Journal of Magnetic Resonance 306 (2019) 195–201

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

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

Structure determination using solution NMR: Is it worth the effort?

Koh Takeuchi [a], Kumaran Baskaran [b], Haribabu Arthanari [c,d,*]

[a] Molecular Profiling Research Center for Drug Discovery (Molprof), National Institute of Advanced Industrial Science and Technology (AIST), Tokyo 135-0064, Japan
[b] Department of Biochemistry, University of Wisconsin-Madison, 433 Babcock Dr, Madison, WI 53706, United States
[c] Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, United States
[d] Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02115, United States

ARTICLE INFO

Article history:
Received 6 May 2019
Revised 10 July 2019
Accepted 11 July 2019
Available online 13 July 2019

Keywords:
Solution NMR
Structure determination
NOE
PRE/PCS

ABSTRACT

It has been almost 40 years since solution NMR joined X-ray crystallography as a technique for determining high-resolution structures of proteins. Since then NMR derived structure has contributed in fundamental ways to our understanding of the function of biomolecules. With the already existing mature field of X-ray crystallography and the emergence of cryo-EM as techniques to tackle high-resolution structures of large protein complexes, the role of NMR in structure determination has been questioned. However, NMR has the unique ability to recapitulate the dynamic motion of proteins in their structures, while size limitations of the biomolecular systems that can be routinely studied still present challenges. The field has continually developed methodology and instrumentation since its introduction, pushing its frontiers and redefining its limits. Here we present a brief overview of NMR-based structure determination over the past 40 years. We outline the current state of the field and look ahead to the challenges that still need to be addressed to realize the future potential of NMR as a structural technique.

The road travelled and oceans crossed: The solution structure of the proteinase inhibitor IIA, determined in 1985, was one of the first structures determined by NMR [1]. Since then solution NMR has made substantial contributions to atomic resolution structures of biological macromolecules and their complexes. These structures in turn have played key roles in our understanding of the biological functions of these complexes and have allowed them to be targeted therapeutically. In addition to providing structural information, NMR is uniquely poised to provide information on inherent molecular motions, or dynamics, which also influence function. NMR structures are unique in comparison to structures determined by other techniques, such as X-ray crystallography and cryo-electron microscopy (cryo-EM), because it is the only technique that allows high-resolution structure determination of biomolecules in solution, at physiological temperatures and without disruption or modification of endogenous conformations. Though the measurement conditions in solution NMR may not be identical to its native environment, inside a cell, for example, measurements can often be made in near-to biologically relevant conditions. In special cases, NMR has even been used for structure determination and characterization of biomolecules inside living mammalian cells [2–5].

The history of NMR has witnessed periodic developments in methodology, including isotopic labeling and instrumentation, which have propelled the field forward and redefined its limits, particularly when it comes to the molecular weight of the systems that can be addressed. Early in its development, NMR was not thought of as a method for atomic-resolution structure determination. This attitude was changed by the determination of two structures, those of proteinase inhibitor IIA and bovine pancreatic trypsin inhibitor (BPTI), in the 1980s [1,6]. Although these proteins are small (<10 kDa) by modern-day standards, the nuclear-Overhauser-effect (NOE)- and three-bond-coupling-based strategy developed for these earliest structures is still an integral part of structure determination by NMR today. The unique nature of the structural information derived from NMR was evident from the very beginning. The observation of aromatic side chain dynamics, which were only possible due to structural *breathing*, fundamentally changed the perception of proteins as rigid molecules [7].

Since then, continuous development of new NMR methods and technologies, including advances in protein expression, isotopic labeling, resonance assignment, structure determination approaches, and detection hardware, has pushed the limits of biomolecular NMR to complex biological systems with higher molecular weights. The NMR constraints used to solve macromolecular structures are no longer limited to the classical NOE-derived distance constraints and J-coupling-derived dihedral

* Corresponding author at: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, United States.
E-mail address: hari@hms.harvard.edu (H. Arthanari).

https://doi.org/10.1016/j.jmr.2019.07.045
1090-7807/

constraints. New experimental methods now provide an array of constraints for structure determination, including long-range distance and angle constraints from paramagnetic relaxation enhancement (PRE), pseudo-contact shift (PCS), and residual dipolar coupling (RDC) measurements (Fig. 1). NMR structures and their associated resonance assignments are collected and shared by the NMR community via the protein data bank (PDB) and the biological magnetic resonance bank (BMRB), respectively. This body of data has allowed the development of empirical strategies such as TALOS [8], CS23D [9], and CS-ROSETTA [10,11] to determine macromolecular structures more rapidly and efficiently. In the era of big data, with the development of high-throughput genome sequencing and structural bioinformatics, sequence-based methods for determining constraints, such as the consideration of residue co-evolution, have proven to be useful in NMR structure determination [12]. NMR-derived constraints can be integrated with data from other structural methods including X-ray crystallography, cryo-EM, small-angle X-ray and neutron scattering (SAXS and SANS, respectively), and electron paramagnetic resonance (EPR), thereby leveraging the individual strengths of different methods. Here we present a brief overview of the methodological developments in NMR structure determination over the past 40 years and their implications; the current state of the field and its limitations; and the outlook for future developments and challenges. Although NMR has played a pivotal role in structure determination of nucleic acids and carbohydrates, for the sake of brevity, this report will be limited to the discussion of structure determination of proteins.

State of the Structural Union: As of April 2019, there are 12,591 NMR structures available in the PDB database. This is far less than the number of structures determined by X-ray crystallography (135,258), and exceeds the 3094 deposited structures determined by EM (Fig. 2A). However, the fraction of structures that are *unique*, (i.e., structures with less than 90% identity to all the other structures determined by that method), is much larger for the NMR structures than for the X-ray crystallography or EM structures (Fig. 2A). There are 9,442 *unique* deposited NMR structures, which corresponds to 75% of the total number of structures determined by NMR. By comparison, the deposited X-ray crystallography and EM structures are only 33% and 52% *unique*, respectively. This could be a result of the fact that molecular replacement is not commonly used in NMR. Furthermore, 6,250 (66%) of the *unique* structures that have been determined by NMR have not been determined using any other structural methods. These statistics demonstrate that NMR has been instrumental in determining the structures of macromolecules, and thus illuminating key structure-function relationships, not readily accessible to other techniques.

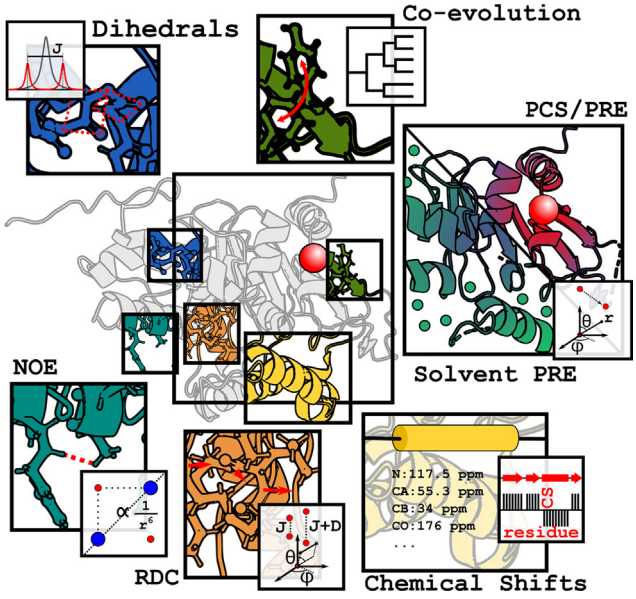

Fig. 1. Schematic representation of various structural constraints experimentally observed or derived from NMR experiments that drive NMR based structure determination of proteins. A cartoon representation of the structural constraint in relation to the structure (large inset) and the NMR data (small inset) is shown.

The percentage of human proteins amongst the NMR-determined structures is also larger when compared to other structural methods. 32% of *unique* protein structures determined by NMR are of human origin, whilst only 18% and 20% of X-ray crystallography and EM structures, respectively, are of human proteins (Fig. 2B). Although human proteins dominate the structures determined by all three major techniques, *E. coli* proteins are the second most frequently determined for X-ray crystallography (6% of all X-ray crystallography structures), whereas for NMR, mouse proteins are the second most common. Thus, while the total number of structures determined by NMR may be smaller than that determined by X-ray crystallography, the contribution of NMR to our understanding of human and vertebrate proteins, and therefore to biomedical research, including the development of therapeutics, should not be underestimated.

Challenges: Major challenges limiting the study of larger proteins and protein systems by NMR are the faster transverse relaxation rates of larger systems, which results in broader linewidths, and overall spectral crowding, due the proportional increase in the number of resonances observed. As a result, structure determination by NMR has typically been limited to proteins smaller than 30 kDa. Before 1994, the average molecular weight (MW) of the structures determined by NMR was 7.4 kDa and there was no structure larger than 30 kDa; however, since then, the average MW has been steadily increasing (Fig. 3A). Actually, the average MW does not accurately reflect the ability of NMR to address large systems. The average MW of the top 5% of deposited structures (by MW) has increased from 17.4 kDa before 1994 to 40.3 kDa in 2019 (Fig. 3B). In the same time, the record of the largest MW structure determined by solution NMR has increased exponentially (Fig. 3C). Although these are by no means commonplace, these represent herculean efforts to tackle proteins larger than 100 kDa by NMR.

Looking under the hood: In this section, we discuss the number and type of constraints that have been used in NMR-based structure determination. NOEs, particularly between hydrogen atoms, have provided the majority of constraints used in NMR structures. Traditionally, the rule of thumb is that approximately 10 constraints per amino acid are needed to determine a high-resolution structure of a protein. According to the BMRB database, for the structure of proteins where the molecular weight is less than 20 kDa, approximately 13 NOE constraints per residue have been used (Fig. 4A). Most of these constraints are obtained from ^{15}N-edited and/or ^{13}C-edited 3D/4D NOESY experiments. The rate-limiting step in determining these type of distance constraints is the assignment of the NOE crosspeaks as the complexity increases non-linearly with the size of the protein. For proteins larger than 30 kDa, however, the number of NOEs per residue drops from 13 to below 4.

The double-edged sword: Two key developments that complement each other propelled the application of NMR to large systems; these were the development of protein deuteration [13,14] and transverse relaxation-optimized spectroscopy (TROSY) [15–17]. Leveraging the TROSY effect helps to counteract the broader linewidths that result from faster transverse relaxation times in larger systems, and both the amide and methyl TROSY effects produce the narrowest lines when the remainder of the protein is perdeuterated except the observing sites. However, perdeuterated

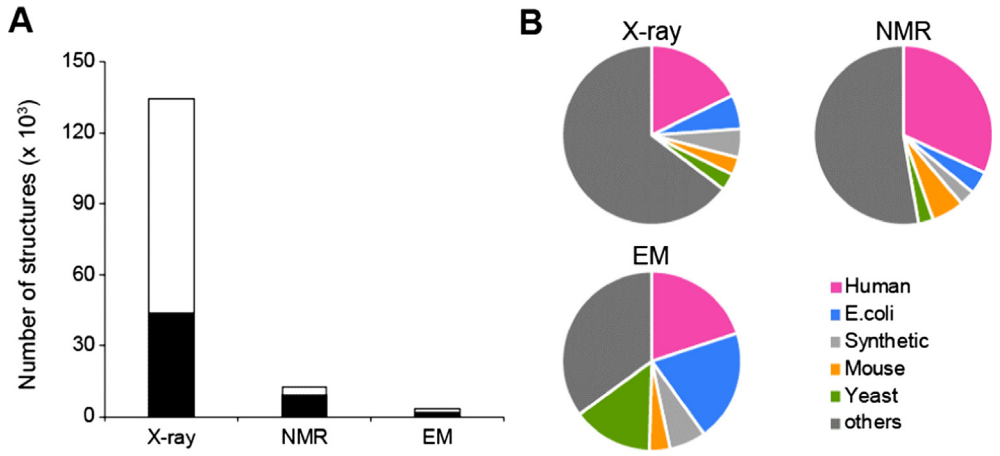

Fig. 2. Numbers and nature of structures determined by three structural methods. (A) Total number (white) and number of *unique* (black) structures determined by each method. (B) The distribution of structures determined by each method among different species.

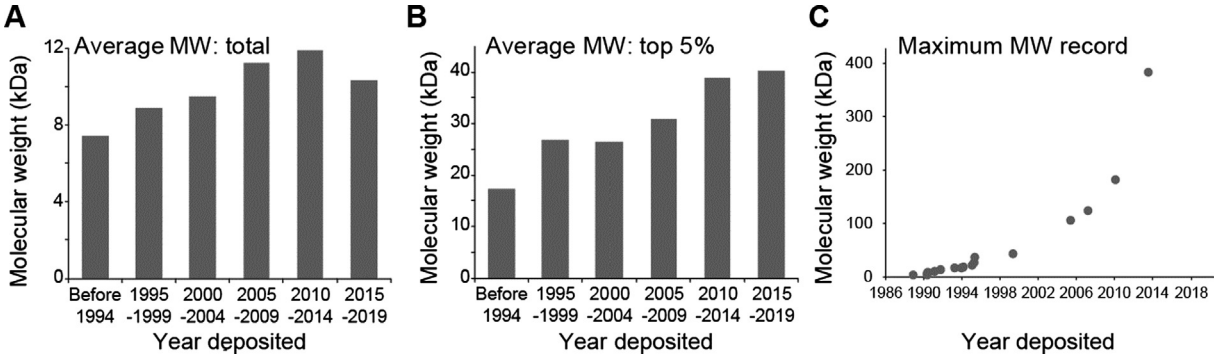

Fig. 3. Molecular weight (MW) of NMR structures in the PDB. (A) The average MW of *unique* NMR structures deposited in the PDB as of the indicated years. (B) The average MW of top 5% NMR structures as of the indicated years. (C) The maximum MW structure solved by solution NMR by year.

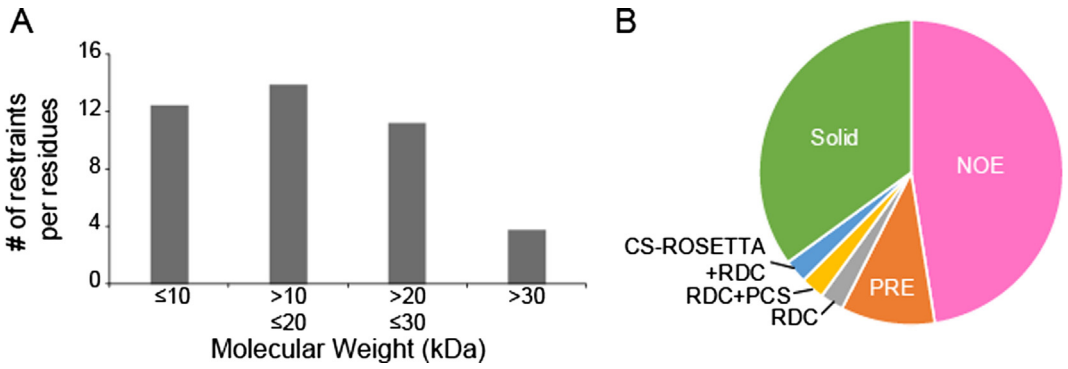

Fig. 4. Information used to drive NMR-based structure determinations. (A) The average number of NOE constraints per residue used structure determination as a function of the MW range of the protein. (B) The type of structural constraints used in the NMR-based structures with MW greater than 50 kDa.

proteins lack hydrogen atoms, which normally provide NOEs, making it increasingly difficult to obtain sufficient numbers of NOE-based distance constraints, especially weak long-range constraints. Therefore, for large perdeuterated proteins, other types of constraints aid structure determination (Fig. 4B).

The proven path: For large proteins, NOEs from methyl groups, aromatic residues, and amides, in an otherwise deuterated background, provide critical constraints for structure determination. The amide [15], methyl [17], and aromatic [16] TROSY effects can be taken advantage of to create narrow resonances. The methyl groups of proteins are particularly helpful as they are three-fold symmetric, and the rapid rotation around the connecting C—C bond mean that all three hydrogens contribute to a single peak, and the rapid rotation offsets some of the size-dependent relaxation loses. In addition, interference between the multiple $^{1}H—^{1}H$ and $^{1}H—^{13}C$ dipolar interactions make certain $^{13}C—^{1}H_3$ coherences to relax much slower than others. In fact, methyl resonances can be clearly observed, even in a megadalton-sized protein [18,19]. Currently, the following amino acids can be site-specifically and stereo-specifically ^{13}C- and ^{1}H-labeled at the terminal position of the amino acid with the remainder of the amino acid deuterated: Ala, Ile, Leu, Met, Thr, and Val. Time-shared 3D and 4D NOESY

experiments with TROSY selection can be used to simultaneously observe NOEs from both amide and methyl resonances [20,21]. Although the aromatic TROSY effect can be leveraged to provide additional NOE constraints, especially to better define a hydrophobic core, it has not been utilized to its full potential. One reason for this is the large $^{13}C—^{13}C$ coupling in aromatic rings (~60 Hz), which limits the resolution in the ^{13}C dimension. Recent developments, which allow the incorporation of aromatic amino acids that have isolated $^{13}C—^{1}H$ bonds [22], may mean that the aromatic TROSY effect may be applied more often in future efforts. Are the NOEs between methyls, aromatics, and amides hydrogens sufficient to determine a *de novo* structure of large proteins remains an unanswered question. With reliable NOE back-calculation programs that are available today this problem can be addressed, theoretically, using existing structures and assignments.

And the Oscars go to: In this section, we spotlight some of the high MW structures determined by NMR as examples of the current state of the art, and to highlight the future possibilities.

Category - Largest structure by NOE-based methods: The largest structure that has been determined by NOE-based methods is the structure of the chaperone SecB in complex with unstructured proPhoA (PDB ID 5JTL; [23]). The total MW of the complex, including unstructured regions, is 119 kDa. The structure was calculated with CYANA [24], using the 7,530 methyl-, aromatic-, and amide-derived NOE constraints measured from 3D NOESY experiments, and 5,193 main-chain dihedral angle restraints (φ and ψ) obtained using the TALOS+ program [8], based on the $^{13}C\alpha$, $^{13}C\beta$, $^{13}C'$, ^{15}N, and $^{1}H_N$ chemical shifts. The structure itself is actually a hybrid structure as the SecB regions remote to the binding sites were kept rigid using the crystal structure coordinates for *E. coli* SecB, while the SecB side chains at or proximal to the interface with proPhoA were allowed to be flexible and their conformation was determined using intermolecular NOEs. Due to the nature of interaction between the molecular chaperone and its substrate, a significant part of the substrate structure remains flexible in the final coordinates, while the chaperone-substrate interfaces are well defined. This structure is also an excellent example of the ability of NMR to determine the structure of complexes with substantial flexibility.

Category - Largest structure by NMR, with support from SAXS/SANS: Currently, the largest structure determined by NMR is the box C/D ribonucleoprotein enzyme, which methylates ribosomal RNA at the 2′-O-ribose (PDB ID 4BY9; [25]). This 390 kDa archaeal RNP enzyme bound to substrate RNA was determined by a combination of NMR spectroscopy and small-angle neutron and X-ray scattering (SANS and SAXS). Sample concentrations were reportedly between 7 and 40 μM, and for structure determination, PRE-based distance constraints were used in combination with multiple SANS experiments. Taking advantage of the long-range (up to 25 Å) distance information provided by PREs, 512 distance restraints were obtained for 8 different spin labeling sites and the SANS/SAXS data was used to select structures that agreed reasonably well with all the curves. The structure calculation was done using CNS [26]. The approach is quite distinct from the classic NOE-based structure determination strategy, which utilizes the network of the short range (typically up to 5 Å) interactions between atoms, and this example represents an integrated strategy for structure determination of large proteins.

Honorable mentions: There are currently 40 NMR structures with MW greater than 50 kDa. Of these structures, 14 (35%) have been determined by solid-state NMR, indicating the strength of the technique for large MW complexes, including fibrils. Among these 14 structures, 11 structures were determined solely using solid-state NMR data, 2 structures were determined by combining NMR data with cryo-EM and one with X-ray crystallography.

19 of the 40 solution NMR structures (48%) were determined using NOE-based strategies. While the strategy that utilizes dihedral angle restraints (φ and ψ) obtained using the TALOS+ program were mostly used in combination with NOE distance constraints, 2 structures used RDC, which provide orientational constraints for bond angles relative to the RDC tensor. At this MW (>50 kDa), amide-amide, amide-methyl, and methyl-methyl NOEs in combination with TROSY-based experiments have been used extensively to obtain structure information. There are also two hybrid structures that make use of X-ray crystallography structures, and one that incorporates negative stain EM data. In the case of the X-ray structures, the X-ray data was used to define the rigid core structure of multidomain proteins or even protein complexes. Sites that did not show chemical shift perturbations upon complex formation or multimerization were defined using the crystal structure coordinates [23,27]. Even low-resolution EM data can be used to determine the oligomeric state [28] or the relative positioning of domains in large protein complexes [29].

Other structures have been determined using PCS, PRE, and/or RDC data with X-ray crystallography, SANS, and/or SAXS data to compensate for insufficient NMR constraints [25,30]. SANS and SAXS data mainly contribute to defining the overall envelope of structures, and are complementary to the short range or local conformation information derived from NMR. It should be noted that the approach to combine NMR data with those from SAXS/SANS was introduced not only for the structural determination of large protein systems but also to define the quaternary structures of the dynamic multidomain systems and to assess the accuracy of the determined structures [31–34], which has been extensively reviewed [35]. Currently there are no structures of molecular weight more 50 kDa that are determined by a combination of NMR and EPR data. However, the potential of NMR data in combination with EPR-based techniques such as double-electron-elec tron-resonance (DEER) should be mentioned. EPR data can provide distance constraints in the 10–70 Å range, which can be used as additional constraints for structure determination, for structure validation, and/or for selecting candidate structures from an ensemble [36–38].

The new kid in the block: It should be noted that there are other types of constraints, besides those described above, that can also be used. Solvent PREs (sPRE) are one such type of constraint that can be combined with limited NOE data. NOEs between backbone amide protons, RDCs, and dihedral angles derived from chemical shift information were used together with sPRE-derived constraints to rapidly obtain the structure of MBP with an approximately 1 Å RMSD [39]. The sPRE-based refinement strategy can also be used to define the relative orientation of protein complexes [40] and protein multimers [41]. sPRE data has also been shown to improve the accuracy of structures predicted by CS-ROSETTA [42].

As demonstrated by the examples above, the molecular weight limit of biomolecular NMR continues to increase and has even exceeded 100 kDa at its uppermost limit. The structural information that is used to tackle larger molecular weight proteins is not just from classical NOE experiments, but also from other measurements such as PREs, PCSs, and RDCs. These latter strategies make use of the sensitive 2D experiments, increasing the size of the proteins that can be studied using these structure determination methods.

What is in a name? Do we need to make a distinction as to what constitutes a structure and what constitutes a model? Aren't all NMR structures models that fit the experimental data? Though these questions border on philosophy, for large proteins, often, we depend on existing high-resolution structures for parts of the protein and use long-range constraints derived from NMR to determine the structure. This is different from the case where the entire structure is determined solely using NMR derived

constraints. However the structural information from either effort does not diminish the contribution of the structure to the understanding of the biology, but one of them requires the existence of high-resolution structure for parts of the protein. It should be also mentioned that insufficient structural information can lead to incorrect a structural model, one should carefully check the correctness of the structures by orthogonal structural method that are not used in the structure determination or functional study such as site directed mutagenesis. Thus, the assessment of the quality of the model needs to be considered more seriously, along with the relevance of showing a small or single set of structural models as representative. This is especially true for systems with functional dynamics, and the unique ability of NMR to access highly dynamic systems means that an ensemble of structures should be considered [43], and not waived in exchange for a static structural model.

The Wall and the Wishful Crystal ball: With ease of solving structures by X-Ray crystallography ("crystallography by FedEx") and recent advent of cryo-EM, there is a feeling that, in future, NMR will have less of a role to play in structure determination of proteins. Though we partly empathize with this pessimistic thought, we think that this opinion is rather misplaced. One has to agree that the NMR structure determination process is far for being fully automated; it requires higher concentrations of proteins compared to X-Ray crystallography and cryo-EM; NMR data acquisition times are in order of weeks, which require the proteins to be stable for long periods of time. For large protein systems, deuteration is essential and this limits the expression system to *E. coli*, though there are a few examples of deuteration in eukaryotic expression systems [44–46]. For large systems the methyls, aromatics, and amides serve as beacons providing distance constraints, however, assignments of these resonances are still a challenge. The methyl resonance for the two largest systems studied by NMR the proteasome and the Hsp-60, where accomplished by mutations. These challenges demand new NMR methods, labeling strategies, and better instrumentation (probe technology) that can increase the sensitivity of NMR experiments.

There are a few new technologies that are at the horizon, which show promise.

(1) Assignment of methyl side chains from NOESY experiments: There are numbers of attempts to assign the methyl resonances from 3D/4D ^{13}C-edited NOESY experiments using computational approaches [47–51]. These NOESY experiments would be less demanding in terms of relaxation, as they involve short transfers and a mixing time period where the magnetization is longitudinal, thus would work for large systems. However, to obtain the resonance assignment by computational methods one needs a good homology model or a related structure to begin with.
(2) Low-γ detection: ^{13}C and ^{15}N detection experiments would be a good alternative for larger protein systems [52–56]. These low γ nuclei, under the right experimental conditions, relax slowly, enabling detection of large systems, but they suffer from low intrinsic sensitivity. The TROSY effect on ^{15}N does not depend deuteration and hence can be leverage for proteins made in eukaryotic expression systems [56,57]. But new NMR methods that leverage these slow relaxing coherences should be developed. If the probe technology can increase the sensitivity of these nuclei by a factor of 10, or solution DNP becomes a reality, NMR would get a new pair of glasses, with which we can look beyond, like never before. ^{13}C–^{13}C NOESY could be alternative to provide structural constraints for large proteins in a deuterated environment [58].
(3) Resolution and Non-Uniform Sampling (NUS): Large protein systems require high-resolution to offset spectral crowding, which is made possible by NUS methods. Better reconstruction methods, which permit sampling a lower percentage of the grid and the ability to reconstruct data with higher fidelity, with respect to peak position and intensity, will greatly assist NMR of large systems. This is especially important in NOESY experiments, which have a large dynamic range in their resonances' intensities, which directly provide the distance constraints.
(4) Labeling Technology: New labeling technology that site selectively introduce isotopic labels in deuterated background will significantly advance the field. SAIL labeling is a great example of this idea where the minimal hydrogens are introduced to give the maximum information [59]. New labeling methods should be affordable, easily adaptable to various expression systems and should not reduce the yield in protein expression. PRE and PCS derived constraints are critical driving structure determination of large molecular weight systems. Here probes that are rigid by themselves and closely resemble native amino acids would provide more accurate constraints. Recent work on ^{13}C-^{19}F TROSY [60] can be leveraged to obtain ^{1}H -^{19}F heteronuclear NOE (HOESY) that will provide unique constraints for structure determination.
(5) Automated resonance assignment and structure calculation tools: In order to scale and bring NMR based structure determination to non-experts, software packages that seamlessly process, assign and analyze NMR data are required. Though there have been a number of significant advances in this front, we still require human intervention for large systems. There wealth of knowledge in digital signal processing, pattern matching and machine learning that NMR spectroscopists can leverage for automation. A workflow platform that can seamlessly integrate NMR derived experimental data, with data from other experimental methods as well as empirical data and evolutionary constraints will have a significant impact.

Though these methods have the potential to overcome some of the roadblocks in structure determination of large systems, newer methods are required in four fronts in NMR methodology; labeling; data acquisition and automation; and instrumentation to propel NMR to access larger systems routinely.

So, why bother with NMR based structure determination? Proteins are moving breathing entities, which often change shape and form to perform function. Though X-Ray crystallography and cryo-EM provide awe-inspiring structures at a remarkably high resolution, NMR often provides the missing pose that is critical to tell a whole story. This is especially true for modular or multi-domain proteins, where orchestrated structural rearrangement is an essential part of the function. NMR is uniquely capable of obtaining structural information on lowly populated *invisible* states of proteins [61]. The synergy and complementarity between the different structural methods is another significant aspect to consider, where NMR can provide vital structural information about the flexible regions. This combination static and dynamic views of structure reveal how proteins truly function and interact in the nonequilibrium environment of living cellular systems and contribute to our global understanding of the molecular milieu. The real power of NMR is in observing and characterizing dynamics over a large range of time scales and determining structures of complexes involving weak and transient interactions inaccessible to other structural methods. Given the challenges in NMR based structure determination, especially for large molecular weight systems, one

should embark on NMR based structure determination only if the system investigated cannot be addressed by X-Ray or cryo-EM, or if there is sufficient evidence that the existing structures does not capture a functional form of the protein. The advantages of complimentary structural methods should be embraced, leveraged and combined with power and uniqueness of NMR derived structural information. When it comes to structure determination of large systems by NMR, the road is not easy, not yet, but choose the journey where effort is worth the adventure and a glimpse at rare vistas. For some games you might get a seat in the bleachers, yet you get to watch a live game and that, makes all the difference.

Acknowledgements

We thank Kendra E. Leigh, Abhinav Dubey and Thibault Viennet for their valuable input and discussions. We thank Rhea Montgomery-Walsh for help with the graphics. HA would like to thank the Claudia Adams Barr Program for Innovative Cancer Research and the Louis B Mayer Foundation for their financial support.

References

[1] M.P. Williamson, T.F. Havel, K. Wüthrich, Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry, J. Mol. Biol. 182 (1985) 295–315.
[2] D. Sakakibara, A. Sasaki, T. Ikeya, J. Hamatsu, T. Hanashima, M. Mishima, M. Yoshimasu, N. Hayashi, T. Mikawa, M. Walchli, B.O. Smith, M. Shirakawa, P. Guntert, Y. Ito, Protein structure determination in living cells by in-cell NMR spectroscopy, Nature 458 (2009) 102–105.
[3] K. Inomata, A. Ohno, H. Tochio, S. Isogai, T. Tenno, I. Nakase, T. Takeuchi, S. Futaki, Y. Ito, H. Hiroaki, M. Shirakawa, High-resolution multi-dimensional NMR spectroscopy of proteins in human cells, Nature 458 (2009) 106–109.
[4] F.-X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H.M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb, P. Selenko, Structural disorder of monomeric α-synuclein persists in mammalian cells, Nature 530 (2016) 45–50.
[5] S. Ogino, S. Kubo, R. Umemoto, S. Huang, N. Nishida, I. Shimada, Observation of NMR signals from proteins introduced into living mammalian cells by reversible membrane permeabilization using a pore-forming toxin, streptolysin O, J. Am. Chem. Soc. 131 (2009) 10834–10835.
[6] G. Wagner, W. Braun, T.F. Havel, T. Schaumann, N. Go, K. Wuthrich, Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN, J. Mol. Biol. 196 (1987) 611–639.
[7] G. Wagner, A. DeMarco, K. Wuthrich, Dynamics of the aromatic amino acid residues in the globular conformation of the basic pancreatic trypsin inhibitor (BPTI). I. ^{1}H NMR studies, Biophys. Struct. Mech. 2 (1976) 139–158.
[8] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, J. Biomol. NMR 44 (2009) 213–223.
[9] D.S. Wishart, D. Arndt, M. Berjanskii, P. Tang, J. Zhou, G. Lin, CS23D: a web server for rapid protein structure generation using NMR chemical shifts and sequence data, Nucleic Acids Res. 36 (2008) W496–W502.
[10] Y. Shen, O. Lange, F. Delaglio, P. Rossi, J.M. Aramini, G. Liu, A. Eletsky, Y. Wu, K. K. Singarapu, A. Lemak, A. Ignatchenko, C.H. Arrowsmith, T. Szyperski, G.T. Montelione, D. Baker, A. Bax, Consistent blind protein structure generation from NMR chemical shift data, Proc. Natl. Acad. Sci. 105 (2008) 4685–4690.
[11] Y. Shen, R. Vernon, D. Baker, A. Bax, De novo protein structure generation from incomplete chemical shift assignments, J. Biomol. NMR 43 (2009) 63–78.
[12] Y. Tang, Y.J. Huang, T.A. Hopf, C. Sander, D.S. Marks, G.T. Montelione, Protein structure determination by combining sparse NMR data with evolutionary couplings, Nat. Methods 12 (2015) 751–754.
[13] D.M. LeMaster, Deuteration in protein proton magnetic resonance, Methods Enzymol. 177 (1989) 23–43.
[14] M.A. Markus, K.T. Dayie, P. Matsudaira, G. Wagner, Effect of deuteration on the amide proton relaxation rates in proteins. Heteronuclear NMR experiments on villin 14T, J. Magn. Reson. B 105 (1994) 192–195.
[15] K. Pervushin, R. Riek, G. Wider, K. Wuthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, Proc. Natl. Acad. Sci. USA 94 (1997) 12366–12371.
[16] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in 13C-Labeled Proteins, J. Am. Chem. Soc. 120 (1998) 6394–6400.
[17] V. Tugarinov, P.M. Hwang, J.E. Ollerenshaw, L.E. Kay, Cross-correlated relaxation enhanced ^{1}H-^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes, J. Am. Chem. Soc. 125 (2003) 10420–10428.
[18] A.M. Ruschak, L.E. Kay, Proteasome allostery as a population shift between interchanging conformers, Proc. Natl. Acad. Sci. U S A 109 (2012) E3454–E3462.
[19] G. Mas, J.-Y. Guan, E. Crublet, E.C. Debled, C. Moriscot, P. Gans, G. Schoehn, P. Macek, P. Schanda, J. Boisbouvier, Structural investigation of a chaperonin in action reveals how nucleotide binding regulates the functional cycle, Sci. Adv. 4 (2018), eaau4196.
[20] S.H. Mishra, B.J. Harden, D.P. Frueh, A 3D time-shared NOESY experiment designed to provide optimal resolution for accurate assignment of NMR distance restraints in large proteins, J. Biomol. NMR 60 (2014) 265–274.
[21] D.P. Frueh, A. Leed, H. Arthanari, A. Koglin, C.T. Walsh, G. Wagner, Time-shared HSQC-NOESY for accurate distance constraints measured at high-field in ^{15}N–^{13}C-ILV methyl labeled proteins, J. Biomol. NMR 45 (2009) 311–318.
[22] A.G. Milbradt, H. Arthanari, K. Takeuchi, A. Boeszoermenyi, F. Hagn, G. Wagner, Increased resolution of aromatic cross peaks using alternate C-13 labeling and TROSY, J. Biomol. NMR 62 (2015) 291–301.
[23] C. Huang, P. Rossi, T. Saio, C.G. Kalodimos, Structural basis for the antifolding activity of a molecular chaperone, Nature 537 (2016) 202–206.
[24] P. Güntert, Automated NMR Structure Calculation With CYANA, in: A.K. Downing (Ed.), Protein NMR Techniques, Humana Press, Totowa, NJ, 2004, pp. 353–378.
[25] A. Lapinaite, B. Simon, L. Skjaerven, M. Rakwalska-Bange, F. Gabel, T. Carlomagno, The structure of the box C/D enzyme reveals regulation of RNA methylation, Nature 502 (2013) 519–523.
[26] A.T. Brunger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Crystallography & NMR system: A new software suite for macromolecular structure determination, Acta Crystallogr. D Biol. Crystallogr. 54 (1998) 905–921.
[27] K. Takeuchi, M.H. Roehrl, Z.Y. Sun, G. Wagner, Structure of the calcineurin-NFAT complex: defining a T cell activation switch using solution NMR and crystal coordinates, Structure 15 (2007) 587–597.
[28] K. Oxenoid, Y. Dong, C. Cao, T. Cui, Y. Sancak, A.L. Markhard, Z. Grabarek, L. Kong, Z. Liu, B. Ouyang, Y. Cong, V.K. Mootha, J.J. Chou, Architecture of the mitochondrial calcium uniporter, Nature 533 (2016) 269–273.
[29] O.R. Lorenz, L. Freiburger, D.A. Rutz, M. Krause, B.K. Zierer, S. Alvira, J. Cuéllar, José M. Valpuesta, T. Madl, M. Sattler, J. Buchner, Modulation of the Hsp90 chaperone cycle by a stringent client protein, Mol. Cell 53 (2014) 941–953.
[30] V. Venditti, C.D. Schwieters, A. Grishaev, G.M. Clore, Dynamic equilibrium between closed and partially closed states of the bacterial Enzyme I unveiled by solution NMR and X-ray scattering, Proc. Natl. Acad. Sci. 112 (2015) 11565–11570.
[31] M.-L. Mattinen, K. Pääkkönen, T. Ikonen, J. Craven, T. Drakenberg, R. Serimaa, J. Waltho, A. Annila, Quaternary structure built from subunits combining NMR and small-angle X-ray scattering data, Biophys. J. 83 (2002) 1177–1183.
[32] M. Sunnerhagen, G.A. Olah, J. Stenflo, S. Forsén, T. Drakenberg, J. Trewhella, The relative orientation of Gla and EGF domains in coagulation factor X Is altered by Ca2+ binding to the first EGF domain. A combined NMR−small angle X-ray scattering study, Biochemistry 35 (1996) 11547–11559.
[33] A. Grishaev, J. Wu, J. Trewhella, A. Bax, Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data, J. Am. Chem. Soc. 127 (2005) 16621–16628.
[34] M. Falb, I. Amata, F. Gabel, B. Simon, T. Carlomagno, Structure of the K-turn U4 RNA: a combined NMR and SANS study, Nucleic Acids Res. 38 (2010) 6274–6285.
[35] T. Madl, F. Gabel, M. Sattler, NMR and small-angle scattering-based structural analysis of protein complexes in solution, J. Struct. Biol. 173 (2011) 472–482.
[36] Y. Yang, T.A. Ramelot, R.M. McCarrick, S. Ni, E.A. Feldmann, J.R. Cort, H. Wang, C. Ciccosanti, M. Jiang, H. Janjua, T.B. Acton, R. Xiao, J.K. Everett, G.T. Montelione, M.A. Kennedy, Combining NMR and EPR methods for homodimer protein structure determination, J. Am. Chem. Soc. 132 (2010) 11910–11913.
[37] R. Ward, M. Zoltner, L. Beer, H. El Mkami, I.R. Henderson, T. Palmer, D.G. Norman, The orientation of a tandem POTRA domain pair, of the beta-barrel assembly protein BamA, determined by PELDOR spectroscopy, Structure 17 (2009) 1187–1194.
[38] Z.-Y.J. Sun, K.J. Oh, M. Kim, J. Yu, V. Brusic, L. Song, Z. Qiao, J.-H. Wang, G. Wagner, E.L. Reinherz, HIV-1 broadly neutralizing antibody extracts its epitope from a kinked gp41 ectodomain region on the viral membrane, Immunity 28 (2008) 52–63.
[39] T. Madl, W. Bermel, K. Zangger, Use of relaxation enhancements in a paramagnetic environment for the structure determination of proteins using NMR spectroscopy, Angew. Chem. Int. Ed. 48 (2009) 8259–8262.
[40] T. Madl, T. Güttler, D. Görlich, M. Sattler, Structural analysis of large protein complexes using solvent paramagnetic relaxation enhancements, Angew. Chem. Int. Ed. 50 (2011) 3993–3997.
[41] Y. Wang, C.D. Schwieters, N. Tjandra, Parameterization of solvent–protein interaction and its use on NMR protein structure determination, J. Magn. Reson. 221 (2012) 76–84.
[42] C. Hartlmüller, C. Göbl, T. Madl, Prediction of protein structure using surface accessibility data, Angew. Chem. Int. Ed. 55 (2016) 11970–11974.
[43] M.R. Jensen, M. Zweckstetter, J.R. Huang, M. Blackledge, Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy, Chem. Rev. 114 (2014) 6632–6660.

[44] W.D. Morgan, A. Kragt, J. Feeney, Expression of deuterium-isotope-labelled protein in the yeast pichia pastoris for NMR studies, J. Biomol. NMR 17 (2000) 337–347.
[45] M. Miyazawa-Onami, K. Takeuchi, T. Takano, T. Sugiki, I. Shimada, H. Takahashi, Perdeuteration and methyl-selective (1)H, (13)C-labeling by using a Kluyveromyces lactis expression system, J. Biomol. NMR 57 (2013) 297–304.
[46] Y. Kofuku, T. Yokomizo, S. Imai, Y. Shiraishi, M. Natsume, H. Itoh, M. Inoue, K. Nakata, S. Igarashi, H. Yamaguchi, T. Mizukoshi, E.-I. Suzuki, T. Ueda, I. Shimada, Deuteration and selective labeling of alanine methyl groups of β2-adrenergic receptor expressed in a baculovirus-insect cell expression system, J. Biomol. NMR 71 (2018) 185–192.
[47] F. Hagn, G. Wagner, Structure refinement and membrane positioning of selectively labeled OmpX in phospholipid nanodiscs, J. Biomol. NMR 61 (2015) 249–260.
[48] V. Venditti, N.L. Fawzi, G.M. Clore, Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear Overhauser enhancement spectroscopy, J. Biomol. NMR 51 (2011) 319–328.
[49] Y. Xu, S. Matthews, MAP-XSII: an improved program for the automatic assignment of methyl resonances in large proteins, J. Biomol. NMR 55 (2013) 179–187.
[50] F.A. Chao, L. Shi, L.R. Masterson, G. Veglia, FLAMEnGO: a fuzzy logic approach for methyl group assignment using NOESY and paramagnetic relaxation enhancement data, J. Magn. Reson. 214 (2012) 103–110.
[51] I. Pritisanac, M.T. Degiacomi, T.R. Alderson, M.G. Carneiro, E. Ab, G. Siegal, A.J. Baldwin, Automatic assignment of methyl-NMR spectra of supramolecular machines using graph theory, J. Am. Chem. Soc. 139 (2017) 9523–9533.
[52] W. Bermel, I. Bertini, L. Duma, I.C. Felli, L. Emsley, R. Pierattelli, P.R. Vasos, Complete assignment of heteronuclear protein resonances by protonless NMR spectroscopy, Angew. Chem., Int. Ed. Engl. 44 (2005) 3089–3092.
[53] W. Bermel, I. Bertini, I.C. Felli, M. Piccioli, R. Pierattelli, ^{13}C-detected protonless NMR spectroscopy of proteins in solution, Prog. Nucl. Magn. Res. Spec. 48 (2006) 25–45.
[54] K. Takeuchi, Z.Y. Sun, G. Wagner, Alternate 13C–12C labeling for complete mainchain resonance assignments using Ca direct-detection with applicability toward fast relaxing protein systems, J. Am. Chem. Soc. 130 (2008) 17210–17211.
[55] K. Takeuchi, D.P. Frueh, S.G. Hyberts, Z.Y. Sun, G. Wagner, High-resolution 3D CANCA NMR experiments for complete mainchain assignments using C(alpha) direct detection, J. Am. Chem. Soc. 132 (2010) 2945–2951.
[56] K. Takeuchi, H. Arthanari, M. Imai, G. Wagner, I. Shimada, Nitrogen-detected TROSY yields comparable sensitivity to proton-detected TROSY for non-deuterated, large proteins under physiological salt conditions, J. Biomol. NMR 64 (2016) 143–151.
[57] K. Takeuchi, H. Arthanari, I. Shimada, G. Wagner, Nitrogen detected TROSY at high field yields high resolution and sensitivity for protein NMR, J. Biomol. NMR 64 (2016) 143–151.
[58] I. Bertini, I.C. Felli, R. Kümmerle, D. Moskau, R. Pierattelli, 13C–13C NOESY: an attractive alternative for studying large macromolecules, J. Am. Chem. Soc. 126 (2004) 464–465.
[59] M. Kainosho, P. Guntert, SAIL–stereo-array isotope labeling, Q. Rev. Biophys. 42 (2009) 247–300.
[60] A. Boeszoermenyi, S. Chhabra, A. Dubey, D.L. Radeva, N.T. Burdzhiev, C.D. Chanev, O.I. Petrov, V.M. Gelev, M. Zhang, C. Anklin, H. Kovacs, G. Wagner, I. Kuprov, K. Takeuchi, H. Arthanari, Aromatic 19F–13C TROSY: a background-free approach to probe biomolecular structure, function, and dynamics, Nat. Methods 16 (2019) 333–340.
[61] A.J. Baldwin, L.E. Kay, NMR spectroscopy brings invisible protein states into focus, Nat. Chem. Biol. 5 (2009) 808–814.