.; Bertini, I.; Capozzi, F.; Felli, I. C.; Lelli, M.; Luchinat, C. Direct Carbon Detection in Paramagnetic Metalloproteins to Further Exploit Pseudocontact Shift Restraints. *J. Am. Chem. Soc.* **2004**, *126*, 10496–10497.
- (48) Banci, L.; Bertini, I.; Luchinat, C. *Nuclear and Electron Relaxation: The Magnetic Nucleus-Unpaired Electron Coupling in Solution*; Wiley, 1991; pp 1–228.
- (49) Bertini, I.; Luchinat, C.; Parigi, G.; Pierattelli, R. NMR Spectroscopy of Paramagnetic Metalloproteins. *ChemBioChem* **2005**, *6*, 1536–1549.
- (50) Bermel, W.; Bertini, I.; Felli, I. C.; Piccioli, M.; Pierattelli, R. ^{13}C -Detected Protonless NMR Spectroscopy of Proteins in Solution. *Prog. Nucl. Magn.* **2006**, *48*, 25–45.
- (51) Ohki, S.; Kainosho, M. Stable Isotope Labeling Methods for Protein NMR Spectroscopy. *Prog. Nucl. Magn.* **2008**, *53* (4), 208–226.
- (52) Matzapetakis, M.; Turano, P.; Theil, E. C.; Bertini, I. ^{13}C - ^{13}C NOESY Spectra of a 480 KDa Protein: Solution NMR of Ferritin. *J. Biomol. NMR* **2007**, *38*, 237–242.
- (53) Serber, Z.; Richter, C.; Dötsch, V. Carbon-Detected NMR Experiments to Investigate Structure and Dynamics of Biological Macromolecules. *ChemBioChem* **2001**, *2*, 247–251.
- (54) Jordan, J. B.; Kovacs, H.; Wang, Y.; Mobli, M.; Luo, R.; Anklin, C.; Hoch, J. C.; Kriwacki, R. W. Three-Dimensional ^{13}C -Detected CH_3 -TOCSY Using Selectively Protonated Proteins: Facile Methyl Resonance Assignment and Protein Structure Determination. *J. Am. Chem. Soc.* **2006**, *128*, 9119–9128.
- (55) Morris, G. A.; Freeman, R. Enhancement of Nuclear Magnetic Resonance Signals by Polarization Transfer. *J. Am. Chem. Soc.* **1979**, *101*, 760–762.
- (56) Mueller, L. Sensitivity Enhanced Detection of Weak Nuclei Using Heteronuclear Multiple Quantum Coherence. *J. Am. Chem. Soc.* **1979**, *101*, 4481–4484.
- (57) Bodenhausen, G.; Ruben, D. J. Natural Abundance Nitrogen-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chem. Phys. Lett.* **1980**, *69*, 185–189.
- (58) Bax, A.; Griffey, R. H.; Hawkins, B. L. Correlation of Proton and Nitrogen-15 Chemical Shifts by Multiple Quantum NMR. *J. Magn. Reson.* **1983**, *55*, 301–315.
- (59) Bermel, W.; Bertini, I.; Felli, I. C.; Pierattelli, R. Speeding Up ^{13}C Direct Detection Biomolecular NMR Spectroscopy. *J. Am. Chem. Soc.* **2009**, *131* (42), 15339–15345.
- (60) Hošek, T.; Gil-Caballero, S.; Pierattelli, R.; Brutscher, B.; Felli, I. C. Longitudinal Relaxation Properties of $^1\text{H}^{\text{N}}$ and $^1\text{H}^{\alpha}$ determined by Direct-Detected ^{13}C NMR Experiments to Study Intrinsically Disordered Proteins (IDPs). *J. Magn. Reson.* **2015**, *254*, 19–26.
- (61) Schanda, P.; Kupce, E.; Brutscher, B. SOFAST-HMQC Experiments for Recording Two-Dimensional Heteronuclear Correlation Spectra of Proteins within a Few Seconds. *J. Biomol. NMR* **2005**, *33*, 199–211.
- (62) Lescop, E.; Schanda, P.; Brutscher, B. A Set of BEST Triple-Resonance Experiments for Time-Optimized Protein Resonance Assignment. *J. Magn. Reson.* **2007**, *187*, 163–169.
- (63) Gil, S.; Hošek, T.; Solyom, Z.; Kümmerle, R.; Brutscher, B.; Pierattelli, R.; Felli, I. C. NMR Spectroscopic Studies of Intrinsically Disordered Proteins at Near-Physiological Conditions. *Angew. Chem., Int. Ed.* **2013**, *52*, 11808–11812.
- (64) Lopez, J.; Schneider, R.; Cantrelle, F. X.; Huvent, I.; Lippens, G. Studying Intrinsically Disordered Proteins under True In Vivo Conditions by Combined Cross-Polarization and Carbonyl-Detection NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2016**, *55*, 7418–7422.
- (65) Yuwen, T.; Skrynnikov, N. R. CP-HISQC: A Better Version of HSQC Experiment for Intrinsically Disordered Proteins under Physiological Conditions. *J. Biomol. NMR* **2014**, *58*, 175–192.
- (66) Eletsky, A.; Moreira, O.; Kovacs, H.; Pervushin, K. A Novel Strategy for the Assignment of Side-Chain Resonances in Completely Deuterated Large Proteins Using ^{13}C Spectroscopy. *J. Biomol. NMR* **2003**, *26*, 167–179.
- (67) Balayssac, S.; Jiménez, B.; Piccioli, M. ^{13}C Direct Detected COCO-TOCSY: A Tool for Sequence Specific Assignment and Structure Determination in Protonless NMR Experiments. *J. Magn. Reson.* **2006**, *182*, 325–329.
- (68) Takeuchi, K.; Frueh, D. P.; Hyberts, S. G.; Sun, Z. Y. J.; Wagner, G. High-Resolution 3D CANCA NMR Experiments for Complete Mainchain Assignments Using Ca Direct Detection. *J. Am. Chem. Soc.* **2010**, *132*, 2945–2951.
- (69) Bermel, W.; Bruix, M.; Felli, I. C.; Kumar, M. V.; Pierattelli, R.; Serrano, S. Improving the Chemical Shift Dispersion of Multidimensional NMR Spectra of Intrinsically Disordered Proteins. *J. Biomol. NMR* **2013**, *55*, 231–237.
- (70) Takeuchi, K.; Sun, Z.-Y. J.; Wagner, G. Alternate ^{13}C – ^{12}C Labeling for Complete Mainchain Resonance Assignments Using C^{α}

Direct-Detection with Applicability Toward Fast Relaxing Protein Systems. *J. Am. Chem. Soc.* **2008**, *130*, 17210–17211.

(71) Takeuchi, K.; Frueh, D. P.; Sun, Z. Y. J.; Hiller, S.; Wagner, G. CACA-TOCSY with Alternate ^{13}C - ^{12}C Labeling: A $^{13}\text{C}^{\alpha}$ Direct Detection Experiment for Mainchain Resonance Assignment, Dihedral Angle Information, and Amino Acid Type Identification. *J. Biomol. NMR* **2010**, *47*, 55–63.

(72) Bermel, W.; Bertini, I.; Felli, I. C.; Kümmerle, R.; Pierattelli, R. Novel ^{13}C Direct Detection Experiments, Including Extension to the Third Dimension, to Perform the Complete Assignment of Proteins. *J. Magn. Reson.* **2006**, *178*, 56–64.

(73) Schiavina, M.; Pontoriero, L.; Uversky, V. N.; Felli, I. C.; Pierattelli, R. The Highly Flexible Disordered Regions of the SARS-CoV-2 Nucleocapsid N Protein within the 1–248 Residue Construct: Sequence-Specific Resonance Assignments through NMR. *Biomol. NMR Assign.* **2021**, *15*, 219–227.

(74) Bermel, W.; Bertini, I.; Felli, I. C.; Pierattelli, R.; Vasos, P. R. A Selective Experiment for the Sequential Protein Backbone Assignment from 3D Heteronuclear Spectra. *J. Magn. Reson.* **2005**, *172*, 324–328.

(75) Csizmek, V.; Felli, I. C.; Tompa, P.; Banci, L.; Bertini, I. Structural and Dynamic Characterization of Intrinsically Disordered Human Securin by NMR Spectroscopy. *J. Am. Chem. Soc.* **2008**, *130*, 16873–16879.

(76) Felli, I. C.; Pierattelli, R.; Glaser, S. J.; Luy, B. Relaxation-Optimised Hartmann–Hahn Transfer Using a Specifically Tailored MOCCA-XY16 Mixing Sequence for Carbonyl–Carbonyl Correlation Spectroscopy in ^{13}C Direct Detection NMR Experiments. *J. Biomol. NMR* **2009**, *43*, 187–196.

(77) Mateos, B.; Conrad-Billroth, C.; Schiavina, M.; Beier, A.; Kontaxis, G.; Konrat, R.; Felli, I. C.; Pierattelli, R. The Ambivalent Role of Proline Residues in an Intrinsically Disordered Protein: From Disorder Promoters to Compaction Facilitators. *J. Mol. Biol.* **2020**, *432*, 3093–3111.

(78) Pantoja-Uceda, D.; Santoro, J. Direct Correlation of Consecutive C'-N Groups in Proteins: A Method for the Assignment of Intrinsically Disordered Proteins. *J. Biomol. NMR* **2013**, *57*, 57–63.

(79) Schubert, M.; Oschkinat, H.; Schmieder, P. Music, Selective Pulses, and Tuned Delays: Amino Acid Type-Selective ^1H - ^{15}N Correlations, II. *J. Magn. Reson.* **2001**, *148*, 61–72.

(80) Schubert, M.; Oschkinat, H.; Schmieder, P. Amino Acid Type-Selective Backbone ^1H - ^{15}N -Correlations for Arg and Lys. *J. Biomol. NMR* **2001**, *20*, 379–384.

(81) Schubert, M.; Oschkinat, H.; Schmieder, P. MUSIC and Aromatic Residues: Amino Acid Type-Selective ^1H - ^{15}N Correlations, III. *J. Magn. Reson.* **2001**, *153*, 186–192.

(82) Bermel, W.; Bertini, I.; Chill, J.; Felli, I. C.; Haba, N.; Vasantha Kumar, M. V.; Pierattelli, R. Exclusively Heteronuclear ^{13}C -Detected Amino-Acid-Selective NMR Experiments for the Study of Intrinsically Disordered Proteins (IDPs). *ChemBioChem.* **2012**, *13*, 2425–2432.

(83) Pantoja-Uceda, D.; Santoro, J. A Suite of Amino Acid Residue Type Classification Pulse Sequences for ^{13}C -Detected NMR of Proteins. *J. Magn. Reson.* **2013**, *234*, 190–196.

(84) O'Hare, B.; Benesi, A. J.; Showalter, S. A. Incorporating ^1H Chemical Shift Determination into ^{13}C -Direct Detected Spectroscopy of Intrinsically Disordered Proteins in Solution. *J. Magn. Reson.* **2009**, *200*, 354–358.

(85) Gołowicz, D.; Kasprzak, P.; Orekhov, V.; Kazimierczuk, K. Fast Time-Resolved NMR with Non-Uniform Sampling. *Prog. Nucl. Magn. Reson.* **2020**, *116*, 40–55.

(86) Kazimierczuk, K.; Zawadzka, A.; Koźmiński, W.; Zhukov, I. Lineshapes and Artifacts in Multidimensional Fourier Transform of Arbitrary Sampled NMR Data Sets. *J. Magn. Reson.* **2007**, *188*, 344–356.

(87) Orekhov, V. Y.; Ibraghimov, I.; Billeter, M. Optimizing Resolution in Multidimensional NMR by Three-Way Decomposition. *J. Biomol. NMR* **2003**, *27*, 165–173.

(88) Kazimierczuk, K.; Orekhov, V. Y. Accelerated NMR Spectroscopy by Using Compressed Sensing. *Angew. Chem., Int. Ed.* **2011**, *50*, 5556–5559.

(89) Kazimierczuk, K.; Zawadzka, A.; Koźmiński, W. Optimization of Random Time Domain Sampling in Multidimensional NMR. *J. Magn. Reson.* **2008**, *192*, 123–130.

(90) Jee, J. G. Application of Non-Uniform Sampling to Three Dimensional Carbon Direct-Detection NMR Experiment. *Bull. Korean Chem. Soc.* **2011**, *32*, 3551–3552.

(91) Kazimierczuk, K.; Zawadzka, A.; Koźmiński, W. Narrow Peaks and High Dimensionalities: Exploiting the Advantages of Random Sampling. *J. Magn. Reson.* **2009**, *197*, 219–228.

(92) Bermel, W.; Felli, I. C.; Gonnelli, L.; Koźmiński, W.; Piai, A.; Pierattelli, R.; Zawadzka-Kazimierczuk, A. High-Dimensionality ^{13}C Direct-Detected NMR Experiments for the Automatic Assignment of Intrinsically Disordered Proteins. *J. Biomol. NMR* **2013**, *57*, 353–361.

(93) Pantoja-Uceda, D.; Santoro, J. New ^{13}C -Detected Experiments for the Assignment of Intrinsically Disordered Proteins. *J. Biomol. NMR* **2014**, *59*, 43–50.

(94) Nováček, J.; Haba, N. Y.; Chill, J. H.; Židek, L.; Sklenář, V. 4D Non-Uniformly Sampled HCBCACON and $^1\text{J}(\text{NC}^{\alpha})$ -Selective HCBCANCO Experiments for the Sequential Assignment and Chemical Shift Analysis of Intrinsically Disordered Proteins. *J. Biomol. NMR* **2012**, *53*, 139–148.

(95) Bermel, W.; Bertini, I.; Felli, I. C.; Gonnelli, L.; Koźmiński, W.; Piai, A.; Pierattelli, R.; Stanek, J. Speeding up Sequence Specific Assignment of IDPs. *J. Biomol. NMR* **2012**, *53*, 293–301.

(96) Nováček, J.; Zawadzka-Kazimierczuk, A.; Papoušková, V.; Židek, L.; Šanderová, H.; Krásný, L.; Koźmiński, W.; Sklenář, V. 5D ^{13}C -Detected Experiments for Backbone Assignment of Unstructured Proteins with a Very Low Signal Dispersion. *J. Biomol. NMR* **2011**, *50*, 1–11.

(97) Nováček, J.; Janda, L.; Dopitová, R.; Židek, L.; Sklenář, V. Efficient Protocol for Backbone and Side-Chain Assignments of Large, Intrinsically Disordered Proteins: Transient Secondary Structure Analysis of 49.2 kDa Microtubule Associated Protein 2c. *J. Biomol. NMR* **2013**, *56*, 291–301.

(98) Piai, A.; Hošek, T.; Gonnelli, L.; Zawadzka-Kazimierczuk, A.; Koźmiński, W.; Brutscher, B.; Bermel, W.; Pierattelli, R.; Felli, I. C. CON-CON Assignment Strategy for Highly Flexible Intrinsically Disordered Proteins. *J. Biomol. NMR* **2014**, *60*, 209–218.

(99) Piai, A.; Calçada, E. O.; Tarenzi, T.; del Grande, A.; Varadi, M.; Tompa, P.; Felli, I. C.; Pierattelli, R. Just a Flexible Linker? The Structural and Dynamic Properties of CBP-ID4 Revealed by NMR Spectroscopy. *Biophys. J.* **2016**, *110*, 372–381.

(100) Eftekharzadeh, B.; Piai, A.; Chiesa, G.; Mungianu, D.; García, J.; Pierattelli, R.; Felli, I. C.; Salvatella, X. Sequence Context Influences the Structure and Aggregation Behavior of a PolyQ Tract. *Biophys. J.* **2016**, *110*, 2361–2366.

(101) Contreras-Martos, S.; Piai, A.; Kosol, S.; Varadi, M.; Bekesi, A.; Lebrun, P.; Volkov, A. N.; Gevaert, K.; Pierattelli, R.; Felli, I. C.; Tompa, P. Linking Functions: An Additional Role for an Intrinsically Disordered Linker Domain in the Transcriptional Coactivator CBP. *Sci. Rep.* **2017**, *7*, 4676.

(102) Schiavina, M.; Salladini, E.; Murrall, M. G.; Tria, G.; Felli, I. C.; Pierattelli, R.; Longhi, S. Ensemble Description of the Intrinsically Disordered N-Terminal Domain of the Nipah Virus P/V Protein from Combined NMR and SAXS. *Sci. Rep.* **2020**, *10*, 19574.

(103) Dziekański, P.; Grudziak, K.; Jarvoll, P.; Koźmiński, W.; Zawadzka-Kazimierczuk, A. ^{13}C -Detected NMR Experiments for Automatic Resonance Assignment of IDPs and Multiple-Fixing SMFT Processing. *J. Biomol. NMR* **2015**, *62*, 179–190.

(104) Piai, A.; Gonnelli, L.; Felli, I. C.; Pierattelli, R.; Kazimierczuk, K.; Grudziak, K.; Koźmiński, W.; Zawadzka-Kazimierczuk, A. Amino Acid Recognition for Automatic Resonance Assignment of Intrinsically Disordered Proteins. *J. Biomol. NMR* **2016**, *64*, 239–253.

(105) Kupče, E.; Freeman, R. Projection–Reconstruction of Three-Dimensional NMR Spectra. *J. Am. Chem. Soc.* **2003**, *125*, 13958–13959.

(106) Hiller, S.; Fiorito, F.; Wuthrich, K.; Wider, G. Automated Projection Spectroscopy (APSY). *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10876–81.

- (107) Murrall, M. G.; Schiavina, M.; Sainati, V.; Bermel, W.; Pierattelli, R.; Felli, I. C. ^{13}C APSY-NMR for Sequential Assignment of Intrinsically Disordered Proteins. *J. Biomol. NMR* **2018**, *70*, 167–175.
- (108) Pontoriero, L.; Schiavina, M.; Murrall, M. G.; Pierattelli, R.; Felli, I. C. Monitoring the Interaction of α -Synuclein with Calcium Ions through Exclusively Heteronuclear Nuclear Magnetic Resonance Experiments. *Angew. Chem., Int. Ed.* **2020**, *132*, 18696–18704.
- (109) Kupče, E.; Kay, L. E.; Freeman, R. Detecting the “Afterglow” of ^{13}C NMR in Proteins Using Multiple Receivers. *J. Am. Chem. Soc.* **2010**, *132*, 18008–18011.
- (110) Kupče, E.; Kay, L. E. Parallel Acquisition of Multi-Dimensional Spectra in Protein NMR. *J. Biomol. NMR* **2012**, *54*, 1–7.
- (111) Schiavina, M.; Murrall, M. G.; Pontoriero, L.; Sainati, V.; Kümmerle, R.; Bermel, W.; Pierattelli, R.; Felli, I. C. Taking Simultaneous Snapshots of Intrinsically Disordered Proteins in Action. *Biophys. J.* **2019**, *117*, 46–55.
- (112) Felli, I. C.; Gonnelli, L.; Pierattelli, R. In-Cell ^{13}C NMR Spectroscopy for the Study of Intrinsically Disordered Proteins. *Nat. Protoc.* **2014**, *9*, 2005–2016.
- (113) Spera, S.; Bax, A. Empirical Correlation between Protein Backbone Conformation and C^α and C^β ^{13}C Nuclear Magnetic Resonance Chemical Shifts. *J. Am. Chem. Soc.* **1991**, *113*, 5490–5492.
- (114) Wishart, D. S.; Sykes, B. D.; Richards, F. M. The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure Through NMR Spectroscopy. *Biochemistry* **1992**, *31*, 1647–1651.
- (115) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: A Hybrid Method for Predicting Protein Backbone Torsion Angles from NMR Chemical Shifts. *J. Biomol. NMR* **2009**, *44*, 213–223.
- (116) Shen, Y.; Bax, A. Protein Structural Information Derived from NMR Chemical Shift with the Neural Network Program TALOS-N. *Methods Mol. Biol.* **2015**, *1260*, 17–32.
- (117) Schwarzing, S.; Kroon, G. J. A.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. Sequence-Dependent Correction of Random Coil NMR Chemical Shifts. *J. Am. Chem. Soc.* **2001**, *123*, 2970–2978.
- (118) Tamiola, K.; Acar, B.; Mulder, F. A. A. Sequence-Specific Random Coil Chemical Shifts of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2010**, *132*, 18000–18003.
- (119) Camilloni, C.; De Simone, A.; Vranken, W. F.; Vendruscolo, M. Determination of Secondary Structure Populations in Disordered States of Proteins Using Nuclear Magnetic Resonance Chemical Shifts. *Biochemistry* **2012**, *51*, 2224–2231.
- (120) Kjaergaard, M.; Poulsen, F. M. Disordered Proteins Studied by Chemical Shifts. *Prog. Nucl. Magn.* **2012**, *60*, 42–51.
- (121) Murrall, M. G.; Piai, A.; Bermel, W.; Felli, I. C.; Pierattelli, R. Proline Fingerprint in Intrinsically Disordered Proteins. *ChemBioChem* **2018**, *19*, 1625–1629.
- (122) Palmer, A. G. NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* **2004**, *104*, 3623–3640.
- (123) Bermel, W.; Bertini, I.; Felli, I. C.; Peruzzini, R.; Pierattelli, R. Exclusively Heteronuclear NMR Experiments to Obtain Structural and Dynamic Information on Proteins. *ChemPhysChem* **2010**, *11*, 689–695.
- (124) Lawrence, C. W.; Showalter, S. A. Carbon-Detected ^{15}N NMR Spin Relaxation of an Intrinsically Disordered Protein: FCP1 Dynamics Unbound and in Complex with RAP74. *J. Phys. Chem. Lett.* **2012**, *3*, 1409–1413.
- (125) Pasat, G.; Zintsmaster, J. S.; Peng, J. W. Direct ^{13}C -Detection for Carbonyl Relaxation Studies of Protein Dynamics. *J. Magn. Reson.* **2008**, *193*, 226–232.
- (126) Bertini, I.; Felli, I. C.; Gonnelli, L.; Vasantha Kumar, M. V.; Pierattelli, R. High-Resolution Characterization of Intrinsic Disorder in Proteins: Expanding the Suite of ^{13}C -Detected NMR Spectroscopy Experiments to Determine Key Observables. *ChemBioChem* **2011**, *12*, 2347–2352.
- (127) Brutscher, B.; Felli, I. C.; Gil-Caballero, S.; Hošek, T.; Kümmerle, P. A.; Pierattelli, R.; Solyom, Z. NMR Methods for the Study of Intrinsically Disordered Proteins Structure, Dynamics, and Interactions: General Overview and Practical Guidelines. *Adv. Exp. Med. Biol.* **2015**, *870*, 49–122.
- (128) Kumari, P.; Frey, L.; Sobol, A.; Lakomek, N. A.; Riek, R. ^{15}N Transverse Relaxation Measurements for the Characterization of μs –ms Dynamics are Deteriorated by the Deuterium Isotope Effect on ^{15}N Resulting from Solvent Exchange. *J. Biomol. NMR* **2018**, *72*, 125–137.
- (129) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **1993**, *17*, 75–86.
- (130) Hsu, S. T. D.; Bertoncini, C. W.; Dobson, C. M. Use of Protonless NMR Spectroscopy to Alleviate the Loss of Information Resulting from Exchange-Broadening. *J. Am. Chem. Soc.* **2009**, *131*, 7222–7223.
- (131) Schanda, P.; Forge, V.; Brutscher, B. HET-SOFAST NMR for Fast Detection of Structural Compactness and Heterogeneity along Polypeptide Chains. *Magn. Reson. Chem.* **2006**, *44*, 177–184.
- (132) Deschamps, M.; Campbell, I. D. Cooling Overall Spin Temperature: Protein NMR Experiments Optimized for Longitudinal Relaxation Effects. *J. Magn. Reson.* **2006**, *178*, 206–211.
- (133) Hwang, T.-L.; van Zijl, P. C. M.; Mori, S. Accurate Quantitation of Water–Amide Proton Exchange Rates Using the Phase-Modulated CLEAN Chemical EXchange (CLEANEX-PM) Approach with a Fast-HSQC (FHSQC) Detection Scheme. *J. Biomol. NMR* **1998**, *11*, 221–226.
- (134) Mori, S.; Johnson, M. O.; Berg, J. M.; van Zijl, P. C. M. Water Exchange Filter (WEX Filter) for Nuclear Magnetic Resonance Studies of Macromolecules. *J. Am. Chem. Soc.* **1994**, *116*, 11982–11984.
- (135) Bertini, I.; Felli, I. C. I. C.; Gonnelli, L.; Vasantha Kumar, M. V. V.; Pierattelli, R. High-Resolution Characterization of Intrinsic Disorder in Proteins: Expanding the Suite of ^{13}C -Detected NMR Spectroscopy Experiments to Determine Key Observables. *ChemBioChem* **2011**, *12*, 2347–2352.
- (136) Thakur, A.; Chandra, K.; Dubey, A.; D’Silva, P.; Atreya, H. S. Rapid Characterization of Hydrogen Exchange in Proteins. *Angew. Chem., Int. Ed.* **2013**, *52*, 2440–2443.
- (137) Mackenzie, H. W.; Hansen, D. F. Arginine Side-Chain Hydrogen Exchange: Quantifying Arginine Side-Chain Interactions in Solution. *ChemPhysChem* **2019**, *20*, 252–259.
- (138) Kateb, F.; Pelupessy, P.; Bodenhausen, G. Measuring Fast Hydrogen Exchange Rates by NMR Spectroscopy. *J. Magn. Reson.* **2007**, *184*, 108–113.
- (139) Segawa, T.; Kateb, F.; Duma, L.; Bodenhausen, G.; Pelupessy, P. Exchange Rate Constants of Invisible Protons in Proteins Determined by NMR Spectroscopy. *ChemBioChem* **2008**, *9*, 537–542.
- (140) Sehgal, A. A.; Duma, L.; Bodenhausen, G.; Pelupessy, P. Fast Proton Exchange in Histidine: Measurement of Rate Constants through Indirect Detection by NMR Spectroscopy. *Chem.-Eur. J.* **2014**, *20*, 6332–6338.
- (141) Dass, R.; Corlianò, E.; Mulder, F. A. A. Measurement of Very Fast Exchange Rates of Individual Amide Protons in Proteins by NMR Spectroscopy. *ChemPhysChem* **2019**, *20*, 231–235.
- (142) Dass, R.; Mulder, F. A. A. Paris-DÉCOR: A Protocol for the Determination of Fast Protein Backbone Amide Hydrogen Exchange Rates. *Methods Mol. Biol.* **2020**, *2141*, 337–344.
- (143) Dass, R.; Corlianò, E.; Mulder, F. A. A. The Contribution of Electrostatics to Hydrogen Exchange in the Unfolded Protein State. *Biophys. J.* **2021**, *120*, 4107–4114.
- (144) Scheurer, C.; Skrynnikov, N. R.; Lienin, S. F.; Straus, S. K.; Brüschweiler, R.; Ernst, R. R. Effects of Dynamics and Environment on ^{15}N Chemical Shielding Anisotropy in Proteins. A Combination of Density Functional Theory, Molecular Dynamics Simulation, and NMR Relaxation. *J. Am. Chem. Soc.* **1999**, *121*, 4242–4251.
- (145) Kurzbach, D.; Canet, E.; Flamm, A. G.; Jhajharia, A.; Weber, E. M. M.; Konrat, R.; Bodenhausen, G. Investigation of Intrinsically Disordered Proteins through Exchange with Hyperpolarized Water. *Angew. Chem., Int. Ed.* **2017**, *56*, 389–392.
- (146) Kim, J.; Liu, M.; Hilty, C. Modeling of Polarization Transfer Kinetics in Protein Hydration Using Hyperpolarized Water. *J. Phys. Chem. B* **2017**, *121*, 6492–6498.

- (147) Kadeřávek, P.; Ferrage, F.; Bodenhausen, G.; Kurzbach, D. High-Resolution NMR of Folded Proteins in Hyperpolarized Physiological Solvents. *Chem.-Eur. J.* **2018**, *24*, 13418–13423.
- (148) Szekely, O.; Olsen, G. L.; Felli, I. C.; Frydman, L. High-Resolution 2D NMR of Disordered Proteins Enhanced by Hyperpolarized Water. *Analy. Chem.* **2018**, *90*, 6169–6177.
- (149) Olsen, G. L.; Szekely, O.; Mateos, B.; Kadeřávek, P.; Ferrage, F.; Konrat, R.; Pierattelli, R.; Felli, I. C.; Bodenhausen, G.; Kurzbach, D.; Frydman, L. Sensitivity-Enhanced Three-Dimensional and Carbon-Detected Two-Dimensional NMR of Proteins Using Hyperpolarized Water. *J. Biomol. NMR* **2020**, *74*, 161–171.
- (150) Prestegard, J. H.; Al-Hashimi, H. M.; Tolman, J. R. NMR Structures of Biomolecules Using Field Oriented Media and Residual Dipolar Couplings. *Q. Rev. Biophys.* **2000**, *33*, 371–424.
- (151) Bax, A.; Kontaxis, G.; Tjandra, N. Dipolar Couplings in Macromolecular Structure Determination. *Methods Enzymol.* **2001**, *339*, 127–174.
- (152) Barbieri, R.; Bertini, I.; Cavallaro, G.; Lee, Y.-M.; Luchinat, C.; Rosato, A. Paramagnetically Induced Residual Dipolar Couplings for Solution Structure Determination of Lanthanide Binding Proteins. *J. Am. Chem. Soc.* **2002**, *124*, 5581–5587.
- (153) Vögeli, B.; Kovacs, H.; Pervushin, K. Measurements of Side-Chain ^{13}C - ^{13}C Residual Dipolar Couplings in Uniformly Deuterated Proteins. *J. Am. Chem. Soc.* **2004**, *126*, 2414–2420.
- (154) Williams, R. V.; Yang, J. Y.; Moremen, K. W.; Amster, I. J.; Prestegard, J. H. Measurement of Residual Dipolar Couplings in Methyl Groups via Carbon Detection. *J. Biomol. NMR* **2019**, *73*, 191–198.
- (155) Cook, E. C.; Sahu, D.; Bastidas, M.; Showalter, S. A. Solution Ensemble of the C-Terminal Domain from the Transcription Factor Pdx1 Resembles an Excluded Volume Polymer. *J. Phys. Chem. B* **2019**, *123*, 106–116.
- (156) Balayssac, S.; Bertini, I.; Luchinat, C.; Parigi, G.; Piccioli, M. ^{13}C Direct Detected NMR Increases the Detectability of Residual Dipolar Couplings. *J. Am. Chem. Soc.* **2006**, *128*, 15042–15043.
- (157) Zeng, L.; Fischer, M. W. F.; Zuiderweg, E. R. P. Study of Protein Dynamics in Solution by Measurement of $^{13}\text{C}^{\alpha}$ - ^{13}CO NOE and ^{13}CO Longitudinal Relaxation. *J. Biomol. NMR* **1996**, *7*, 157–162.
- (158) Fischer, M. W. F.; Zeng, L.; Zuiderweg, E. R. P. Use of ^{13}C - ^{13}C NOE for the Assignment of NMR Lines of Larger Labeled Proteins at Larger Magnetic Fields. *J. Am. Chem. Soc.* **1996**, *118*, 12457–12458.
- (159) Pantoja-Uceda, D.; Santoro, J. Selective Observation of Asp and Glu Resonances in ^{13}CO Detected Experiments. *Magn. Reson. Chem.* **2011**, *49*, 558–561.
- (160) Nielsen, M. S.; Vorum, H.; Lindersson, E.; Jensen, P. H. Ca^{2+} Binding to α -Synuclein Regulates Ligand Binding and Oligomerization. *J. Biol. Chem.* **2001**, *276*, 22680–22684.
- (161) Binolfi, A.; Rasia, R. M.; Bertocini, C. W.; Ceolin, M.; Zweckstetter, M.; Griesinger, C.; Jovin, T. M.; Fernández, C. O. Interaction of α -Synuclein with Divalent Metal Ions Reveals Key Differences: A Link between Structure, Binding Specificity and Fibrillation Enhancement. *J. Am. Chem. Soc.* **2006**, *128*, 9893–9901.
- (162) Lautenschläger, J.; Stephens, A. D.; Fusco, G.; Ströhl, F.; Curry, N.; Zacharopoulou, M.; Michel, C. H.; Laine, R.; Nespovitya, N.; Fantham, M.; Pinotsi, D.; Zago, W.; Fraser, P.; Tandon, A.; St George-Hyslop, P.; Rees, E.; Phillips, J. J.; de Simone, A.; Kaminski, C. F.; Schierle, G. S. K. C-Terminal Calcium Binding of α -Synuclein Modulates Synaptic Vesicle Interaction. *Nat. Commun.* **2018**, *9*, 712.
- (163) Banci, L.; Bertini, I.; Cantini, F.; Felli, I. C.; Gonnelli, L.; Hadjiladis, N.; Pierattelli, R.; Rosato, A.; Voulgaris, P. The Atx1-Ccc2 Complex Is a Metal-Mediated Protein-Protein Interaction. *Nat. Chem. Biol.* **2006**, *2*, 367–368.
- (164) Werbeck, N. D.; Kirkpatrick, J.; Hansen, D. F. Probing Arginine Side-Chains and Their Dynamics with Carbon-Detected NMR Spectroscopy: Application to the 42 KDa Human Histone Deacetylase 8 at High PH. *Angew. Chem., Int. Ed.* **2013**, *52*, 3145–3147.
- (165) Gerecht, K.; Figueiredo, A. M.; Hansen, D. F. Determining Rotational Dynamics of the Guanidino Group of Arginine Side Chains in Proteins by Carbon-Detected NMR. *Chem. Commun.* **2017**, *53*, 10062–10065.
- (166) Zhuravleva, A.; Orekhov, V. Yu. Divided Evolution: A Scheme for Suppression of Line Broadening Induced by Conformational Exchange. *J. Am. Chem. Soc.* **2008**, *130*, 3260–3261.
- (167) Mackenzie, H. W.; Hansen, D. F. A ^{13}C -Detected ^{15}N Double-Quantum NMR Experiment to Probe Arginine Side-Chain Guanidinium $^{15}\text{N}^{\eta}$ Chemical Shifts. *J. Biomol. NMR* **2017**, *69*, 123–132.
- (168) Jaipuria, G.; Lobo, N. P.; Shet, D.; Atreya, H. S. High Resolution Methyl Selective ^{13}C -NMR of Proteins in Solution and Solid State. *J. Biomol. NMR* **2012**, *54*, 33–42.
- (169) Baldwin, R. L. The Search for Folding Intermediates and the Mechanism of Protein Folding. *Annual Review of Biophysics* **2008**, *37*, 1–21.
- (170) Pal, D.; Chakrabarti, P. *Cis* Peptide Bonds in Proteins: Residues Involved, Their Conformations, Interactions and Locations. *J. Mol. Biol.* **1999**, *294*, 271–288.
- (171) Dugave, C.; Demange, L. *Cis-Trans* Isomerization of Organic Molecules and Biomolecules: Implications and Applications. *Chem. Rev.* **2003**, *103*, 2475–2532.
- (172) Felli, I. C.; Bermel, W.; Pierattelli, R. Exclusively Heteronuclear NMR Experiments for the Investigation of Intrinsically Disordered Proteins: Focusing on Proline Residues. *Magn. Res.* **2021**, *2*, 511–522.
- (173) Calçada, E. O.; Felli, I. C.; Hošek, T.; Pierattelli, R. The Heterogeneous Structural Behavior of E7 from HPV16 Revealed by NMR Spectroscopy. *ChemBioChem* **2013**, *14*, 1876–1882.
- (174) Hošek, T.; Calçada, E. O.; Nogueira, M. O.; Salvi, M.; Pagani, T. D.; Felli, I. C.; Pierattelli, R. Structural and Dynamic Characterization of the Molecular Hub Early Region 1A (E1A) from Human Adenovirus. *Chem.-Eur. J.* **2016**, *22*, 13010–13013.
- (175) Kumar, A.; Bellstedt, P.; Wiedemann, C.; Wißbrock, A.; Imhof, D.; Ramachandran, R.; Ohlenschläger, O. NMR experiments on the transient interaction of the intrinsically disordered N-terminal peptide of cystathionine- β -synthase with heme. *J. Magn. Reson.* **2019**, *308*, 106561.
- (176) Haba, N. Y.; Gross, R.; Novacek, J.; Shaked, H.; Zidek, L.; Barda-Saad, M.; Chill, J. H. NMR Determines Transient Structure and Dynamics in the Disordered C-Terminal Domain of WASp Interacting Protein. *Biophys. J.* **2013**, *105*, 481–493.
- (177) Elazari-Shalom, H.; Shaked, H.; Esteban-Martin, S.; Salvatella, X.; Barda-Saad, M.; Chill, J. H. New Insights into the Role of the Disordered WIP N-Terminal Domain Revealed by NMR Structural Characterization. *FEBS J.* **2015**, *282* (4), 700–714.
- (178) Gibbs, E. B.; Cook, E. C.; Showalter, S. A. Application of NMR to Studies of Intrinsically Disordered Proteins. *Arch. Biochem. Biophys.* **2017**, *628*, 57–70.
- (179) Gustafson, C. L.; Parsley, N. C.; Asimgil, H.; Lee, H. W.; Ahlback, C.; Michael, A. K.; Xu, H.; Williams, O. L.; Davis, T. L.; Liu, A. C.; Partch, C. L. A Slow Conformational Switch in the BMAL1 Transactivation Domain Modulates Circadian Rhythms. *Mol. Cell* **2017**, *66*, 447–457.
- (180) Parigi, G.; Ravera, E.; Luchinat, C. Magnetic Susceptibility and Paramagnetism-Based NMR. *Prog. Nucl. Magn.* **2019**, *114–115*, 211–236.
- (181) Bertini, I.; Lee, Y. M.; Luchinat, C.; Piccioli, M.; Poggi, L. Locating the Metal Ion in Calcium-Binding Proteins by Using Cerium(III) as a Probe. *ChemBioChem* **2001**, *2*, 550–558.
- (182) Caillat-Saguy, C.; Delepiere, M.; Lecroisey, A.; Bertini, I.; Piccioli, M.; Turano, P. Direct-Detected ^{13}C NMR to Investigate the Iron(III) Hemophore HasA. *J. Am. Chem. Soc.* **2006**, *128*, 150–158.
- (183) Balayssac, S.; Jiménez, B.; Piccioli, M. Assignment Strategy for Fast Relaxing Signals: Complete Amino acid Identification in Thulium Substituted Calbindin D9K. *J. Biomol. NMR* **2006**, *34*, 63–73.
- (184) Abriata, L. A.; Ledesma, G. N.; Pierattelli, R.; Vila, A. J. Electronic Structure of the Ground and Excited States of the CuA Site by NMR Spectroscopy. *J. Am. Chem. Soc.* **2009**, *131*, 1939–1945.
- (185) Pochapsky, S. S.; Sunshine, J. C.; Pochapsky, T. C. Completing the Circuit: Direct-Observe ^{13}C , ^{15}N Double-Quantum Spectroscopy Permits Sequential Resonance Assignments near a Paramagnetic Center in Acireductone Dioxxygenase. *J. Am. Chem. Soc.* **2008**, *130*, 2156–2157.

- (186) Mainz, A.; Bardiaux, B.; Kuppler, F.; Multhaupt, G.; Felli, I. C.; Pierattelli, R.; Reif, B. Structural and Mechanistic Implications of Metal Binding in the Small Heat-Shock Protein AB-Crystallin. *J. Biol. Chem.* **2012**, *287*, 1128–1138.
- (187) Turano, P.; Lalli, D.; Felli, I. C.; Theil, E. C.; Bertini, I. NMR Reveals Pathway for Ferric Mineral Precursors to the Central Cavity of Ferritin. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 545–550.
- (188) Lin, I. J.; Xia, B.; King, D. S.; Machonkin, T. E.; Westler, W. M.; Markley, J. L. Hyperfine-Shifted ^{13}C and ^{15}N NMR Signals from *Clostridium Pasteurianum* Rubredoxin: Extensive Assignments and Quantum Chemical Verification. *J. Am. Chem. Soc.* **2009**, *131*, 15555–15563.
- (189) Cacciatore, S.; Piccioli, M.; Turano, P. Electron Self-Exchange of Cytochrome c Measured via ^{13}C Detected Protonless NMR. *J. Porph. Phthal.* **2013**, *17*, 142–149.
- (190) Pintacuda, G.; John, M.; Su, X. C.; Otting, G. NMR Structure Determination of Protein–Ligand Complexes by Lanthanide Labeling. *Acc. Chem. Res.* **2007**, *40*, 206–212.
- (191) Lindorff-Larsen, K.; Kristjansdottir, S.; Teilum, K.; Fieber, W.; Dobson, C. M.; Poulsen, F. M.; Vendruscolo, M. Determination of an Ensemble of Structures Representing the Denatured State of the Bovine Acyl-Coenzyme A Binding Protein. *J. Am. Chem. Soc.* **2004**, *126*, 3291–3299.
- (192) Bernadó, P.; Bertoncini, C. W.; Griesinger, C.; Zweckstetter, M.; Blackledge, M. Defining Long-Range Order and Local Disorder in Native α -Synuclein Using Residual Dipolar Couplings. *J. Am. Chem. Soc.* **2005**, *127*, 17968–17969.
- (193) Bertoncini, C. W.; Jung, Y. S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. Release of Long-Range Tertiary Interactions Potentiates Aggregation of Natively Unstructured α -Synuclein. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 1430–1435.
- (194) Mittag, T.; Forman-Kay, J. D. Atomic-Level Characterization of Disordered Protein Ensembles. *Curr. Opin. Struct. Biol.* **2007**, *17*, 3–14.
- (195) Kragelj, J.; Blackledge, M.; Jensen, M. R. Ensemble Calculation for Intrinsically Disordered Proteins Using NMR Parameters. *Adv. Exp. Med. Biol.* **2015**, *870*, 23–147.
- (196) Kurzbach, D.; Kontaxis, G.; Coudeville, N.; Konrat, R. NMR Spectroscopic Studies of the Conformational Ensembles of Intrinsically Disordered Proteins. *Adv. Exp. Med. Biol.* **2015**, *870*, 149–185.
- (197) Göbl, C.; Madl, T.; Simon, B.; Sattler, M. NMR Approaches for Structural Analysis of Multidomain Proteins and Complexes in Solution. *Prog. Nucl. Magn.* **2014**, *80*, 26–63.
- (198) Madl, T.; Felli, I. C.; Bertini, I.; Sattler, M. Structural Analysis of Protein Interfaces from ^{13}C Direct-Detected Paramagnetic Relaxation Enhancements. *J. Am. Chem. Soc.* **2010**, *132*, 7285–7287.
- (199) Mateos, B.; Konrat, R.; Pierattelli, R.; Felli, I. C. NMR Characterization of Long-Range Contacts in Intrinsically Disordered Proteins from Paramagnetic Relaxation Enhancement in ^{13}C Direct-Detection Experiments. *ChemBioChem.* **2019**, *20*, 335–339.
- (200) Gibbs, E. B.; Showalter, S. A. Quantification of Compactness and Local Order in the Ensemble of the Intrinsically Disordered Protein FCP1. *J. Phys. Chem. B* **2016**, *120*, 8960–8969.
- (201) Madl, T.; Bermel, W.; Zangger, K. Use of Relaxation Enhancements in a Paramagnetic Environment for the Structure Determination of Proteins Using NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2009**, *48*, 8259–8262.
- (202) Theillet, F. X.; Binolfi, A.; Liokatis, S.; Verzini, S.; Selenko, P. Paramagnetic Relaxation Enhancement to Improve Sensitivity of Fast NMR Methods: Application to Intrinsically Disordered Proteins. *J. Biomol. NMR* **2011**, *51*, 487–495.
- (203) Alik, A.; Bouguechtouli, C.; Julien, M.; Bermel, W.; Ghouil, R.; Zinn-Justin, S.; Theillet, F. Sensitivity-Enhanced ^{13}C -NMR Spectroscopy for Monitoring Multisite Phosphorylation at Physiological Temperature and pH. *Angew. Chem., Int. Ed.* **2020**, *59*, 10411–10415.
- (204) Bertini, I.; Jiménez, B.; Pierattelli, R.; Wedd, A. G.; Xiao, Z. Protonless ^{13}C Direct Detection NMR: Characterization of the 37 KDa Trimeric Protein CutA1. *Proteins* **2008**, *70*, 1196–1205.
- (205) Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. TROSY in Triple-Resonance Experiments: New Perspectives for Sequential NMR Assignment of Large Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 13585–13590.
- (206) Schütz, S.; Sprangers, R. Methyl TROSY Spectroscopy: A Versatile NMR Approach to Study Challenging Biological Systems. *Prog. Nucl. Magn.* **2020**, *116*, 56–84.
- (207) Yamaguchi, Y.; Wälchli, M.; Nagano, M.; Kato, K. A ^{13}C -Detection NMR Approach for Large Glycoproteins. *Carbohydr. Res.* **2009**, *344*, 535–538.
- (208) Schwalbe, H. Editorial: New 1.2 GHz NMR Spectrometers—New Horizons? *Angew. Chem. Int. Ed. - International Edition* **2017**, *56*, 10252.
- (209) Banci, L.; Barbieri, L.; Calderone, V.; Cantini, F.; Cerofolini, L.; Ciofi-Baffoni, S.; Felli, I. C.; Fragai, M.; Lelli, M.; Luchinat, C.; Luchinat, E.; Parigi, G.; Piccioli, M.; Pierattelli, R.; Ravera, E.; Rosato, A.; Tenori, L.; Turano, P. Biomolecular NMR at 1.2 GHz. *arXiv: 1910.07462* **2019**.
- (210) Dyson, H. J.; Wright, P. E. Nuclear Magnetic Resonance Methods for Elucidation of Structure and Dynamics in Disordered States. *Methods Enzymol.* **2001**, *339*, 258–270.
- (211) Pérez, Y.; Gairí, M.; Pons, M.; Bernadó, P. Structural Characterization of the Natively Unfolded N-Terminal Domain of Human c-Src Kinase: Insights into the Role of Phosphorylation of the Unique Domain. *J. Mol. Biol.* **2009**, *391*, 136–148.
- (212) Theillet, F.-X.; Smet-Nocca, C.; Liokatis, S.; Thongwichian, R.; Kosten, J.; Yoon, M.-K.; Kriwacki, R. W.; Landrieu, I.; Lippens, G.; Selenko, P. Cell Signaling, Post-Translational Protein Modifications and NMR Spectroscopy. *J. Biomol. NMR* **2012**, *54*, 217–236.
- (213) Showalter, S. A.; Gibbs, E. B. Phosphorylation Induces Sequence-Specific Conformational Switches in the RNA Polymerase II C-Terminal Domain. *Biophys. J.* **2017**, *112*, 509a.
- (214) Bertini, I.; Felli, I. C.; Gonnelli, L.; Vasantha Kumar, M. V.; Pierattelli, R. ^{13}C Direct-Detection Biomolecular NMR Spectroscopy in Living Cells. *Angew. Chem., Int. Ed.* **2011**, *50*, 2339–2341.
- (215) Mitrea, D. M.; Kriwacki, R. W. Phase Separation in Biology; Functional Organization of a Higher Order. *Cell Commun. Signal.* **2016**, *14*, 1.
- (216) Das, S.; Lin, Y. H.; Vernon, R. M.; Forman-Kay, J. D.; Chan, H. S. Comparative Roles of Charge, π , and Hydrophobic Interactions in Sequence-Dependent Phase Separation of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 28795–28805.
- (217) Forman-Kay, J. D.; Kriwacki, R. W.; Seydoux, G. Phase Separation in Biology and Disease. *J. Mol. Biol.* **2018**, *430*, 4603–4606.
- (218) Elbaum-Garfinkle, S.; Kriwacki, R. W. Phase Separation in Biology & Disease: The next Chapter. *J. Mol. Biol.* **2021**, *433*, 166990.
- (219) Mukrasch, M. D.; Biernat, J.; von Bergen, M.; Griesinger, C.; Mandelkow, E.; Zweckstetter, M. Sites of Tau Important for Aggregation Populate β -Structure and Bind to Microtubules and Polyanions. *J. Biol. Chem.* **2005**, *280* (26), 24978–24986.
- (220) Ambadipudi, S.; Reddy, J. G.; Biernat, J.; Mandelkow, E.; Zweckstetter, M. Residue-Specific Identification of Phase Separation Hot Spots of Alzheimer's-Related Protein Tau. *Chem. Sci.* **2019**, *10*, 6503–6507.
- (221) Kim, T. H.; Tsang, B.; Vernon, R. M.; Sonenberg, N.; Kay, L. E.; Forman-Kay, J. D. Phospho-Dependent Phase Separation of FMRP and CAPRIN1 Recapitulates Regulation of Translation and Deadenylation. *Science* **2019**, *365*, 825–829.
- (222) Stoyas, C. A.; la Spada, A. R. The CAG–Polyglutamine Repeat Diseases: A Clinical, Molecular, Genetic, and Pathophysiologic Nosology. *Handb. Clin. Neurol.* **2018**, *147*, 143–170.
- (223) Escobedo, A.; Topal, B.; Kunze, M. B. A.; Aranda, J.; Chiesa, G.; Mungianu, D.; Bernardo-Seisdedos, G.; Eftekharzadeh, B.; Gairí, M.; Pierattelli, R.; Felli, I. C.; Diercks, T.; Millet, O.; García, J.; Orozco, M.; Crehuet, R.; Lindorff-Larsen, K.; Salvatella, X. Side Chain to Main Chain Hydrogen Bonds Stabilize a Polyglutamine Helix in a Transcription Factor. *Nat. Commun.* **2019**, *10*, 2034.
- (224) Kosol, S.; Contreras-Martos, S.; Piai, A.; Varadi, M.; Lazar, T.; Bekesi, A.; Lebrun, P.; Felli, I. C.; Pierattelli, R.; Tompa, P. Interaction between the Scaffold Proteins CBP by IQGAP1 Provides an Interface

between Gene Expression and Cytoskeletal Activity. *Sci. Rep.* **2020**, *10*, 5753.

(225) Sicorello, A.; Kelly, G.; Oregioni, A.; Nováček, J.; Sklenář, V.; Pastore, A. The Structural Properties in Solution of the Intrinsically Mixed Folded Protein Ataxin-3. *Biophys. J.* **2018**, *115*, 59–71.

(226) Urbanek, A.; Popovic, M.; Morató, A.; Estaña, A.; Elena-Real, C. A.; Mier, P.; Fournet, A.; Allemand, F.; Delbecq, S.; Andrade-Navarro, M. A.; Cortés, J.; Sibille, N.; Bernadó, P. Flanking Regions Determine the Structure of the Poly-Glutamine in Huntingtin through Mechanisms Common among Glutamine-Rich Human Proteins. *Structure* **2020**, *28*, 733–746.

(227) Mier, P.; Elena-Real, C.; Urbanek, A.; Bernadó, P.; Andrade-Navarro, M. A. The Importance of Definitions in the Study of PolyQ Regions: A Tale of Thresholds, Impurities and Sequence Context. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 306–313.