JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Hydrogen Exchange Study on the Hydroxyl Groups of Serine and Threonine Residues in Proteins and Structure Refinement Using NOE Restraints with Polar Side-Chain Groups

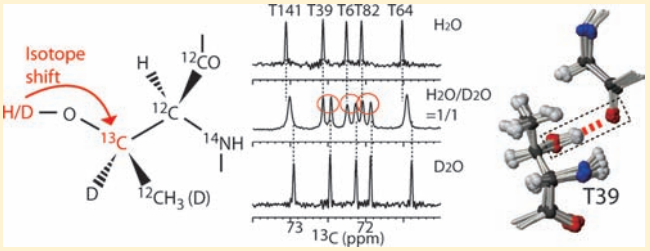
Mitsuhiro Takeda,[†] JunGoo Jee,[‡] Akira M. Ono,[‡] Tsutomu Terauchi,[‡] and Masatsune Kainosho*[,†,‡]

[†]Structural Biology Research Center, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan

[‡]Center for Priority Areas, Graduate School of Science and Technology, Tokyo Metropolitan University, 1-1 minami-ohsawa, Hachioji 192-0397, Japan

Supporting Information

ABSTRACT: We recently developed new NMR methods for monitoring the hydrogen exchange rates of tyrosine hydroxyl (Tyr−OH) and cysteine sulfhydryl (Cys−SH) groups in proteins. These methods facilitate the identification of slowly exchanging polar side-chain protons in proteins, which serve as sources of NOE restraints for protein structure refinement. Here, we have extended the methods for monitoring the hydrogen exchange rates of the OH groups of serine (Ser) and threonine (Thr) residues in an 18.2 kDa protein, EPPIb, and thus demonstrated the usefulness of NOE restraints with slowly exchanging OH protons for refining the protein structure. The slowly exchanging Ser/Thr−OH groups were readily identified by monitoring the $^{13}C_{\beta}$-NMR signals in an H_2O/D_2O (1:1) mixture, for the protein containing Ser/Thr residues with ^{13}C, ^{2}H-double labels at their β carbons. Under these circumstances, the OH groups exist in equilibrium between the protonated and deuterated isotopomers, and the $^{13}C_{\beta}$ peaks of the two species are resolved when their exchange rate is slower than the time scale of the isotope shift effect. In the case of EPPIb dissolved in 50 mM sodium phosphate buffer (pH 7.5) at 40 °C, one Ser and four Thr residues were found to have slowly exchanging hydroxyl groups ($k_{ex} < \sim 40\ s^{-1}$). With the information for the slowly exchanging Ser/Thr−OH groups in hand, we could collect additional NOE restraints for EPPIb, thereby making a unique and important contribution toward defining the spatial positions of the OH protons, and thus the hydrogen-bonding acceptor atoms.



INTRODUCTION

The structure of a protein is stabilized by a large number of hydrogen bonds between polar atoms in the polypeptide chain.[1−3] Therefore, it is quite important to investigate these specific interactions to advance our understanding of the nature of proteins.

Protein hydrogen exchange studies have played an important role in investigations of hydrogen bonds.[4] Especially, NMR studies have made large contributions, because they can provide atom-resolved information.[5] However, because a conventional NMR hydrogen exchange study involves the observation of exchanging protons, this method is not optimal to study polar groups with intrinsically rapid exchange rates, which is the case for side-chain OH and SH groups.[6−9] In addition, even when a slowly exchanging OH/SH proton exists in a protein, it is difficult to observe, because conventional 2D experiments, such as $^{1}H-^{15}N$ and $^{1}H-^{13}C$ HSQC, cannot be applied. Therefore, a large number of such observable OH/SH proton peaks probably have been overlooked or mistakenly interpreted.

The situation would undoubtedly be improved, if the respective hydrogen exchange rates of OH/SH groups could be monitored independently by another method. In an effort to realize this, we have recently developed NMR methods for monitoring the proton/deuterium hydrogen exchange rates of Tyr−OH and Cys−SH groups in proteins.[10,11] In these methods, the NMR resonances of the carbon atoms attached to the polar groups are observed in a 1:1 mixture of H_2O and D_2O, where the side-chain polar groups exist in equilibrium between the protonated and deuterated isotopomers. Because the chemical shifts of the carbon resonances are slightly different between the two isotopomers, due to an isotope shift effect,[12] the isotopomer-resolved carbon resonances are separately observed when the exchange rate is slower than the inverse of the isotope shift.[13] The key behind these methods is that the line-widths of the target carbon atoms, $^{13}C_{\zeta}$ of Tyr and $^{13}C_{\beta}$ of Cys, should be reduced by using Tyr and Cys with optimized isotope labeling patterns, (2*S*,3*R*)-[0,α,β,ζ-$^{13}C_4$; β_2,$\varepsilon_{1,2}$-$^{2}H_3$; ^{15}N]-Tyr (ζ-SAIL Tyr)[14] and L-[β-^{13}C;β,β-$^{2}H_2$] Cys. In the case of an 18.2 kDa protein, *E. coli* peptidyl prolyl cis−trans isomerase b (EPPIb), the exchange rates of two Tyr−OH groups and the two Cys−SH groups were readily found to be slower than the order of 10 s^{-1} by this method.[10,11]

Received: July 27, 2011
Published: September 28, 2011

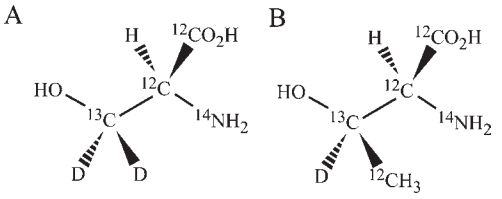
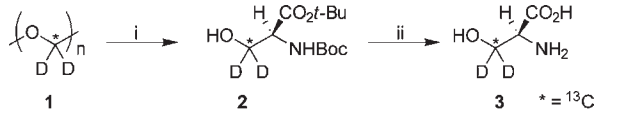


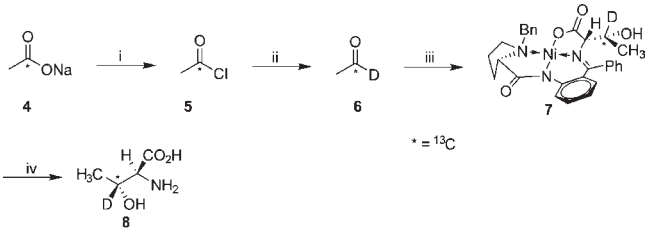
Figure 1. Chemical structures of L-[β-^{13}C;β,β-^{2}H$_2$] serine and L-[β-^{13}C;β-^{2}H] threonine.

Scheme 1. Preparation of L-[β-^{13}C;β,β-^{2}H$_2$]Ser[a]



[a] Reaction conditions: (i) see ref 11, (ii) H+, Δ, then ion exchange.

Scheme 2. Preparation of L-[β-^{13}C;β-^{2}H] Thr[a]



[a] Reaction conditions: (i) phthaloyl dichloride, (ii) Bu_3SnD, $Pd(PPh_3)_4$, (iii) Ni-(S)-BPB-Gly complex, MeONa, MeOH, (iv) H+, Δ, then ion exchange.

In this study, we have extended the method for monitoring the hydrogen exchange rates of the Ser−OH and Thr−OH groups. To reduce the $^{13}C_\beta$ line-widths of Ser and Thr, which are sensitive reporters for the OH/OD states of hydroxyl groups, we synthesized L-[β-^{13}C;β,β-^{2}H$_2$] Ser and L-[β-^{13}C;β-^{2}H] Thr (Figure 1) and incorporated them into EPPIb by an *E. coli* cell-free expression system.[15,16] In the case of EPPIb dissolved in 50 mM sodium phosphate buffer (pH 7.5) at 40 °C, we could identify one Ser and four Thr residues with slowly exchanging hydroxyl groups. Subsequently, NOE restraints involving the slowly exchanging OH and SH protons, identified in this (Ser, Thr) and previous studies (Tyr, Cys),[10,11] were implemented in restrained simulated annealing molecular dynamics for the NMR structure calculation of EPPIb, based on EPPIb composed exclusively of SAIL amino acids. The effect of the supplemental restraints on the quality of the NMR structure was evaluated.

MATERIAL AND METHODS

Chemical Syntheses of L-[β-^{13}C;β,β -^{2}H$_2$] Serine and L-[β-^{13}C;β -^{2}H] Threonine. L-[β-^{13}C;β,β-^{2}H$_2$]Ser (Scheme 1) was prepared by acid hydrolysis of the intermediate for the synthesis of L-[β-^{13}C;β,β-^{2}H$_2$]Cys.[11] The synthesis of L-[β-^{13}C;β-^{2}H]Thr (Scheme 2) was accomplished via aldol condensation of acetaldehyde, which was prepared from sodium [1-^{13}C]acetate, with a Ni-BPB-Gly complex[17] to obtain the diastereomerically pure threonine.

Preparation of *E. coli* Peptidyl-Prolyl Cis−Trans Isomerase b, EPPIb, Selectively Labeled with Isotope-Labeled Ser and Thr. The protein samples of EPPIb selectively labeled with each of the isotope-labeled Ser and Thr were produced by using an *E. coli* cell-free system, as described previously.[15,16] The cell-free reactions were performed in a vessel composed of 5 mL of an inner reaction solution and 20 mL of an outer feeding solution, with shaking for 8 h at 37 °C. The synthesized Ser and Thr were each present at 0.5 mM, and all of the other unlabeled amino acids were present at 1 mM, for both the inner and the outer solutions. The produced EPPIb protein was purified according to the previously reported method.[18] Each NMR sample contained 0.2−0.7 mM EPPIb in 50 mM sodium phosphate buffer, containing 100 mM NaCl and 0.1 mM NaN_3, pH 7.5 (meter reading).

NMR Measurements and Analysis. All NMR experiments were recorded on a DRX600 spectrometer (Bruker Biospin; 600.3 MHz for ^{1}H) equipped with a TCI cryogenic triple resonance probe at 40 °C, unless otherwise stated. The ^{13}C direct observations of EPPIb samples were performed with proton and/or deuterium decoupling by a WALTZ16 ^{1}H and/or ^{2}H decoupling scheme.[19] The carrier frequencies of carbon were 62 ppm for the Ser $^{13}C_\beta$ observation and 72 ppm for the Thr $^{13}C_\beta$ observation. The sweep widths were both 4500 Hz, and the repetition time was 5 s. In the experiment in the 100% H_2O solution, a 4.1 mm o.d. Shigemi tube containing the protein solution was inserted into a 5 mm o.d. outer tube containing pure D_2O, for the ^{2}H lock signal. All of the acquired 1D-^{13}C NMR data were processed without a window function, to facilitate the comparison of the line-widths of the protonated and deuterated β-carbon signals. All of the protein samples dissolved in H_2O/D_2O (1:1) and 100% D_2O solutions were incubated for at least 1 day at 40 °C to ensure that the hydrogen−deuterium equilibrium was completed before the NMR measurements.

Collection of NOEs with Slowly Exchanging Hydroxyl Protons. The assignments of the Ser/Thr hydroxyl protons and their NOE peaks were performed by using ^{13}C-edited NOESY-HSQC and ^{15}N-edited NOESY-HSQC data, which were acquired with 0.3−0.5 mM of the SAIL-EPPIb sample, exclusively composed of SAIL amino acids[20] (Figure S1). The SAIL-EPPIb protein contains (2*S*,3*S*)-[ul-^{13}C;^{15}N;-β-^{2}H] Ser and L-[ul-^{13}C;^{15}N;γ,γ-^{2}H$_2$] Thr. The ^{15}N-edited 3D-NOESY-HSQC spectra were measured with a Bruker DRX600 spectrometer equipped with a TXI cryogenic probe. The data points and the spectral width of the ^{15}N-edited NOESY-HSQC were 256 (t_1) × 48 (t_2) × 1024 (t_3) points and 8400 Hz (ω_1, ^{1}H) × 1800 Hz (ω_2, ^{15}N) × 8400 Hz (ω_3, ^{1}H), respectively. The number of scans/FID was 32, and the NOE mixing time was 200 ms. The repetition time was 1 s. The ^{13}C-edited 3D-NOESY-HSQC experiments were performed with a DRX800 spectrometer equipped with a normal TXI probe, and the experiments were acquired for two different regions, the aliphatic and aromatic regions. In the first set for the aliphatic region, the data points and the spectral width were 210 (t_1) × 58 (t_2) × 1024 (t_3) points and 12 000 Hz (ω_1, ^{1}H) × 6700 Hz (ω_2, ^{13}C) × 10 000 Hz (ω_3, ^{1}H), respectively. The number of scans/FID was 32. The repetition time was 1 s. The carrier frequency of carbon was set to 40 ppm. In the second set for the aromatic region, the data points and the spectral width were 200 (t_1) × 40 (t_2) × 1,024 (t_3) points and 12 000 Hz (ω_1, ^{1}H) × 5800 Hz (ω_2, ^{13}C) × 10 000 Hz (ω_3, ^{1}H), respectively. The number of scans/FID was 48. The repetition time was 1 s. The carrier frequency of carbon was set to 125.5 ppm. The mixing time for both of the ^{13}C-edited NOESY-HSQC spectra was 145 ms.

Structure Calculation with Distance Restraints Involving Hydroxyl and Sulfhydryl Protons. In total, 34 NOE peaks involving Ser−OH, Thr−OH, and Tyr−OH and 10 Cys−SH peaks were assigned, and their peak intensities were converted to upper distance restraints by using a calibration protocol implemented in CYANA.[21] In the conversion of the peak intensity to the upper distance restraint, the exchange-relaxation effect during the NOE mixing time was not considered, assuming that exchange processes on the order of 10^{-1} s^{-1} did not appreciably affect the intensity.[6]

The structural calculation of the EPPIb protein was performed in the presence and absence of NOE restraints involving the OH and SH

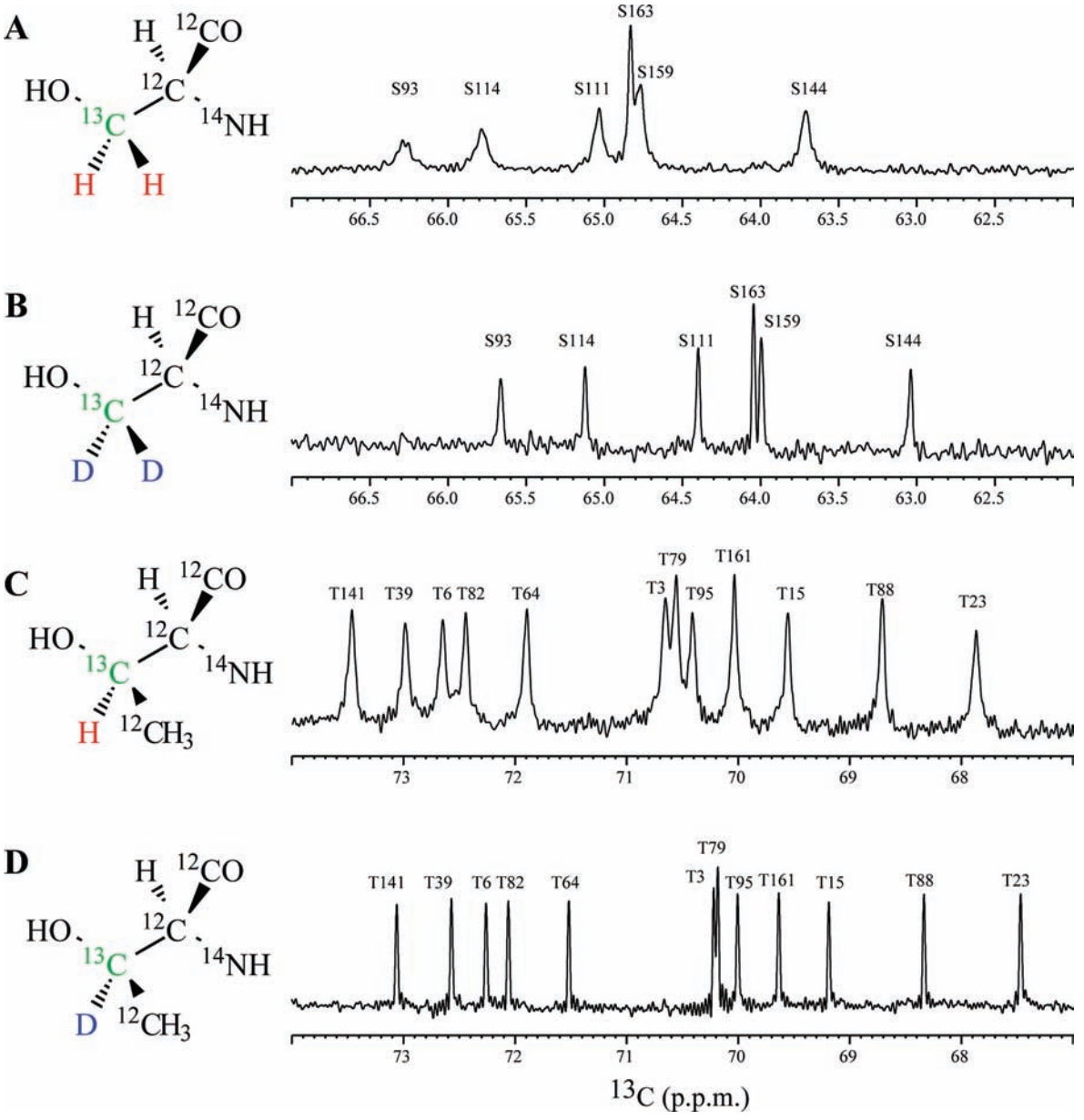


Figure 2. 150.9 MHz [^{2}H,^{1}H]-decoupled 1D ^{13}C NMR spectrum of EPPIb, selectively labeled with L-[β-^{13}C] serine (A), L-[β-^{13}C;β,β-^{2}H$_2$] serine (B), L-[β-^{13}C] threonine (C), and L-[β-^{13}C;β-^{2}H] threonine (D). These spectra were acquired at 40 °C at 150.9 MHz of ^{13}C frequency. In (B) and (D), proton and deuterium decoupling was applied during acquisition. The chemical shift differences between L-[β-^{13}C] serine and L-[β-^{13}C;β,β-^{2}H$_2$] serine and those between L-[β-^{13}C] threonine and L-[β-^{13}C;β-^{2}H] threonine are due to the deuterium isotope shift effect arising from the proton/deuterium substitution of methylene protons. Each peak is labeled with its assignment.

protons. The structural restraints common to the two calculations were the NOE restraints previously collected from SAIL-EPPIb samples,[20] which is exclusively composed of SAIL amino acids, as listed in Supporting Information Figure 1, and the backbone torsion angle restraints obtained from database searches with the program TALOS.[22] Torsion angle dynamics with 20 000 steps per conformer by CYANA generated 100 structures that lacked significant violations against the experimental restraints. These structures were further refined by the AMBER 9 software package[23] with an all-atom force field (ff99SB)[24] and a generalized Born (GB) implicit solvent model.[25] The refinement was comprised of three stages: minimization, molecular dynamics, and minimization. Minimization and molecular dynamics consisted of 1500 steps and a 20 ps duration, respectively. We used the force constants of 50 kcal/(mol × Å^2) and 200 kcal/(mol × rad^2) for the distance and angle restraints, respectively. Among the structures that did not significantly violate the experimental restraints, we selected the 20 lowest energy structures for analysis. The hydrogen acceptor atoms of the side-chain hydrogen bonds were searched using the program MOLMOL[26] with the following criteria: for the distance between the OH/SH protons, maximum distance between the proton and acceptor atom, 2.4 Å; the maximum angle O(S)−H···(acceptor atom), 45°; and the minimum number of conformers that satisfy the criteria, 10 (out of 20).

The structural quality was evaluated with PROCHECK-NMR.[27] The program MOLMOL[26] was used to visualize the structures. The coordinates of the 20 energy-refined conformers of SAIL-EPPIb have been deposited in the Protein Data Bank (accession code 2rs4). The chemical shifts of SAIL-EPPIb, including the assignments of the OH and SH

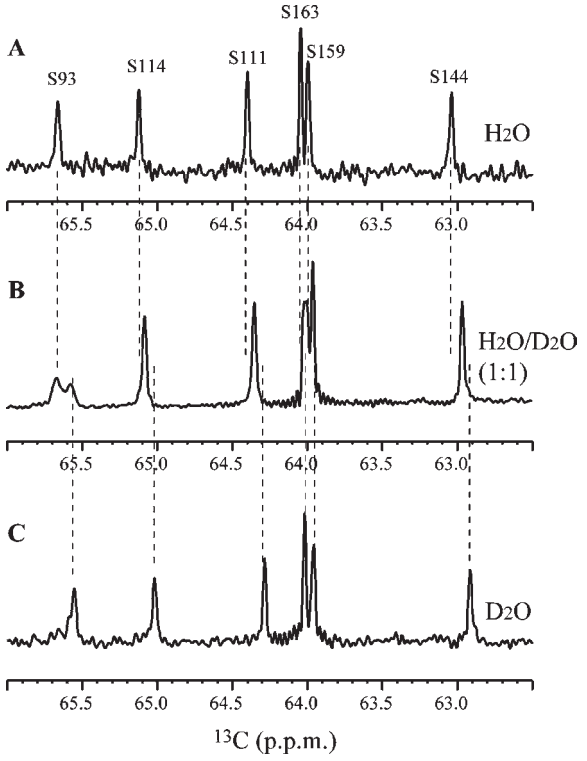


Figure 3. 150.9 MHz [^{2}H,^{1}H]-decoupled 1D ^{13}C NMR spectra of EPPIb, selectively labeled with L-[β-^{13}C;β,β-^{2}H$_2$] serine. The labeled EPPIb was dissolved in H_2O (A), H_2O/D_2O (1:1) (B), and D_2O (C) buffers to concentrations ranging from 0.2 to 0.7 mM for each solution at pH 7.5. Peaks are labeled with their assignments. These NMR experiments were performed at 40 °C. The $^{13}C_\beta$ resonances for Ser-93 were observed as two signals in an H_2O/D_2O (1:1) buffer, due to the isotope effect from the side-chain hydroxyl groups. Overall, the line-width of each peak in an H_2O/D_2O (1:1) buffer is slightly larger than that in H_2O or D_2O, due to the unresolved isotope effect from the backbone amide group.

protons and the distance restraints, have been deposited in the BioMagResBank (accession code 11451).

RESULTS AND DISCUSSION

Optimization of the Isotope Labeling Patterns in Ser and Thr Residues. We previously demonstrated for the EPPIb protein that the exchange rates of the Tyr−OH and Cys−SH groups can be estimated in an aqueous solution, through NMR observations of the carbon atoms attached to the OH/SH groups in an H_2O/D_2O (1:1) solution.[11,12] The success of the previous studies motivated us to apply this approach to an analysis of the Ser−OH and Thr−OH groups in EPPIb.

In Ser and Thr residues, a practically measurable isotope shift effect on their OH group is observed for their $^{13}C_\beta$ atoms. Thus, a high-resolution observation sufficient to resolve the isotopomer-lines of the $^{13}C_\beta$ resonance was a key for success. For this reason, we designed and synthesized L-[β-^{13}C;β,β-^{2}H$_2$] Ser and L-[β-^{13}C;β-^{2}H] Thr (Scheme 1). In these amino acids, the carbon atom at the β position is selectively enriched by ^{13}C, and the hydrogen atoms attached to the $^{13}C_\beta$ atom are deuterated, such that the dipole and scalar couplings involving the $^{13}C_\beta$ atom are extensively eliminated.

These Ser and Thr residues were efficiently incorporated into the EPPIb protein by using an *E. coli* cell-free expression system.[15,16] The prepared sample was dissolved in 100% H_2O buffer, and 1D-^{13}C NMR spectra were acquired at 40 °C. The $^{13}C_\beta$ peaks of 6 Ser and 12 Thr residues were observed at chemical shifts ranging from 63 to 66 ppm and 66 to 73 ppm, respectively (Figure 2B,D). These peaks were assigned according to previously established assignments.[18] Overall, the $^{13}C_\beta$ resonances were very sharp, about 4 Hz, due to the optimized isotope labeling pattern. Especially, the deuteration of the attached protons greatly contributed to the line-narrowing, which was confirmed through a comparison with the line-widths of EPPIb samples selectively labeled with L-[β-^{13}C] Ser and L-[β-^{13}C] Thr, about 10 Hz (Figure 2A,C) (Tables S1,S2).

Observation of Isotope Shift Effects on the $^{13}C_\beta$ Atoms in Ser/Thr Residues. The line-narrowing achieved by the special stable isotope labeling pattern facilitated the detection of isotopomer-resolved $^{13}C_\beta$ peaks in an H_2O/D_2O (1:1) solution. However, caution should be taken in the analysis, because the two-bond isotope shift effect on the side-chain OH group and the three-bond isotope shift effect on the backbone amide (NH) group are observed for the $^{13}C_\beta$ resonances of Ser/Thr residues. Therefore, their line-shapes in an H_2O/D_2O mixture are governed by the exchange rates of the two polar groups, as in the case for the $^{13}C_\beta$ peak of a Cys residue.[11]

In the case of EPPIb, four types of line-shapes were observed for the $^{13}C_\beta$ peaks of Ser/Thr residues in an H_2O/D_2O solution: a four-line peak (Thr-23), a two-line peak with relatively large splitting, 0.08−0.11 ppm (Thr-6, Thr-39, Thr-82, and Ser-93), a two-line peak with relatively small splitting, 0.02 ppm (Thr-161-Ser-163), and a single-line peak (others) (Figures 3, 4; Table S3). Because the size of a three-bond isotope shift effect is often smaller than that of a two-bond shift effect,[11,28,29] and the splitting width of the second group is close to the typical value of the two-bond isotope shift effect on OH groups in sugar rings, about 0.09−0.12 ppm,[28] we concluded that the splits observed for the second and third groups arose from side-chain OH and backbone amide groups, respectively. Taken together, both OH/OD and NH/ND, only OH/OD, and only NH/ND isotopomers were resolved in the first, second, and third groups, respectively. It should be noted that the line-width of the fourth group is slightly broader than that of 100% H_2O or 100% D_2O buffer, which reflects the unresolved isotope effect. Therefore, the exchange rates of the OH groups of Thr-6, Thr-23, Thr-39, Thr-82, and Ser-93 are relatively slower, as compared to the size of the isotope shift effect.

As judged from the spectral simulation, the exchange rate at which the resolution of the isotopomer is observed for $^{13}C_\beta$ resonances is slower than the order of 10 s^{-1}. On the basis of the quantitative estimation by a ^{13}C EXSY experiment[10,30] or a line-shape analysis for the split peaks, the exchange rates of the slowly exchanging OH groups are different from each other, which is correlated with the frequency of the solvent exposure, governed by the dynamic behavior of a protein (Thr-6, <0.2 s^{-1}; Thr-23, 0.5 ± 0.2 s^{-1}; Thr-39, 0.3 ± 0.1 s^{-1}; Thr-82, <0.2 s^{-1}; Ser-93, 26 ± 4 s^{-1}) (Figure S2).

Assignment of Slowly Exchanging OH Protons in Ser/Thr Residues. The exchange rates of slowly exchanging OH groups that give isotopomer-resolved peaks in an H_2O/D_2O solution are slow enough to allow their OH protons to generate NMR-observable ^{1}H resonances.[10] We thus searched for the OH proton signals by NOE and scalar coupling-based experiments.

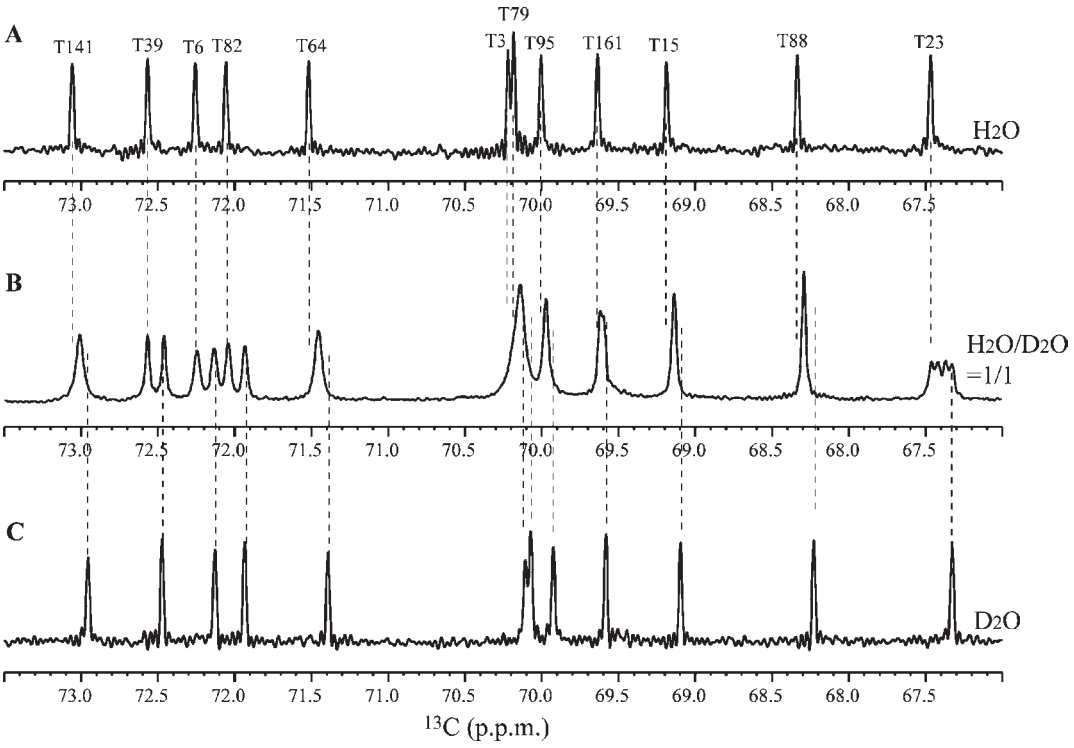


Figure 4. 150.9 MHz [^{2}H,^{1}H]-decoupled 1D ^{13}C NMR spectra of EPPIb, selectively labeled with L-[β-^{13}C;β-^{2}H] threonine in H_2O (A), H_2O/D_2O (1:1) (B), and D_2O (C) buffers. Sample concentrations were 0.2−0.7 mM, and all experiments were performed at 40 °C. Peaks are labeled with their assignments. The $^{13}C_\beta$ resonances for Thr-6, Thr-39, and Thr-82 were observed as two signals in an H_2O/D_2O (1:1) buffer, due to the isotope effect from the side-chain hydroxyl groups. The $^{13}C_\beta$ resonance for Thr-23 was observed as the four-line signals in H_2O/D_2O (1:1), due to the dual isotope effect from the backbone amide and side-chain hydroxyl groups.

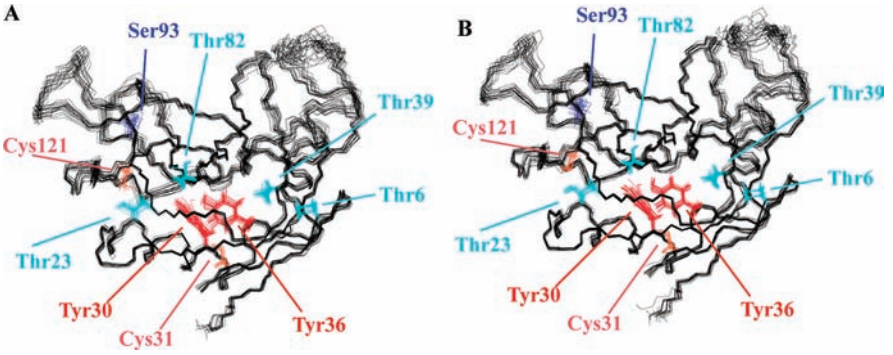


Figure 5. The 20 solution structures of SAIL-EPPIb, determined without (A) and with (B) distance restraints involving observable hydroxyl/sulfhydryl protons, superimposed on a well-defined region (a.a. 1−87, 92−110, 118−136, and 153−162). The side-chains of the Ser, Thr, Cys, and Tyr residues are colored blue, cyan, orange, and red, respectively.

Through the strong intraresidue NOE peaks between the OH protons and the β protons, the discrete OH proton peaks could readily be found at chemical shift values of 6.29, 8.22, 5.67, 6.00, and 6.02 ppm for Thr-6, Thr-23, Thr-39, Thr-82, and Ser-93, respectively (Figure S3). Furthermore, for the scalar-based experiment, 2D HSQC-TOCSY was employed, and correlation peaks between $^{13}C_\beta$ and OH protons were detected for Thr-6, Thr-39, and Thr-82 (Figure S4). However, the correlation peaks for Thr-23 and Ser-93 could not be detected, possibly due to the smaller scalar coupling between H_β and OH, the side-chain motion, and/or the relatively faster hydrogen exchange. While the assignment of the OH protons was performed using SAIL proteins in this case, the assignment could also be accomplished with a conventional ^{13}C, ^{15}N-double labeled protein.

Structure Refinement by NOE Restraints with Hydroxyl and Sulfhydryl Protons. The identification of slowly exchanging OH/SH protons in EPPIb[10,11] prompted us to refine the NMR structure of EPPIb by applying NOE restraints with them. Through an analysis of the NOESY data of SAIL EPPIb proteins,[20] a total of 44 NOE peaks involving OH/SH protons were assigned, and the corresponding NOE restraints were created. It is noteworthy that, based on a spectral simulation, an OH/SH proton with an exchange rate on the order of $10-10^3$ s^{-1} gives rise to a discrete proton signal, while the isotopomer peaks

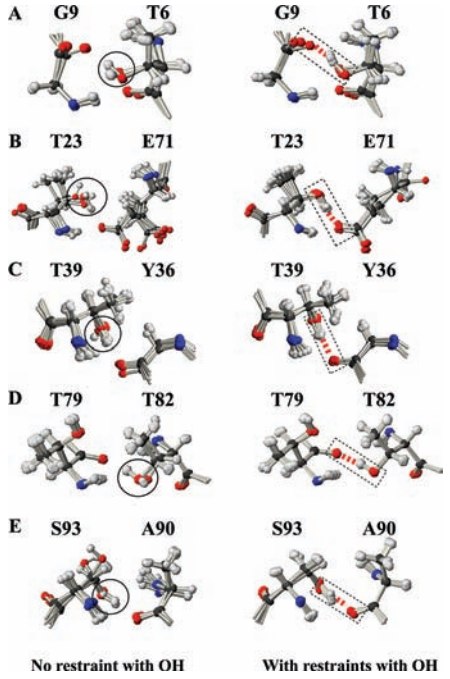


Figure 6. The 10 lowest energy structures of SAIL-EPPIb, superimposed for regions including Thr-6 (A), Thr-23 (B), Thr-39 (C), Thr-82 (D), and Ser-93 (E). For each panel, structures determined without (left) and with (right) NOE distance restraints involving OH are shown. For clarity, conformers were again superimposed for each displayed region. Hydrogen, carbon, nitrogen, and oxygen atoms are colored white, black, blue, and red, respectively. Side-chain hydroxyl groups are circled or, when hydrogen bonds are explicitly formed, enclosed by a dashed box together with its hydrogen-bond acceptor atom. These figures were produced with the MolMol software.[26]

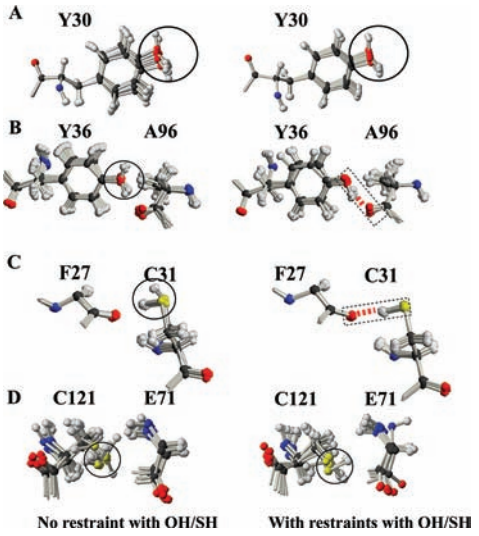


Figure 7. The 10 lowest energy structures of SAIL-EPPIbEPPIb, superimposed for the regions including Tyr-30 (A), Tyr-36 (B), Cys-31 (C), and Cys-121 (D). For each panel, structures determined without (left) and with (right) distance restraints with OH/SH are shown. For clarity, conformers were again superimposed for each displayed region. Side-chain hydroxyl and sulfhydryl groups are circled or, when hydrogen bonds are explicitly formed, enclosed by a dashed box together with its hydrogen-bond acceptor atom.

are unresolved in the observation of the $^{13}C_{\beta}$ signal.[6,7,10] Therefore, we searched for observable NOE peaks involving rapidly exchanging OH groups. In the present case, however, no observable OH protons could be found in the NOESY spectra.

To evaluate the effect of the OH/SH restraints, we performed the restrained simulated annealing with the AMBER force field[24] for the structure of EPPIb in the presence and absence of the OH/SH restraints. By virtue of the large number of NOE restraints,[20] both calculations produced high-quality NMR structures, where the positions of the oxygen/sulfur atoms in the OH/SH groups were fairly well-defined (Figure 5; Table S4). However, the locations of the attached OH/SH protons were ambiguous in the structure calculated without the OH/SH restraints, due to the peculiar hybrid orbital of the oxygen/sulfur atom, which allows three rotamers about the C−O/C−S bond. On the other hand, the positions of the OH/SH protons were well-defined in the structures calculated with the OH/SH restraints, in most cases (Figures 6,7; Table 1).

In a conventional structure calculation, the determination of the rotameric conformations of the OH/SH groups depends only on structural packing effects and/or energy minimization, using a sophisticated force field. The present study highlights the unique and important contribution of the OH/SH restraints to determine the rotameric conformations of the side-chain polar groups (Table 1). Of course, the determination of the conformation by the OH/SH restraints only works well when the positions of the surrounding atoms are precisely defined. In this regard, the accurate determination of the side-chain conformations is also important.

The solution structure of EPPIb obtained by the SAIL method, with the supplementary NOE information associated with the side-chain exchangeable OH and SH groups, showed only a slight improvement with these supplementary NOEs, because the original structure was very accurately determined by the SAIL method. The final structure was quite similar to that obtained by X-ray crystallography (rmsd for the core region is 0.85 Å) (Figure S5). Once the positions of the OH/SH protons are determined, the hydrogen-bond acceptor atoms could readily be identified, based on the relative positions of the atoms in the structure (see Materials and Methods). In the case of EPPIb, seven side-chain hydrogen bonds were determined, and all of them are embedded in the structure of EPPIb (Figure 8, Table 1). A large proportion of the side-chain hydrogen bonds are local, and some of them belong to recurring structural elements (Figure 8C−E,G). For example, the OH groups of Thr-39, Thr-82, and Ser-93 are hydrogen bonded to the backbone carbonyl carbons three residues behind them in the primary sequence (Figure 8C−E). In these cases, the intervening residues are likely to adopt a right-handed α helix-like structure, and the χ_1 angle of the Ser/Thr residue is almost always restricted to the g^+ conformation.[31] The hydrogen bond between the SH group of Cys-31 and the carbonyl carbon three residues behind in the α-helix is also regarded as a typical pattern of an SH-involved in a local hydrogen bond (Figure 8G).[3] On the other hand, long-range hydrogen bonds seem to play an important role in maintaining the tertiary arrangement of the structural elements, as in the case of the hydrogen bond between Tyr−OH and the backbone carbonyl carbon of Ala-96 (Figure 8F). It is particularly interesting that bifurcated hydrogen bonds are formed between Thr-23 and Glu-71 (Figure 8B). In this situation, the side-chain carboxylate of Glu-71 serves as an N-terminal cap, through the

Table 1. Conformations of Side-Chain Hydroxyl and Sulfhydryl Groups in SAIL-EPPIb Structures Determined in the Presence and Absence of Distance Restraints Involving OH/SH

		conformation of side-chain polar group[b]	
polar group	relative solvent accessibility (%)[a]	without restraints with OH/SH	with restraints with OH/SH
Ser-93 OH	0	g^+, 0; g^-, 16; *t*, 0[c]	g^+, 0; g^-, **20**; *t*, 0
Thr-6 OH	0	g^+, 0; g^-, 12; *t*, 8	g^+, 0; g^-, **20**; *t*, 0
Thr-23 OH	0	g^+, 10; g^-, 4; *t*, 6	g^+, **20**; g^-, 0; *t*, 0
Thr-39 OH	0	g^+, 3; g^-, 17; *t*, 0	g^+, 0; g^-, **20**; *t*, 0
Thr-82 OH	0	g^+, 0; g^-, 16; *t*, 4	g^+, 0; g^-, **20**; *t*, 0
Tyr-30 OH	0	*E*,[d] 4; *Z*, 16	*E*, 6; *Z*, 14
Tyr-36 OH	2.3	*E*, 15; *Z*, 5	*E*, **20**; *Z*, 0
Cys-31 SH	0	g^+, 18; g^-, 0; *t*, 2	g^+, **20**; g^-, 0; *t*, 0
Cys-121 SH	0	g^+, 2; g^-, 9; *t*, 9	g^+, 1; g^-, 12; *t*, 7

[a] The relative solvent accessibility was evaluated for each oxygen/sulfur atom by using EPPIb crystal structure[34] (PDB code: 2NUL). [b] The definitions of g^+, g^-, and *t* for the side-chain conformations are identical to those reported by Ippolito.[35] [c] Four out of 20 conformers have different χ^1 angles from the others and are not included. [d] Because the Tyr OH protons are prone to be located in the aromatic ring plane,[33] the conformation of the OH proton is defined by an *E*/*Z* designation. When the Tyr OH proton is proximal to $C_{\varepsilon 1}$, as compared to $C_{\varepsilon 2}$, it is defined as the *Z* conformation; otherwise, it is the *E* conformation. Here, however, the $C_{\delta 1}$ is defined as carbons for which the $C_\alpha-C_\beta-C_\gamma-C_\delta$ dihedral angle has a positive value, 0° to +180°, and $C_{\delta 2}$ has a negative value, −180° to 0°. This definition differs from the previous definition, in which the $C_{\delta 1}$ is defined as the carbon for which the absolute value of the $C_\alpha-C_\beta-C_\gamma-C_\delta$ dihedral angle is smaller than that involving the other C_δ carbon.[36]

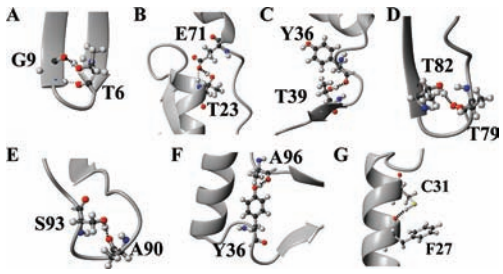


Figure 8. The side-chain hydrogen bonds in EPPIb. The side-chain hydrogen bonds with the hydroxyl groups of Thr-6 (A), Thr-23 (B), Thr-39 (C), Thr-82 (D), Ser-93 (E), Tyr-36 (F), and the sulfhydryl group of Cys-31 (G), those were determined in this study, are depicted as black dashed lines on the lowest-energy SAIL-EPPIb NMR structure. In (B), the hydrogen bond between the Thr-23 amide and the Glu-71 carboxylate is shown by a gray dashed line, to clarify the bifurcated hydrogen bonds.

formation of a hydrogen bond with the amide of Thr-23. The hydrogen bond between the Thr-23 OH group and the Glu-71 carboxylate stabilizes the interaction.

Side-Chain Hydrogen Bonds Involving Slowly Exchanging OH Groups Are Likely To Play an Important Role in Stabilizing the Structure of a Protein. As shown for EPPIb, proteins sometimes contain slowly exchanging and NMR-observable OH/SH protons in their interiors.[6,32] Although their population is small, they are probably involved in structurally or functionally important hydrogen bonds in proteins. It is especially noteworthy that the hydrogen bonds involving the OH groups of Thr-23, Thr-39, and Ser-93 in EPPIb are conserved in structurally homologous proteins (Figure S6), raising the possibility that long-lived side-chain hydrogen bonds are likely to be conserved in homologous proteins, due to their structural importance. Such conservation of water-inaccessible side-chain polar groups in the interiors of homologous proteins has been reported for proteins that satisfy the hydrogen-bonding potential.[33]

CONCLUSION

We have developed a method for monitoring the hydrogen exchange rates of Ser−OH and Thr−OH groups. The method only requires a protein selectively labeled with [β-^{13}C; β,β-^{2}H$_2$]-Ser and/or [β-^{13}C; β-^{2}H]Thr, which can be prepared by conventional cellular expression systems, making it feasible for most NMR laboratories. Using EPPIb, we demonstrated that NOE restraints with NMR-observable OH and SH protons are quite valuable for the refinement of the NMR structure, especially for the residues bearing such OH and SH groups. Because the chemical shift regions of $^{13}C_\beta$ in Ser, Thr, and Cys, and $^{13}C_\zeta$ of Tyr do not overlap each other, it is possible to prepare proteins simultaneously labeled with the four types of amino acids and to perform ^{13}C NMR analyses.[10,11] It should also be pointed out that our approach can be used to detect intermolecular side-chain hydrogen bonds.

ASSOCIATED CONTENT

Supporting Information. ^{13}C NMR EXSY experiment for Ser/Thr−OH groups; structural statistics of the SAIL-EPPIb structure determined with/without distance restraints with OH and SH protons; $^{13}C_\beta$ chemical shifts of Ser/Thr in different isotope-labeling patterns and in 100% H_2O, H_2O/D_2O (1:1), 100% D_2O; experiment for detecting OH and SH protons based on ^{1}H−^{1}H vicinal couplings; side-chain hydrogen bonds in homologous proteins; chemical structures of SAIL amino acids used to prepare SAIL-EPPIb proteins; and comparison between the NMR and crystal structures of EPPIb. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
kainosho@tmu.ac.jp

ACKNOWLEDGMENT

This work was supported by the Targeted Protein Research Program (MEXT) to M.K., a Grant-in-Aid for Young Scientists (B) (23770109) to M.T., and a Grant-in-Aid in Innovative Areas (4104) to M.K. and J.J.

REFERENCES

(1) (a) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97–109. (b) Sticke, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D. *J. Mol. Biol.* **1992**, *226*, 1143–1159.
(2) Eswar, N.; Ramakrishnan, C. *Protein Eng.* **2000**, *13*, 227–238.
(3) (a) Gregoret, L. M.; Rader, S. D.; Fletterick, R. J.; Cohen, F. E. *Proteins* **1991**, *9*, 99–107. (b) Zhou, P.; Tian, F.; Lv, F.; Shang, Z. *Proteins* **2009**, *76*, 151–63.

(4) (a) Havdt, A.; Nielsen, S. O. *Adv. Protein Chem.* **1966**, *21*, 287–386. (b) Englander, S. W.; Downer, N. W.; Teitelbaum, H. *Annu. Rev. Biochem.* **1972**, *41*, 903–924. (c) Woodward, C.; Simon, I.; Tuchsen, E. *Mol. Cell. Biochem.* **1982**, *48*, 135–160. (d) Englander, S. W.; Mayne, L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 243–265. (e) Baldwin, R. L. *Curr. Opin. Struct. Biol.* **1993**, *3*, 84–91. (f) Englander, S. W.; Kallenbach, N. *Q. Rev. Biophys.* **1984**, *16*, 521–655. (g) Bai, Y.; Sosnick, T. R.; Mayne, L.; Englander, S. W. *Science* **1995**, *269*, 192–197. (h) Englander, S. W.; Sosnick, T. R.; Englander, J. J.; Mayne, L. *Curr. Opin. Struct. Biol.* **1996**, *6*, 18–23. (i) Li, R.; Woodward, C. *Protein Sci.* **1999**, *8*, 1571–1590. (j) Krishna, M. M. G.; Hoang, L.; Lin, Y.; Englander, S. W. *Methods* **2004**, *34*, 51–64.

(5) (a) Wagner, G.; Wüthrich, K. *J. Mol. Biol.* **1979**, *130*, 31–37. (b) Wagner, G.; Wüthrich, K. *J. Mol. Biol.* **1979**, *134*, 75–94. (c) Wagner, G.; Wüthrich, K. *J. Mol. Biol.* **1982**, *160*, 343–361. (d) Roder, H.; Wagner, G.; Wüthrich, K. *Biochemistry* **1985**, *24*, 7396–7407.

(6) Liepinsh, E.; Otting, G.; Wüthrich, K. *J. Biomol. NMR* **1992**, *2*, 447–465.

(7) Liepinsh, E.; Otting, G. *Magn. Reson. Med.* **1996**, *35*, 30–42.

(8) Otting, G.; Wüthrich, K. *J. Am. Chem. Soc.* **1989**, *111*, 1871–1875.

(9) Otting, G.; Liepinsh, E.; Wüthrich, K. *J. Am. Chem. Soc.* **1991**, *113*, 4363–4364.

(10) Takeda, M.; Jee, J.; Ono, A. M.; Terauchi, T.; Kainosho, M. *J. Am. Chem. Soc.* **2009**, *131*, 18556–18562.

(11) Takeda, M.; Jee, J.; Terauchi, T.; Kainosho, M. *J. Am. Chem. Soc.* **2010**, *132*, 6254–6260.

(12) (a) Hansen, P. E. *Annu. Rep. NMR Spectrosc.* **1983**, *15*, 105–234. (b) Dziembowska, T.; Hansen, P. E.; Rozwadowski, Z. *Prog. Nucl. Magn. Reson. Spectrosc.* **2004**, *45*, 1–29.

(13) (a) Hawkes, G. E.; Randall, E. W.; Hull, W. E.; Gattegno, D.; Conti, F. *Biochemistry* **1978**, *17*, 3986–3993. (b) Feeney, J.; Partington, P.; Roberts, G. C. K. *J. Magn. Reson.* **1974**, *13*, 268–274. (c) Kainosho, M.; Tsuji, T. *Biochemistry* **1982**, *21*, 6273–6279. (d) Kainosho, M.; Nagao, H.; Tsuji, T. *Biochemistry* **1987**, *26*, 1068–1075. (e) Markley, J. L.; Kainosho, M. In *Stable Isotope Labeling and Resonance Assignments in Larger Proteins: NMR of Macromolecules*; Roberts, G. C. K., Ed.; Oxford University Press: New York, 1993; pp 101−152. (f) Uchida, K.; Markley, J. L.; Kainosho, M. *Biochemistry* **2005**, *44*, 11811–11820.

(14) Kainosho, M.; Torizawa, T.; Iwashita, Y.; Terauchi, T.; Ono, A. M.; Güntert, P. *Nature* **2006**, *440*, 52–57.

(15) Takeda, M.; Ikeya, T.; Güntert, P.; Kainosho, M. *Nat. Protoc.* **2007**, *2*, 2896–2902.

(16) Torizawa, T.; Shimizu, M.; Taoka, M.; Miyano, H.; Kainosho, M. *J. Biomol. NMR* **2004**, *30*, 311–325.

(17) Yuri, N.; Belokon, Y. N.; Bulychev, A. G.; Vitt, S. V.; Struchkov, Y. T.; Batsanov, A. S.; Timofeeva, T. V.; Tsyryapkin, V. A.; Ryzhov, M. C.; Lysova, L. A.; Bakhmutov, V. I.; Belikov, V. M. *J. Am. Chem. Soc.* **1985**, *107*, 4252–4259.

(18) Kariya, E.; Ohki, S.; Hayano, T.; Kainosho, M. *J. Biomol. NMR* **2000**, *18*, 75–76.

(19) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1969**, *52*, 335–338.

(20) Takeda, M.; Terauchi, T.; Ono, A. M.; Kainosho, M. *J. Biomol. NMR* **2010**, *46*, 45–49.

(21) (a) Mumenthaler, C.; Güntert, P.; Braun, W.; Wüthrich, K. *J. Biomol. NMR* **1997**, *10*, 351–362. (b) Güntert, P. *Methods Mol. Biol.* **2004**, *278*, 353–378. (c) Herrmann, T.; Güntert, P.; Wüthrich, K. *J. Mol. Biol.* **2002**, *319*, 209–227.

(22) Cornilescu, G.; Delaglio, F.; Bax, A. *J Biomol. NMR* **1999**, *13*, 289–302.

(23) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(24) (a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197. (b) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–25.

(25) Tsui, V.; Case, D. A. *Biopolymers* **2000**, *56*, 275–291.

(26) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51–55.

(27) Laskowski, R. A.; Rullmann, J. A. C.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J. Biomol. NMR* **1996**, *8*, 477–486.

(28) Bosco, M.; Picotti, F.; Radoicovich, A.; Rizzo, R. *Biopolymers* **2000**, *53*, 272–280.

(29) Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 8070–8075.

(30) (a) Fischer, M. W. F.; Zeng, L.; Zuiderweg, E. R. P. *J. Am. Chem. Soc.* **1996**, *118*, 12457–12458. (b) Bertini, I.; Felli, I. C.; Kümmerle, R.; Moskau, D.; Pierattelli, R. *J. Am. Chem. Soc.* **2004**, *126*, 464–465.

(31) Eswar, N.; Ramakrishnan, C. *Protein Eng.* **1999**, *12*, 447–455.

(32) (a) Löhr, F.; Mayhew, S. G.; Rüterjans, H. *J. Am. Chem. Soc.* **2000**, *122*, 9289–9295. (b) Plevin, M. J.; Hayashi, I.; Ikura, M. *J. Am. Chem. Soc.* **2008**, *130*, 14918–14919.

(33) Worth, C. L.; Blundell, T. L. *Proteins* **2009**, *75*, 413–429.

(34) Edwards, K. J.; Ollis, D. L.; Dixon, N. E. *J. Mol. Biol.* **1997**, *271*, 258–265.

(35) Ippolito, J. A.; Alexandera, R. S.; Christianson, D. W. *J. Mol. Biol.* **1990**, *215*, 457–471.

(36) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. *J. Mol. Biol.* **1998**, *280*, 933–952.