Bioinformatics, 32(4), 2016, 611–613
doi: 10.1093/bioinformatics/btv595
Advance Access Publication Date: 26 October 2015
Applications Note

Structural bioinformatics

NMRe: a web server for NMR protein structure refinement with high-quality structure validation scores

Hyojung Ryu[1,2], GyuTae Lim[1,2], Bong Hyun Sung[3] and Jinhyuk Lee[1,2,*]

[1]Korean Bioinformation Center (KOBIC), Korea Research Institute of Bioscience and Biotechnology, Yuseong-gu, Daejeon 305-806, the Republic of Korea, [2]Department of Nanobiotechnology and Bioinformatics, University of Science and Technology, Yuseong-Gu, Daejeon 305-350, the Republic of Korea and [3]Bioenergy and Biochemical Research Center, Korea Research Institute of Bioscience and Biotechnology, Yuseong-gu, Daejeon 305-806, the Republic of Korea

*To whom correspondence should be addressed.
Associate Editor: Anna Tramontano

Received on June 9, 2015; revised on September 23, 2015; accepted on October 10, 2015

Abstract

Summary: Protein structure refinement is a necessary step for the study of protein function. In particular, some nuclear magnetic resonance (NMR) structures are of lower quality than X-ray crystallographic structures. Here, we present NMRe, a web-based server for NMR structure refinement. The previously developed knowledge-based energy function STAP (Statistical Torsion Angle Potential) was used for NMRe refinement. With STAP, NMRe provides two refinement protocols using two types of distance restraints. If a user provides NOE (Nuclear Overhauser Effect) data, the refinement is performed with the NOE distance restraints as a conventional NMR structure refinement. Additionally, NMRe generates NOE-like distance restraints based on the inter-hydrogen distances derived from the input structure. The efficiency of NMRe refinement was validated on 20 NMR structures. Most of the quality assessment scores of the refined NMR structures were better than those of the original structures. The refinement results are provided as a three-dimensional structure view, a secondary structure scheme, and numerical and graphical structure validation scores.
Availability and implementation: NMRe is available at http://psb.kobic.re.kr/nmre/
Contact: jinhyuk@kribb.re.kr
Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

The determination of three-dimensional (3D) protein structures is an important challenge in structural biology because structure is closely related to function. NMR (Nuclear Magnetic Resonance) spectroscopy is one method of protein structure determination, and it also provides information about the dynamics of proteins. However, protein structures determined by NMR spectroscopy are of lower quality than X-ray structures. Thus, refinement is a necessary step in NMR structure study. Three NMR structure refinement databases have been constructed: DRESS (a Database of REfined Solution NMR Structures; Nabuurs *et al.*, 2004), RECOORD (REcalculated COORdinates Database; Nederveen *et al.*, 2005) and STAP (Statistical Torsion Angle Potential; Yang *et al.*, 2012). NMR structure refinement using the NOE (Nuclear Overhauser Effect) distance data obtained by NMR is performed in many refinement studies (Bertini *et al.*, 2011; Chen *et al.*, 2004; Mao *et al.*, 2014). However, NOE distance data exist in various formats, such as CYANA, CNS and XPLOR. Therefore, NOE data manipulation is difficult for general researchers. Another problem in improving NMR structures is the ambiguity and sparseness of NOE data (Nilges, 1997). The NMRe web server provides simple and fast NMR structure refinement with NOE-like distance data using the

inter-hydrogen distance restraints obtained from a given original structure. Conventional refinement is also performed using NOE distance data in the XPLOR/CYANA format. For acquiring better structure validation scores during the refinement process, a knowledge-based energy function (STAP) is added into the target energy potential. We expect the NMRe web server will contribute to NMR structure studies by providing accurate refinements.

2 Materials and methods

2.1 Two developed energy potentials

The STAP is a grid-type, knowledge-based energy potential that focuses on the torsion angle combinations (φ-ψ, φ-$\chi1$, ψ-$\chi1$ and $\chi1$-$\chi2$) of high-resolution X-ray structures under 2.0 Å (Kim *et al.*, 2013). The efficiency of STAP was demonstrated in earlier studies via NMR structure refinement (Yang *et al.*, 2012). Another energy potential is the structure-derived distance potential, which was generated by inter-hydrogen distances below 7 Å obtained from a given NMR structure, and this potential has greatly enhanced NMR structure refinement without experimental information (Ryu *et al.*, 2014). The potential functions consisted of two parameters: the equilibrium distance of two interacting hydrogen atoms and widths from 0 to 10 Å at 1 Å intervals. In a previous study, we found that the optimal width parameter is 4 Å. The optimal width is the default width for NMRe refinement calculations. This potential function prevents serious deviation from the current state of the original structure during the refinement.

2.2 Running a refinement

Whole process of NMRe refinement is shown in Figure 1A. First, NMRe converts the input Protein Data Bank (PDB) file into a CHARMM-readable format and all ensemble structures are separated for simulation. Then, the 'NMRe_dist' is performed with distance restraints obtained from the given input PDB structure. When a user optionally uploads the NOE restraint file, the 'NMRe_noe' is performed with 'NMRe_dist'. In the 'NMRe_noe', the uploaded NOE file is used to generate the restraining potential of CHARMM (Brooks *et al.*, 2009). The simulated annealing (SA) simulation is conducted with the CHARMM default energy and the EEF1.1 solvation energy (Lazaridis and Karplus, 1999) using CHARMM (see the Supplementary Information for more information). After refinement, NMRe evaluates structural quality to compare the refined structure and the original structure using structure validation scores.

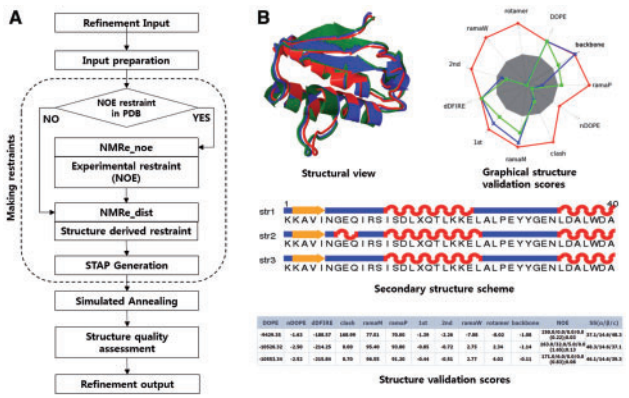

Fig 1. (**A**) Flow chart of the NMRe refinement process. (**B**) Output of a refinement. The output page contains a 3D structure viewer, a secondary structure scheme and numerical/graphical structure validation scores

The detailed structure validation scores are described in Supplementary Information.

3 Results

NMRe refinement was tested on 20 structures obtained from PDB that have both NMR and X-ray structures (Supplementary Table S1). The X-ray structures were used as a reference for measuring backbone similarity. The structure quality was measured using backbone similarity, NOE violation and structure validation scores (Supplementary Tables S2 and S3 and Supplementary Fig. S3). In the 'NMRe_dist', the TM-score was increased from 0.79 to 0.80, and three energy scores (DOPE, nDOPE and dDFIRE) were stabilized. The clash score was also improved from 53.81 to 0.40. The Ramachandran score and several parameters of the WHAT_CHECK *Z*-score were significantly improved in the 'NMRe_noe'. In particular, the PROCHECK Ramachandran score was improved from 80.98 to 95.36. In Supplementary Table S3, comprehensive NOE restraint violations were reported. The number of violated NOE, Root Mean Square Deviation (RMSD) and Maximum NOE violation were improved in 'NMRe_noe'. However, 'NMRe_dist' has a slightly increased NOE restraints violation in comparison with original. Because the experimental measurement of NOE distance may have a distance error (Nilges, 1997) of ~0.5–1.0 Å, these results are not regarded as a serious defect. Additionally, we demonstrated that dihedral angle restraints do not contribute much in this protocol (Supplementary Table S5). One of example target (2L4S) and comparison of the performance between the NMRe and other known refinement protocols are described in Supplementary Information.

4 Web server

NMRe uses PDB formats as an input 3D structure, and a NOE restraint is optional. The XPLOR or CYANA format is allowed. NMRe provides outputs in four ways (Fig. 1B and Supplementary Fig. S4) and is presented in the 'Analysis' page of the status table. The 3D structure viewer shows structural changes caused by the refinement simulation. The secondary structure scheme presents the creation of the secondary structure, and when a user clicks on the secondary structure shape, it is highlighted in the 3D structure viewer. The 'Structural scores' table shows the structure validation scores used to quantify the quality of various protein structure properties. The structure validation scores of the original/refined structures are compared with the normalized structure validation scores of 18 347 X-ray structures in a radar chart (Fig. 1B). The shaded region of the chart indicates greater quality than an average X-ray structure (upper 50% in structure quality). The refined structures can be downloaded through download links.

Funding

This study was supported by grants from the Korean Research Institute of Bioscience and the Biotechnology Research Initiative Program. The National Research Foundation of Korea funded this work under grant numbers 2012R1A1A2002676, 2012-0006002 and 20120006011. The Pioneer Research Center Program is funded by the Ministry of Science, ICT & Future Planning (2013M3C1A3064780).

Conflict of Interest: none declared.

Supplementary Data

Supplementary data are available at *Bioinformatics* online.

References

Bertini,I. *et al*. (2011) A grid-enabled web portal for NMR structure refinement with AMBER. *Bioinformatics*, **27**, 2384–2390.

Brooks,B.R. *et al*. (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.*, **30**, 1545–1614.

Chen,J. *et al*. (2004) Refinement of NMR structures using implicit solvent and advanced sampling techniques. *J. Am. Chem. Soc.*, **126**, 16038–16047.

Kim,T.R. *et al*. (2013) Statistical torsion angle potential energy functions for protein structure modeling: a bicubic interpolation approach. *Proteins*, **81**, 1156–1165.

Lazaridis,T. and Karplus,M. (1999) Effective energy function for proteins in solution. *Proteins*, **35**, 133–152.

Mao,B. *et al*. (2014) Protein NMR structures refined with Rosetta have higher accuracy relative to corresponding X-ray crystal structures. *J. Am. Chem. Soc.*, **136**, 1893–1906.

Nabuurs,S.B. *et al*. (2004) DRESS: a database of REfined solution NMR structures. *Proteins*, **55**, 483–486.

Nederveen,A.J. *et al*. (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins*, **59**, 662–672.

Nilges,M. (1997) Ambiguous distance data in the calculation of NMR structures. *Fold. Des.*, **2**, S53–S57.

Ryu,H. *et al*. (2014) Protein NMR structures refined without NOE data. *PloS One*, **9**, e108888.

Yang,J.S. *et al*. (2012) STAP refinement of the NMR database: a database of 2405 refined solution NMR structures. *Nucleic Acids Res.*, **40**, D525–D530.