Biochemical and Biophysical Research Communications 457 (2015) 467–472

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

# General order parameter based correlation analysis of protein backbone motions between experimental NMR relaxation measurements and molecular dynamics simulations

Qing Liu [a], Chaowei Shi [a], Lu Yu [a, b], Longhua Zhang [a], Ying Xiong [a, *], Changlin Tian [a, b, *]

[a] Hefei National Laboratory for Physical Sciences at The Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, 230026, PR China
[b] High Magnetic Field Laboratory, Chinese Academy of Science, Hefei, Anhui, 230031, PR China

ARTICLE INFO

*Article history:*
Received 20 December 2014
Available online 17 January 2015

*Keywords:*
Internal motion
Molecular dynamic simulation
NMR relaxation
General order parameters
Model-free approach
Charge states

ABSTRACT

Internal backbone dynamic motions are essential for different protein functions and occur on a wide range of time scales, from femtoseconds to seconds. Molecular dynamic (MD) simulations and nuclear magnetic resonance (NMR) spin relaxation measurements are valuable tools to gain access to fast (nanosecond) internal motions. However, there exist few reports on correlation analysis between MD and NMR relaxation data. Here, backbone relaxation measurements of $^{15}$N-labeled SH3 (Src homology 3) domain proteins in aqueous buffer were used to generate general order parameters ($S^2$) using a model-free approach. Simultaneously, 80 ns MD simulations of SH3 domain proteins in a defined hydrated box at neutral pH were conducted and the general order parameters ($S^2$) were derived from the MD trajectory. Correlation analysis using the Gromos force field indicated that $S^2$ values from NMR relaxation measurements and MD simulations were significantly different. MD simulations were performed on models with different charge states for three histidine residues, and with different water models, which were SPC (simple point charge) water model and SPC/E (extended simple point charge) water model. $S^2$ parameters from MD simulations with charges for all three histidines and with the SPC/E water model correlated well with $S^2$ calculated from the experimental NMR relaxation measurements, in a site-specific manner.



## 1. Introduction

Internal motions play a key role in the function of biomolecules such as proteins and nucleic acids [1–3]. The loop regions of proteins, which have evident conformational dynamics, are often involved in mediating specific protein–ligand, protein–protein and protein–DNA interactions [4,5]. These internal motions occur on a wide range of time scales, from femtoseconds to seconds [6], and numerous experimental and theoretical (simulation) approaches have been applied to understand this phenomenon [7–9]. These include nuclear magnetic resonance (NMR) spin relaxation measurements and molecular dynamics (MD) simulations, which have proved particularly valuable for accessing the fast (sub-nanosecond) internal motions. These techniques have been applied to a vast number of protein systems to reveal crucial information on molecular dynamics that consequently helps to elucidate the molecular mechanisms of biological processes [10,11]. However, there exist few reports that correlate MD simulation and NMR relaxation data, possibly due to the absence of a recognized, shared reference value between theses methods. In the present study, the general order parameter ($S^2$) was derived from solution NMR backbone relaxation analysis and MD simulation data, and $S^2$-based comparative analysis of protein backbone internal motions was subsequently performed.

Src homology 3 (SH3) domains are small, soluble protein modules that have been well-characterized using both solution NMR and MD simulation. SH3 domains contain five β-sheets and one $3_{10}$ helix linked via three loops that are referred to as the RT, n-Src and distal loop regions, respectively (Fig. 1A) [12]. Solution NMR relaxation experiments were carried out on the $^{15}$N-labeled SH3

* Corresponding authors. Hefei National Laboratory for Physical Sciences at The Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, 230026, PR China. Fax: +86 551 63600441.
*E-mail addresses:* yxiong73@ustc.edu.cn (Y. Xiong), cltian@ustc.edu.cn (C. Tian).

http://dx.doi.org/10.1016/j.bbrc.2015.01.018
0006-291X/

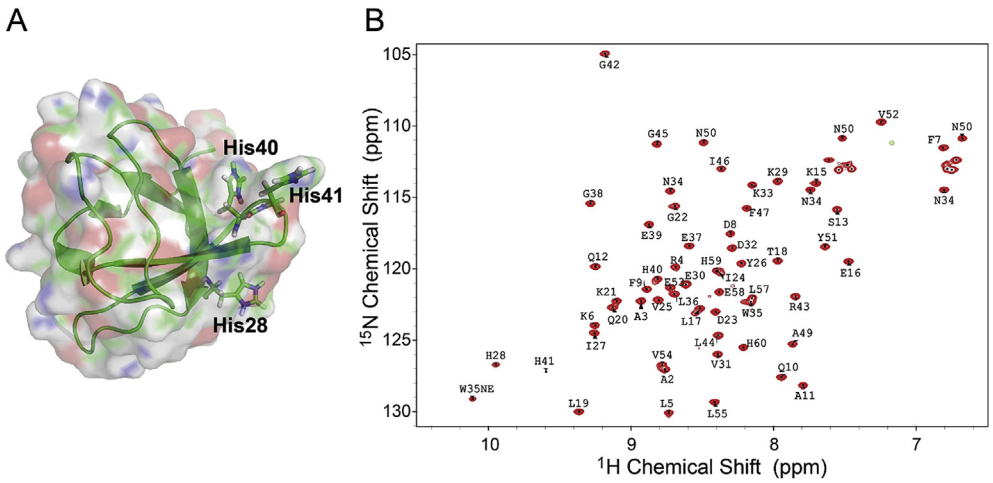


**Fig. 1.** (A) Backbone ribbon structure of SH3. The side chains of the three histidine residues (His28, His40 and His41) for which the charge state was varied are shown; (B) HSQC spectra of $^{15}$N-labeled SH3 at 600 MHz and 25 °C. Sequential assignments are indicated adjacent to each peak.

domain of human vinexin, and backbone amide N–H relaxation data were used to generate a the general order parameter ($S^2$) using a model-free approach [13,14]. Additionally, MD simulations of SH3 domains were performed on models with different charge states for the three histidines (His28, His40 and His41) and with different water structures. A general order parameter-based comparison of protein backbone motions using these NMR relaxation and MD simulation data was performed. MD simulation showed that analysis of the internal motion of the protein backbone was strongly influenced by both the charge states of surface histidines and the water structure of the model.

## 2. Materials and methods

### 2.1. $^{15}$N-labeled SH3 backbone relaxation

Uniformly $^{15}$N-labeled SH3 domain of human vinexin was expressed and purified as described previously [12]. A 1 mM solution of purified $^{15}$N-labeled SH3 in 50 mM PBS pH 6.5 (90% $H_2O$/10% $D_2O$) was used to measure backbone amide $^{15}$N $T_1$ and $T_2$ relaxation data and $^1$H–$^{15}$N steady-state NOE (Nuclear Overhauser Effect) relaxation data on a Bruker AVANCE 500 MHz or 600 MHz NMR spectrometer equipped with a cryoprobe. $T_1$ values were determined from a series of $^1$H–$^{15}$N correlation spectra with 11.7, 61.6, 142, 243, 364, 525, 757 and 1150 ms relaxation delays. $T_2$ values were obtained from spectra with 0, 17.6, 35.2, 52.8, 70.4, 105.6 and 140.8 ms delays. Steady-state $^1$H–$^{15}$N NOE values were determined from peak ratios from spectra with and without a 3 s proton presaturation.

### 2.2. Model-free order parameter ($S^2$) calculation

According to the relaxation theory of Bloch, Wangness and Redfield, the spin-lattice relaxation rate ($R_1 = 1/T_1$), the spin–spin relaxation rate ($R_2 = 1/T_2$) and Heteronuclear Overhauser Enhancement (NOE) are given by:

$$R_1 = d/4*[J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H + \omega_C)] + cJ(\omega_C)$$

$$R_2 = d/8*[4J(0) + J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H) + 6J(\omega_H + \omega_C)] + c/6*[4J(0) + 3J(\omega_C)]$$

$$NOE = 1 + d/4R_1*\gamma_C/\gamma_H*[6J(\omega_H + \omega_C) - J(\omega_H - \omega_C)]$$

where $d = (\mu_0^2/4\pi^2)(h^2\gamma_C^2\gamma_H^2/r_{CH}^6)$, $c = \omega_C^2\Delta\sigma_C^2/3$, and $J(\omega)$ represents the spectra density.

In the model-free approach of Lipari-Szabo, the internal and overall motions are assumed to be independent, and the spectra density is described by

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_c}{1+\tau_c^2\omega^2} + \frac{(1-S^2)\tau}{1+\tau^2\omega^2}\right]$$

When $\tau^{-1} = \tau_c^{-1} + \tau_e^{-1}$, $\tau_c$ is the rotation correlation time, $\tau_e$ is the effective correlation time, and $S^2$ is the order parameter [13,14]. Extended order parameter ($S^2$, $S_f^2$, $S_s^2$), the effective correlation time for fast internal motions ($\tau_e$) and $^{15}$N exchange broadening contribution (Rex) values were obtained using the relax program and the protocol described [15–17].

### 2.3. Molecular dynamic simulation of SH3 with different histidine charge states and water models

The starting SH3 structure was derived from NMR experiments (PDB ID: 2NWM) and Gromacs-4.5.3 simulation software was applied [18]. Three histidine residues were assigned different charge states (neutral or positive) and the Gromos96-53a6 force field was used to generate the model [19]. A 3.2 ns equilibrium and 80 ns MD simulations were conducted during computation. The periodic boundary condition was added and the box was set as a dodecahedron. During the 80 ns MD simulations, the temperature and pressure were set to 303 K and 1 bar, respectively. Protein and non-protein atoms were coupled to their own temperatures and pressures, and v-rescale and Parrinello–Rahman were chosen for temperature- and pressure-coupling methods, respectively [20,21]. The coupling time constants were set as 0.1 ps and 1.0 ps for temperature and pressure, respectively, compressibility was set to $4.5e^{-5}$ $bar^{-1}$, and 2 fs time steps were employed for a total of 40,000,000 calculation steps. The neighbor grid searching method was applied and the neighbor list was updated every ten steps. Twin range cut-offs were chosen for Lennard–Jones interactions, and the distance for the Lennard–Jones cut-off was equal to the cut-off distance for the short-range neighbor list (1.4 nm). Long-range dispersion corrections were applied for energy and

pressure. Fast Smooth Particle-Mesh Ewald (PME) electrostatics were used for coulomb interactions, with 0.1 nm for Fourier spacing and 4 for PME order [22]. The distance for the coulomb cut-off was also 1.4 nm. All bonds were constrained by the LINCS algorithm. For the whole system, the center of mass translation was removed every ten steps. Finally, the leap-frog algorithm was used for the 80 ns MD simulation [23].

A total of five sets of MD simulations were conducted for four different histidine charge states (neutral or positive) and/or two different water models (spc or spc/e) as follows: $His28^{+}His40^{0}His41^{0}$/spc, $His28^{0}His40^{+}His41^{0}$/spc, $His28^{0}His40^{0}His41^{0}$/spc, $His28^{+}His40^{+}His41^{+}$/spc, and $His28^{+}His40^{+}His41^{+}$/(spc/e).

### 2.4. Order parameter ($S^2$) calculation from molecular dynamic simulation calculations

MD simulation trajectories were subjected to removing periodic boundary conditions, molecular translations and rotations, and molecular rotational correlation functions for backbone N–H bonds were calculated from the first 10–80 ns of the simulations. Finally, general order parameters ($S^2$) were derived from the molecular rotational functions using a dedicated perl script.

### 2.5. Comparative analysis of protein internal backbone motions from NMR and MD experiments

General order parameters derived from solution NMR or MD simulation were tabulated in a site-specific manner, and $S^2$ from five different MD simulations were compared with NMR experimental measurements.

## 3. Results and discussion

### 3.1. Solution NMR backbone amide relaxation analysis of SH3

A high-quality heteronuclear single quantum correlation (HSQC) spectrum for SH3 in aqueous solution was obtained for further assignment and analysis (Fig. 1B). All 58 backbone amide ($^{1}H$, $^{15}N$) resonances were assigned for the 58 residue SH3 domain.

### 3.2. Model-free order parameter ($S^2$) calculation from NMR relaxation data

Backbone $^{15}N$ $R_1$ (the spin-lattice relaxation rate, $R_1 = 1/T_1$), $R_2$ (the spin–spin relaxation rate, $R_2 = 1/T_2$) and hetero-nuclear

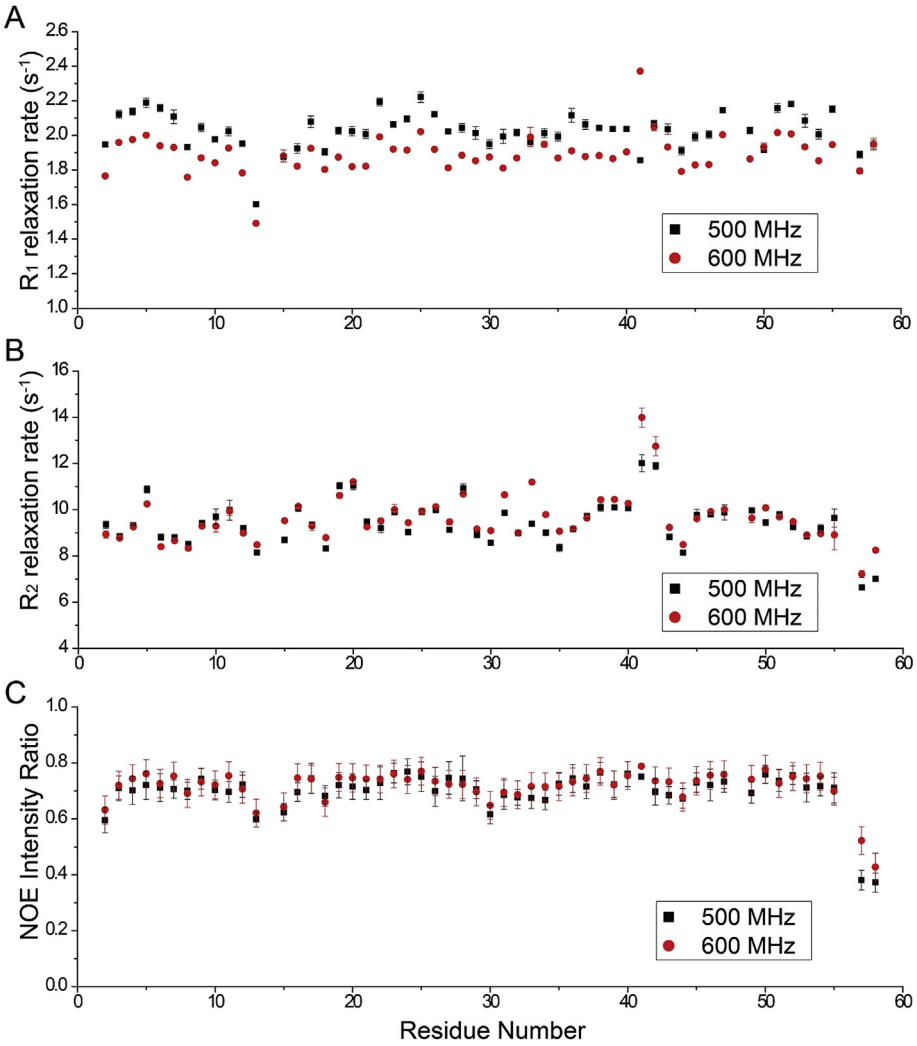


**Fig. 2.** Backbone amide relaxation data collected at 500 MHz (black squares) or 600 MHz (red circles). (A) Amide $^{15}N$ $R_1$ longitudinal relaxation rate; (B) Amide $^{15}N$ $R_2$ transverse relaxation rate; (C) $^{1}H$–$^{15}N$ steady NOE intensity ratio. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$^{1}H-^{15}N$ NOE relaxation data from SH3 in aqueous buffer were acquired at 500 and 600 MHz (Fig. 2). Examples of decay of cross-peak intensities as a function of inversion recovery delays in $T_1$ experiments and CPMG delays in $T_2$ experiments for two resonances (Ser13 and Lys33) are shown in the supporting information (Supplementary Fig. S1). $R_1$, $R_2$ and NOEs values for all 52 non-proline $^{15}N$ sites were calculated (Fig. 2) and data from 500 to 600 MHz spectra were correlated. $R_1$ values were reasonably constant throughout the sequence except for residues Ser13 and His41, and $R_2$ values were also uniformly distributed except for residues His41 and Gly42, and were larger than $R_1$ values. The proposed RT loop (Lys6-Asp23) and n-Src loop (His28-Trp35), visible in the solution NMR structure of SH3, were not significantly abundant, according to the amplitude of $R_1$, $R_2$ and steady-state NOE values. However, the distal loop (His41-Gly42) was apparent and the large $R_2$ values indicated high flexibility.

$R_1$, $R_2$ and NOE values were subsequently used to derive order parameters ($S^2$) which offer an intuitive way to characterize the amplitude of internal picosecond-to-nanosecond timescale motions of bond vectors resulting from NMR relaxation experiments [24]. $S^2$ values are always between 0 (the highest degree of disorder for a bond vector) and 1 (a completely restricted bond vector). Order parameters derived from NMR relaxation experiments using the model-free formalism were later used as references for evaluating the accuracy of the MD simulations.

### 3.3. Molecular dynamics simulations of SH3 domains with different histidine charge states and water models

Molecular dynamic simulation of SH3 proteins with different histidine charge states and/or water models were conducted using the Gromacs-4.5.3 simulation software with the Gromos96-53a6

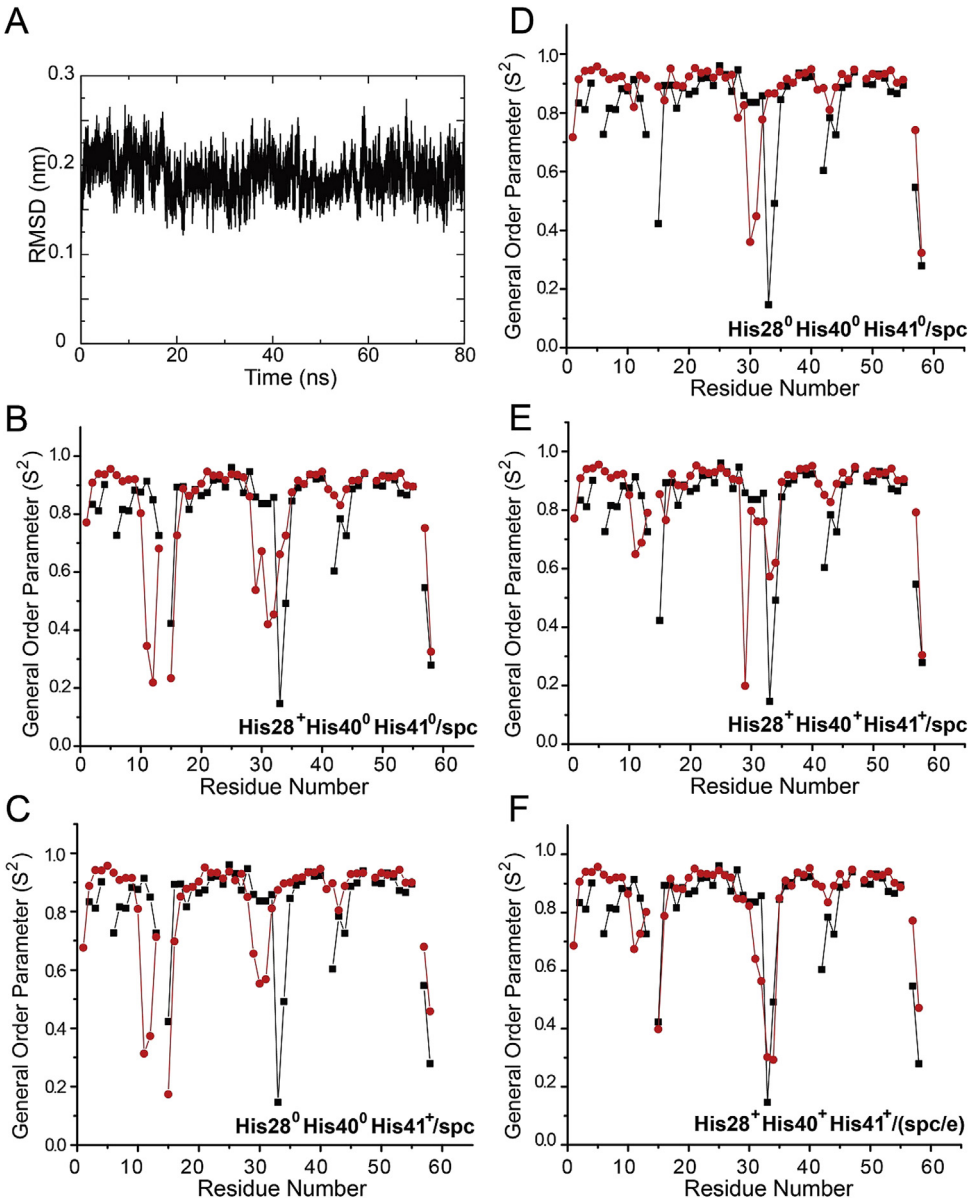


**Fig. 3.** (A) Molecular dynamic simulations (80 ns) of SH3; (B–F) Comparison of order parameters ($S^2$) between model-free calculations from NMR relaxation experiments and molecular dynamic simulations using different water models (B, C, D, E: spc; F: spc/e) and charge states of the three histidine residues (B: His28 charged; C: His41 charged; D: all three histidines not charged; E, F: all three histidines charged).

force field. The backbone $C_\alpha$ RMSD fluctuated slightly (0.2 nm) during the 80 ns MD simulations (Fig. 3A), indicating a proper equilibrium state for this system. Therefore, the simulation trajectories could be used directly for general order parameter derivation. Since molecular dynamics simulations could not mimic the ionization equilibrium, different histidine charge states were explored to identify the likely state present in the experimental data. $S^2$ values were subsequently derived from the simulation results.

### 3.4. Comparison of NMR relaxation data and molecular dynamic simulations using $S^2$ values

Four different histidine charge states were applied in the simulation: H28 charged ($His28^+His40^0His41^0$/spc), His40 charged ($His28^0His40^+His41^0$/spc), all uncharged ($His28^0His40^0His41^0$/spc) and all charged ($His28^+His40^+His41^+$/spc). General order parameters from the four histidine charge states and the SPC water model were compared with the NMR data (Fig. 3B–E). The $S^2$ values derived from the two methods were in close agreement for most residues except for three intervals where this value fell below 0.8. Upon inspection these intervals were found to coincide with the three loop regions (Lys6-Asp23, His28-Trp35, His41-Gly42) visible in the solution NMR structure. The NMR and MD $S^2$ differences were mostly located in these loop regions for all four histidine charge states, indicating that the correct fitting of these loop regions was crucial for determining the likely charge state. Correlation analysis between different series of $S^2$ values and experimental values yielded $R^2$ values that were used to evaluate the correct histidine charge state (Supplementary Fig. S2). An $R^2$ value of 1 represents a perfect correlation between the data, and the $R^2$ values were 0.2794, 0.2022, 0.1233, and 0.1177 for the all charged, His28 charged, His41 charged and all uncharged states, respectively. The all-charged state gave the highest correlation, and the ($His28^+His40^+His41^+$/spc) SH3 domain therefore mimics the experimental conditions most accurately.

To evaluate the influence of different water models on the MD simulation $S^2$ values, simulations for the SH3 all-charged state were conducted with the SPC and SPC/E water models. The SPC/E water model is a more advanced version of the SPC water model that describes the water molecule structure more accurately but at a higher expense in computational power [25,26]. The simulation with the SPC/E water model gave a superior correlation with the experimental data, as indicated by the $S^2$ values (Fig. 3F). The $R^2$ value for the all-charged state with the SPC/E water model was 0.5431, the highest correlation between NMR experimental data and MD simulation data.

The $R^2$ correlation values were generally low for all simulations, which may be due to the force field and parameters chosen during set-up. Gromos is not an all-atom force field and may treat some atoms details vaguely. Most of the parameters were empirical or approximate, and may not accurately describe the true molecular dynamics.

In summary, general order parameters ($S^2$) were calculated from $^{15}N$-labeled SH3 NMR backbone relaxation data using a model-free approach and compared with $S^2$ values derived from the trajectory of SH3 MD simulations using different histidine charge states and/or different water models in a defined hydrated box at neutral pH. $S^2$-based correlation analysis was subsequently used to evaluate the protein backbone internal motions. Interestingly, both histidine charge states and water models heavily influenced the simulation results. The all-charged state and SPC/E water model gave the best correlation between experimental NMR relaxation and simulation data.

## Acknowledgements

This work was supported by the Chinese Key Research Plan—-Protein Science (2013CB910202) and the Chinese Natural Science Foundation (U1332138 and U1432136).

## Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.bbrc.2015.01.018.

## Transparency document

Transparency document related to this article can be found online at http://dx.doi.org/10.1016/j.bbrc.2015.01.018.

## References

[1] Y.P. Chu, C.H. Chang, J.H. Shiu, Y.T. Chang, C.Y. Chen, W.J. Chuang, Solution structure and backbone dynamics of the DNA-binding domain of FOXP1: insight into its domain swapping and DNA binding, Protein. Sci. 20 (2011) 908–924.
[2] V.R. Moorman, K.G. Valentine, S. Bedard, V. Kasinath, J. Dogan, F.M. Love, A.J. Wand, Dynamic and thermodynamic response of the Ras protein Cdc42Hs upon association with the effector domain of PAK3, J. Mol. Biol. 426 (2014) 3520–3538.
[3] K. Petzold, E. Duchardt, S. Flodell, G. Larsson, K. Kidd-Ljunggren, S. Wijmenga, J. Schleucher, Conserved nucleotides in an RNA essential for hepatitis B virus replication show distinct mobility patterns, Nucleic. Acids. Res. 35 (2007) 6854–6861.
[4] C.E. Chua, B.L. Tang, The role of the small GTPase Rab31 in cancer, J. Cell. Mol. Med. 19 (2015) 1–10.
[5] L. Ma, Y.Y. Sham, K.J. Walters, H.C. Towle, A critical role for the loop region of the basic helix-loop-helix/leucine zipper protein Mlx in DNA binding and glucose-regulated transcription, Nucleic. Acids. Res. 35 (2007) 35–44.
[6] M. Guenneugues, B. Gilquin, N. Wolff, A. Menez, S. Zinn-Justin, Internal motion time scales a small, highly stable disulfide-rich protein: a 15N, 13C NMR and molecular dynamics study, J. Biomol. NMR. 14 (1999) 47–66.
[7] D. Toledo Warshaviak, V.V. Khramtsov, D. Cascio, C. Altenbach, W.L. Hubbell, Structure and dynamics of an imidazoline nitroxide side chain with strongly hindered internal motion in proteins, J. Magn. Reson. 232 (2013) 53–61.
[8] O. Julien, P. Mercier, L. Spyracopoulos, J.E. Corrie, B.D. Sykes, NMR studies of the dynamics of a bifunctional rhodamine probe attached to troponin C, J. Am. Chem. Soc. 130 (2008) 2602–2609.
[9] B. Reif, Y. Xue, V. Agarwal, M.S. Pavlova, M. Hologne, A. Diehl, Y.E. Ryabov, N.R. Skrynnikov, Protein side-chain dynamics observed by solution- and solid-state NMR: comparative analysis of methyl 2H relaxation data, J. Am. Chem. Soc. 128 (2006) 12354–12355.
[10] B. Stauch, J. Orts, T. Carlomagno, The description of protein internal motions aids selection of ligand binding poses by the INPHARMA method, J. Biomol. NMR. 54 (2012) 245–256.
[11] A.J. Wand, V.R. Moorman, K.W. Harpole, A surprising role for conformational entropy in protein function, Top. Curr. Chem. 337 (2013) 69–94.
[12] J. Zhang, X. Li, B. Yao, W. Shen, H. Sun, C. Xu, J. Wu, Y. Shi, Solution structure of the first SH3 domain of human vinexin and its interaction with vinculin peptides, Biochem. Biophys. Res. Commun. 357 (2007) 931–937.
[13] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, J. Am. Chem. Soc. 104 (1982) 4546–4559.
[14] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results, J. Am. Chem. Soc. 104 (1982) 4559–4570.
[15] M. Bieri, E.J. d'Auvergne, P.R. Gooley, RelaxGUI: a new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and mus motion of proteins, J. Biomol. NMR. 50 (2011) 147–155.
[16] E.J. d'Auvergne, P.R. Gooley, Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces, J. Biomol. NMR. 40 (2008) 107–119.
[17] E.J. d'Auvergne, P.R. Gooley, Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor, J. Biomol. NMR. 40 (2008) 121–133.
[18] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[19] U. Stocker, W.F. van Gunsteren, Molecular dynamics simulation of hen egg white lysozyme: a test of the GROMOS96 force field against nuclear magnetic resonance data, Proteins 40 (2000) 145–153.
[20] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 014101.

[21] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (1981).
[22] M.J. Abraham, J.E. Gready, Optimization of parameters for molecular dynamics simulation using smooth particle-mesh Ewald in GROMACS 4.5, J. Comput. Chem. 32 (2011) 2031–2040.
[23] R.W. Hockney, S.P. Goel, J.W. Eastwood, Quiet high-resolution computer models of a plasma, J. Comput. Phys. 14 (1974).
[24] T. Reddy, J.K. Rainey, Interpretation of biomolecular NMR spin relaxation parameters, Biochem. Cell. Biol. 88 (2010) 131–142.
[25] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: B. Pullman, D. Reidel (Eds.), Intermolecular Forces, Publishing Company Dordrecht, 1981, pp. 331–342.
[26] H.J.C. Berendsen, J.R. Grigera, T. Straatsma, The missing term in effective pair potentials,, J. Phys. Chem. 91 (1987) 6269–6271.