Computational and Structural Biotechnology Journal 15 (2017) 328–339

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

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

Mini Review

Modern Technologies of Solution Nuclear Magnetic Resonance Spectroscopy for Three-dimensional Structure Determination of Proteins Open Avenues for Life Scientists

Toshihiko Sugiki *, Naohiro Kobayashi, Toshimichi Fujiwara

Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

ARTICLE INFO

Article history:
Received 13 January 2017
Received in revised form 31 March 2017
Accepted 3 April 2017
Available online 13 April 2017

Keywords:
Nuclear magnetic resonance (NMR) spectroscopy
Solution NMR
Automation
Protein
Structure determination
Validation

ABSTRACT

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique for structural studies of chemical compounds and biomolecules such as DNA and proteins. Since the NMR signal sensitively reflects the chemical environment and the dynamics of a nuclear spin, NMR experiments provide a wealth of structural and dynamic information about the molecule of interest at atomic resolution. In general, structural biology studies using NMR spectroscopy still require a reasonable understanding of the theory behind the technique and experience on how to recorded NMR data. Owing to the remarkable progress in the past decade, we can easily access suitable and popular analytical resources for NMR structure determination of proteins with high accuracy. Here, we describe the practical aspects, workflow and key points of modern NMR techniques used for solution structure determination of proteins. This review should aid NMR specialists aiming to develop new methods that accelerate the structure determination process, and open avenues for non-specialist and life scientists interested in using NMR spectroscopy to solve protein structures.

Contents

1. Introduction

Tertiary structure determination of biomolecules such as nucleic acids and proteins at atomic resolution provides essential insight into the function of bioactive molecules. X-ray crystallography [1] and nuclear magnetic resonance (NMR) spectroscopy [2] have been the primary methods over the past few decades to obtain high-resolution structures. More recently, the rapid technological growth of cryo-electron microscopy [3] has seen this technique emerge as a third major approach to solve biomacromolecular structures at atomic resolution. Solution NMR offers a number of distinct features for structural biology studies: 1) Dynamics of protein folding, structural fluctuations, internal mobility

* Corresponding author at: Laboratory of Molecular Biophysics, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan.
E-mail address: sugiki@protein.osaka-u.ac.jp (T. Sugiki).

http://dx.doi.org/10.1016/j.csbj.2017.04.001

and chemical exchange of target molecules can be investigated over a wide range of timescales (i.e., between picoseconds and sub-millisecond timescales) [4,5]. The most successful application of NMR to study biomolecular dynamics has been carried out on intrinsic disordered proteins (IDPs) such as zinc finger proteins that usually do not yield crystals, which provide sufficient quality X-ray diffraction data for high-resolution structure determination [6–11]. Moreover, these proteins are too small or too flexible to obtain strong contrast images by modern cryo-EM analysis. 2) Studies of protein-protein or protein–ligand interactions can be performed under physiological conditions. The affinity and the location of the interaction sites between the target protein and its binding partner molecules can be determined accurately and sensitively even if the interaction is very weak (K_d = ~mM) by performing simple NMR experiments, e.g., chemical shift perturbation analysis by measuring two-dimensional (2D) heteronuclear single-quantum coherence (HSQC) spectra of target proteins with and without the binding partner, while such an interaction study by co-crystallization or co-immunoprecipitation can be challenging [12–14]. Therefore, techniques based on using solution NMR can be readily used to screen chemical fragment libraries for preliminary hits, which have the potential to act as seed compounds in drug development. Such approaches usually identify weak binders for the target protein, but represent a powerful starting point for structure-based drug discovery and development studies.

Among the various applications of solution NMR to study protein function, as aforementioned, tertiary structure determination of proteins by NMR remains a steadfast and powerful application of this spectroscopy. Recent development of software and a number of web-based resources, which reduce the burden of complex NMR data analysis, have contributed to the systematic integration of sophisticated semi-automated NMR platforms for structure determination of biomolecules [15–19]. Such advances, along with improvements in NMR hardware, have lowered the knowledge barrier to facilitate entry into the field of structural biology by NMR for non-specialists. However, only a handful of articles that describe the workflow and practical aspects of protein structure determination by solution NMR spectroscopy, which covers the latest method developments, have been published despite the practicality of this technique.

2. Workflow of Protein Structure Determination by Solution NMR

The standard approach to protein structure determination by solution NMR comprises several steps: (i) preparation of a protein sample labeled with stable NMR spin-1/2 isotopes, ^{13}C and ^{15}N; (ii) acquisition of NMR spectra; (iii) data processing and assignments of signals; (iv) generation of a chemical shift table that gives the signal assignments derived from the analysis of NMR spectra such that the analyst or an automated program can subsequently assign NOE signals to generate

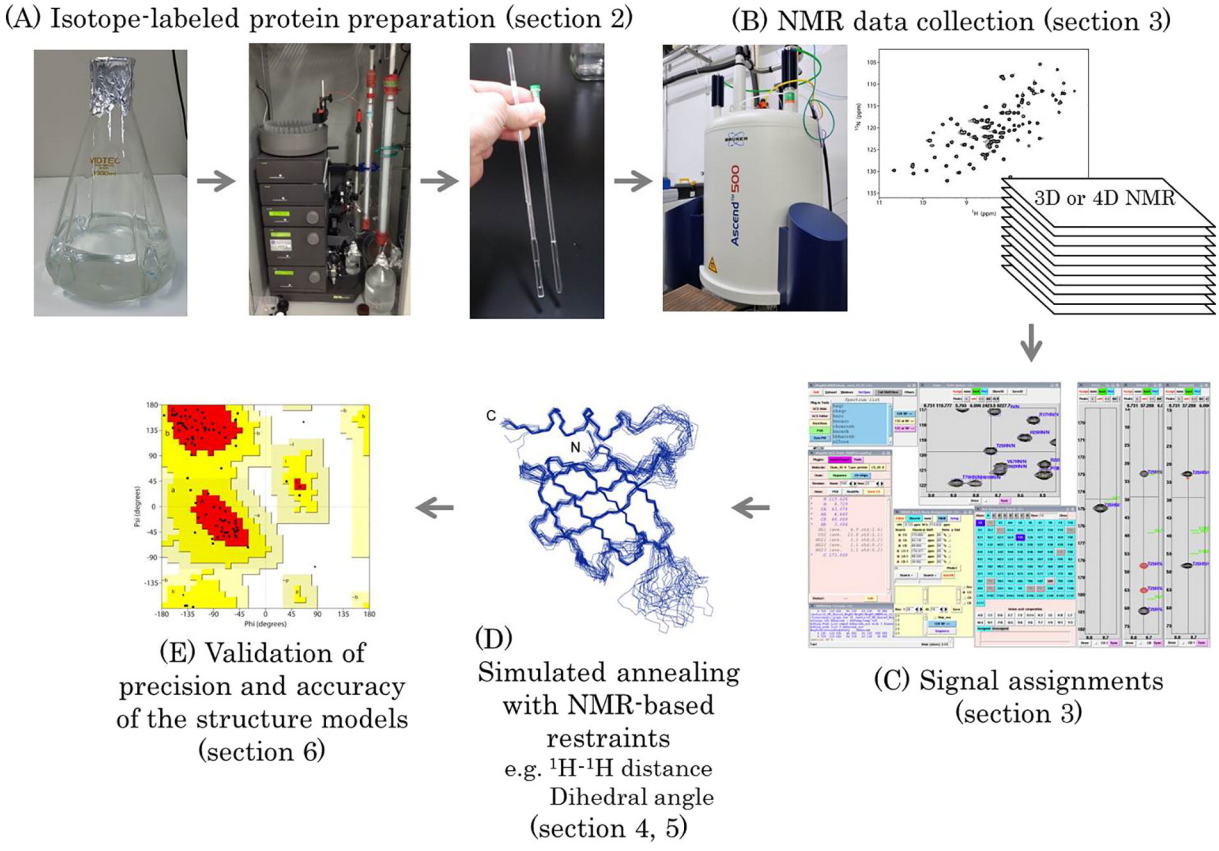

Fig. 1. Workflow of protein structure determination by solution NMR spectroscopy. (A) First, the target protein is uniformly labeled with NMR active isotopes (^{13}C and ^{15}N) and purified by chromatography processes. In many cases, it is essential to optimize the solution composition of the NMR sample (e.g., buffer, pH, type and concentration of salt, and other additives to prevent aggregation and/or to improve thermal stability of the target protein) and NMR parameters (e.g., sample temperature during NMR data collection, strength of static magnetic field, pulse sequences, inversion recovery delay) prior to collecting all the multidimensional heteronuclear NMR experiments. (B) Generally, various sample conditions and NMR parameters are verified by measuring 2D ^{1}H–^{15}N HSQC NMR spectra and/or 2D projection spectra of 3D HNCACB (e.g., 2D HN(CA)CB) until optimal conditions are found. If sample/NMR data collection optimization is performed by only ^{15}N-based NMR experiments, only a ^{15}N-labeled target protein (without ^{13}C-labeling) is required. (C) The NMR signals of all ^{1}H, ^{13}C and ^{15}N nuclei are assigned by analyzing multidimensional heteronuclear NMR spectra using NMR software that facilitates assignment of the signals. The right lower panel of this figure is a screenshot of a NMR signal assignment using the Kujira and MagRO software, developed by Prof. Naohiro Kobayashi (Osaka University, Japan), which is available at PDBj-BMRB http://bmrbdep.pdbj.org/en/nmr_tool_box/magro_nmrview.html. (D) After completing the assignment process, the solution structures of the target protein are determined by performing simulated annealing using NMR-based restraints such as ^{1}H–^{1}H distances and dihedral angles derived from NOEs and chemical shifts, respectively. (E) Finally, the precision and the accuracy of the determined structure are verified statistically and experimentally (e.g., by analyzing the Ramachandran plot and measuring RDCs of the target protein, respectively).