SUPPORTING INFORMATION FOR:

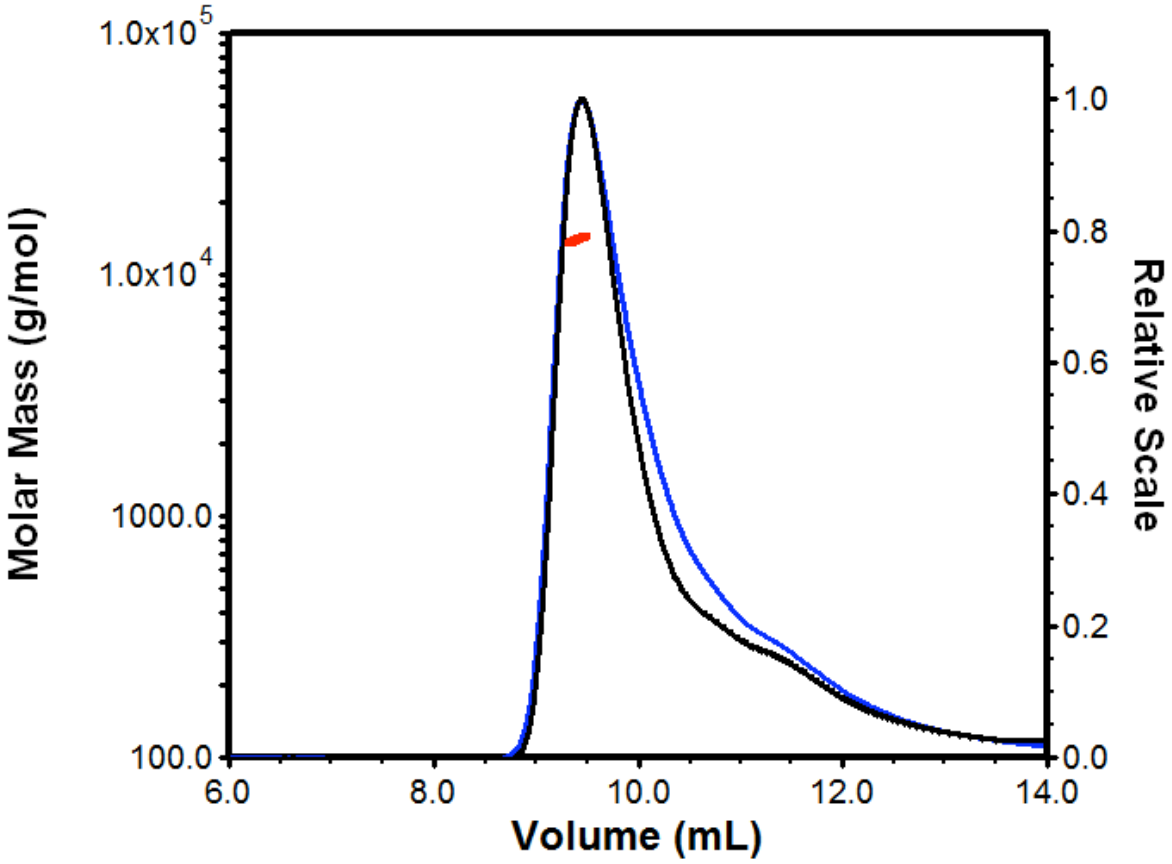
Solution NMR structure of Alr2454 from *Nostoc sp.* PCC 7120, the first structural representative of Pfam domain family PF11267

James M. Aramini,[1] Donal Petrey,[2] Dong Yup Lee,[1] Haleema Janjua,[1] Rong Xiao,[1] Thomas B. Acton,[1] John K. Everett,[1] and Gaetano T. Montelione[1,3]

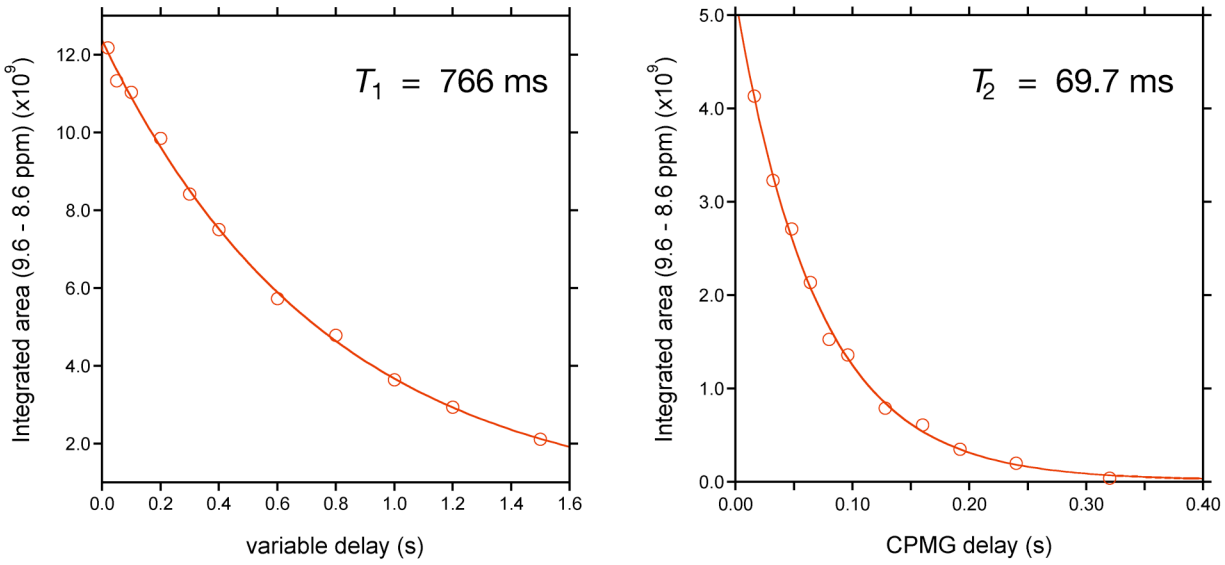
[1] *Center for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854, USA, and the Northeast Structural Genomics Consortium*

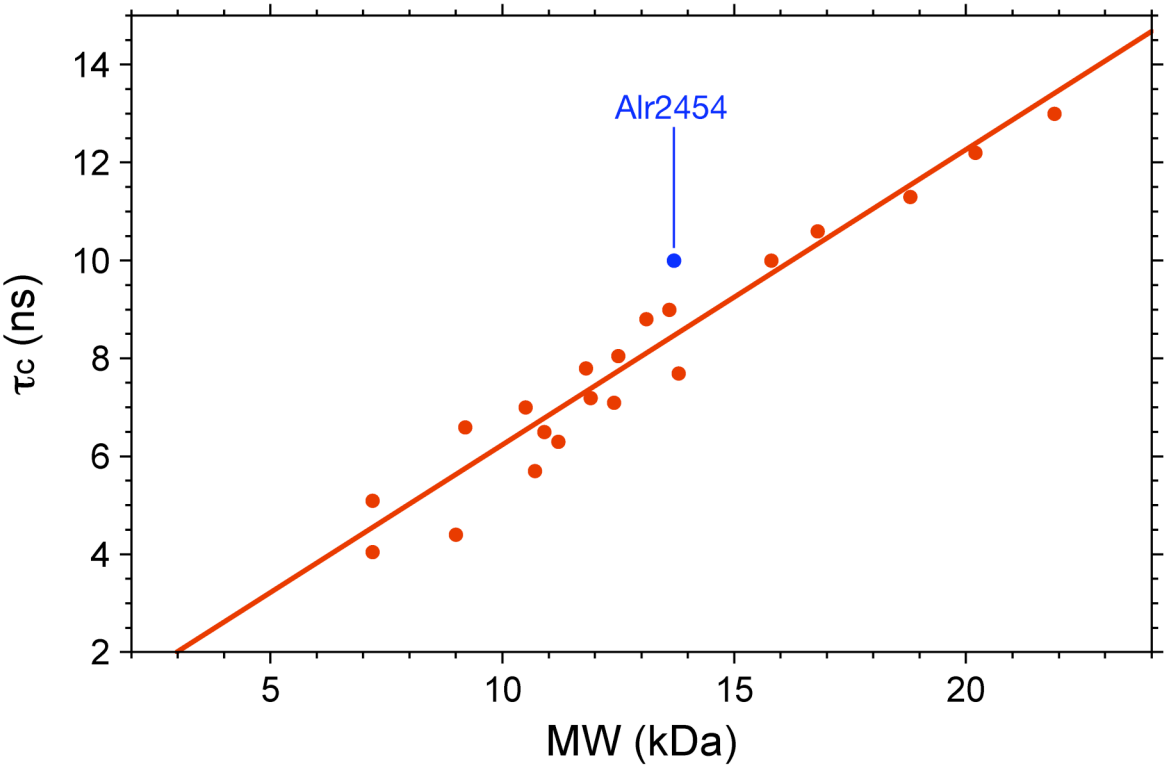
[2] *Department of Biochemistry and Molecular Biophysics, Center for Computational Biology and Bioinformatics, Columbia University, New York, New York 10032, USA, and the Northeast Structural Genomics Consortium*

[3] *Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854, USA*



Suppl. Fig. S1. Static light scattering results for *Nostoc sp.* Alr2454. Data were collected on a miniDAWN (TREOS) Light Scattering Instrument (Wyatt Technology) at room temperature on an NMR sample (30 µl) of [*U*-^{13}C,^{15}N]-Alr2454 (0.8 mM in 20 mM ammonium acetate, 100 mM NaCl, 10 mM DTT, 5 mM $CaCl_2$, 0.02% (w/v) NaN_3, 10% ^2H_2O at pH 4.5). The sample was injected onto an analytical gel filtration column (Protein KW-802.5, Shodex, Japan; flow rate, 0.5 ml/min) with the effluent monitored by UV absorbance (black trace; 280 nm) and 90° static light scattering (blue trace) detectors. The resulting experimental molecular weight of Alr2454 is 13.8 kDa (red); the expected MW including affinity tag and isotope enrichment is 13.7 kDa.

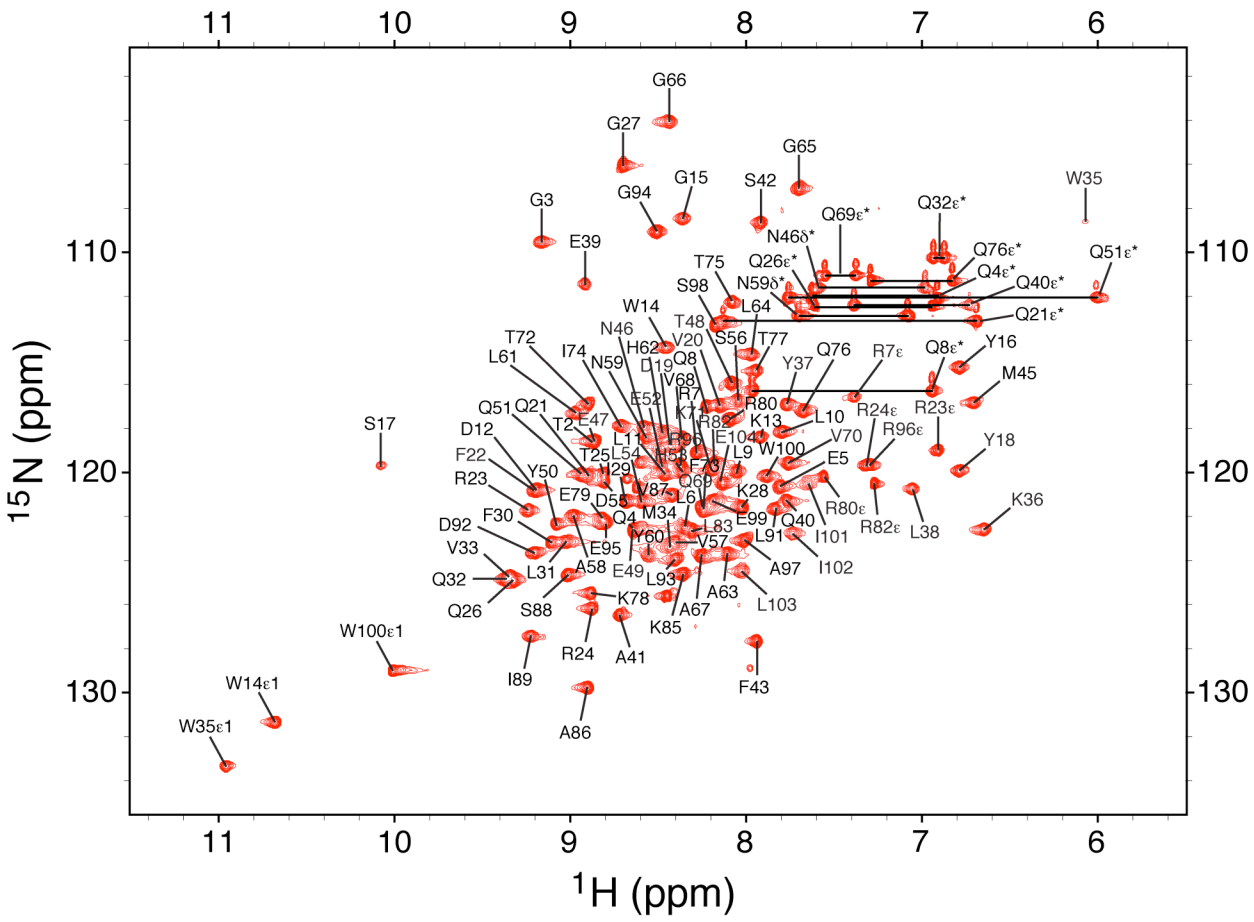

T_1 = 766 ms
Integrated area (9.6 - 8.6 ppm) (x10^9)
variable delay (s)
T_2 = 69.7 ms
Integrated area (9.6 - 8.6 ppm) (x10^9)
CPMG delay (s)


Alr2454
τc (ns)
MW (kDa)

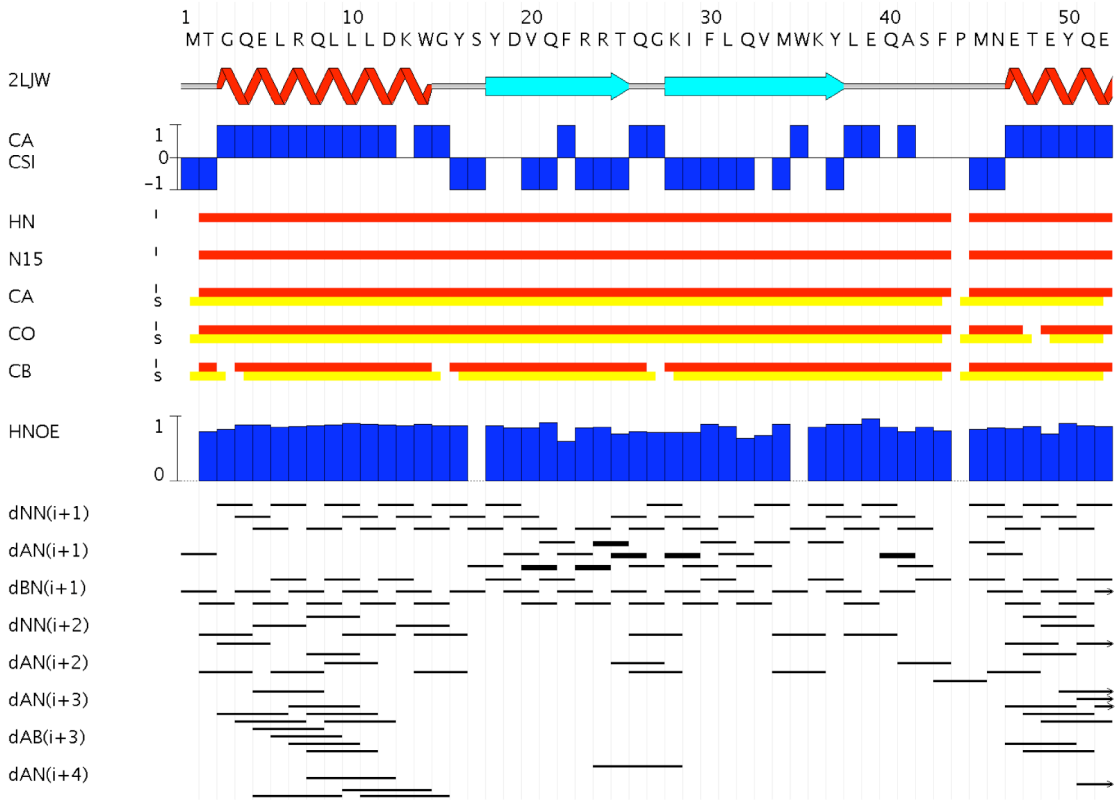
Suppl. Fig. S2. 1D ^{15}N T_1 and T_2 relaxation data for [*U*-^{13}C,^{15}N]-Alr2454. The data were acquired on a Bruker AVANCE 600 MHz spectrometer at 298 K using pseudo-2D ^{15}N T_1 and T_2 gradient experiments [1]. T_1 spectra were acquired with delays, T = 20, 50, 100, 200, 300, 400, 600, 800, 1000, 1200 and 1500 ms, and a relaxation delay of 3 s. T_2 spectra were acquired with CPMG delays, T = 16, 32, 48, 64, 80, 96, 128, 160, 192, 240, and 320 ms, and with a relaxation delay of 1.5 s. (Top): ^{15}N T_1 and T_2 values were extracted by plotting the decay of integrated $^{1}H^{N}$ intensity between $\delta \approx 8.6$ to 9.6 ppm and fitting the curves with standard exponential equations using the program 't1guide' within TopSpin 2.1 (Bruker BioSpin). (Bottom): Plot of rotational correlation time, τ_c (ns), versus protein molecular weight (kDa) for known monomeric NESG targets of ranging size (taking into account isotope enrichment as well as affinity tags in the sequence). ^{15}N T_1/T_2 data for all monomeric proteins used for the τ_c vs. MW plot were obtained on the same Bruker 600 MHz spectrometer at 298 K, and analyzed as described above. For each protein, the τ_c was calculated from the ^{15}N T_1/T_2 ratio using the following approximation of literature relaxation equations [2, 3]:

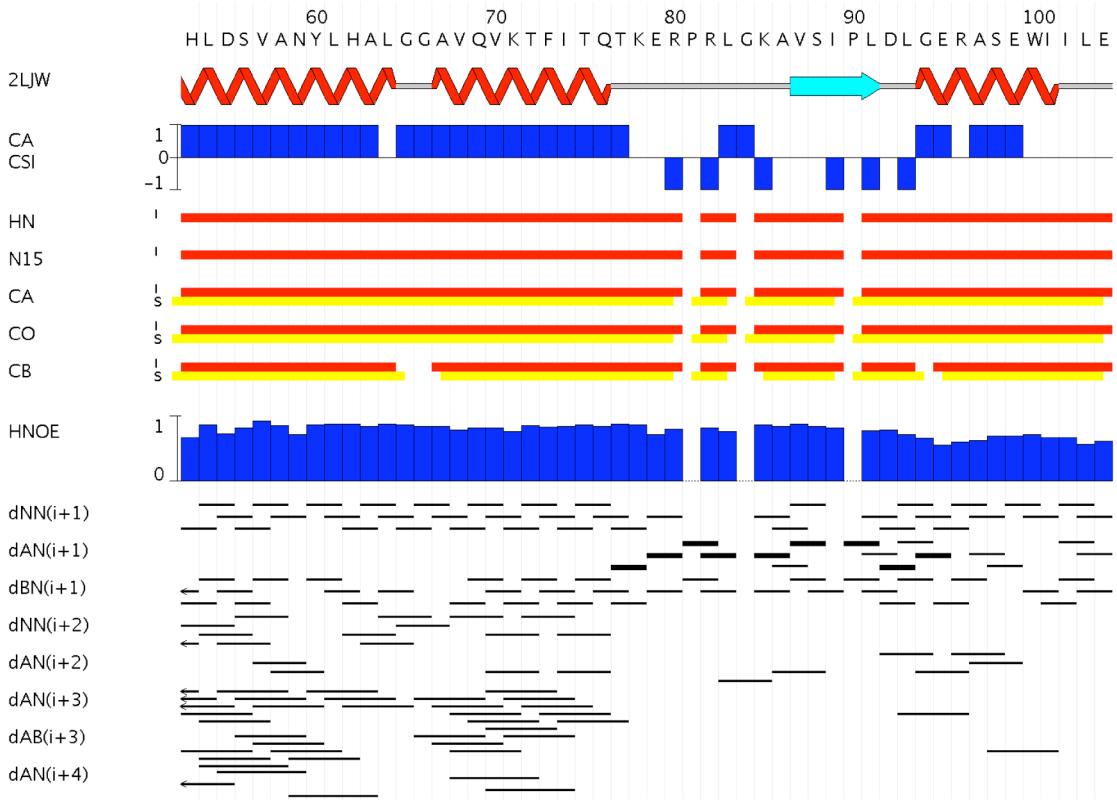
$$\tau_c \approx \left(\sqrt{\frac{6T_1}{T_2} - 7} \right) / 4\pi\nu_N \qquad (1)$$

where ν_N is the resonance frequency of ^{15}N in Hz. Using this approach, we obtain a τ_c of 10.0 ns for [*U*-^{13}C,^{15}N]-Alr2454, shown in blue, which is consistent with a monomer (expected MW = 13.7 kDa, including C-terminal affinity tag). Protocols for this analysis are available on the NESG Wiki protocol site (http://www.nmr2.buffalo.edu/nesg.wiki/Main_Page).



Suppl. Fig. S3. Two dimensional ^{1}H-^{15}N HSQC spectrum of 1.01 mM [*U*-^{13}C,^{15}N]-Alr2454 in 90% H_2O / 10% $^{2}H_2O$ solution containing 20 mM ammonium acetate, 100 mM NaCl, 10 mM DTT, 5 mM $CaCl_2$, 0.02% (w/v) NaN_3, at pH 4.5 collected at 25^{o}C on a Bruker AVANCE 800 MHz spectrometer. Backbone resonance assignments are labeled with one-letter amino acid codes followed by their sequence numbers. Assigned side chain NH resonances of Trp and Arg (aliased) as well as side chain NH_2 resonances of Asn and Gln are also indicated. Note that the backbone amide resonances of S17 and W35 are extremely weak, indicative of local conformational exchange.


1 10 20 30 40 50
MTGQELRQLLLDKWGYSYDVQFRRTQGKI FLQVMWKYLEQASFPMNETEYQE
2LJW
CA
CSI
1
0
-1
HN
N15
CA
CO
CB
HNOE
dNN(i+1)
dAN(i+1)
dBN(i+1)
dNN(i+2)
dAN(i+2)
dAN(i+3)
dAB(i+3)
dAN(i+4)


60 70 80 90 100
HLDSVANYLHALGGAVQVKTFI TQTKERPRLGKAVSI PLDLGERASEWI I LE
2LJW
CA
CSI
1
0
-1
HN
N15
CA
CO
CB
HNOE
dNN(i+1)
dAN(i+1)
dBN(i+1)
dNN(i+2)
dAN(i+2)
dAN(i+3)
dAB(i+3)
dAN(i+4)

Suppl. Fig. S4. NMR connectivity map, generated using AutoAssign software [4], summarizing data used to determine resonance assignments and secondary structure for *Nostoc sp.* Alr2454. The final six unassigned histidines in the C-terminal tag have been omitted. Intraresidue (i) and sequential (s) connectivities for the three-rung assignment strategy [4] matching intraresidue and sequential C′, C^{α}, and C^{β} resonances are shown as horizontal red and yellow lines, respectively. Interresidue NOE connectivities are shown as thin, medium, and thick black lines, corresponding to weak, medium, and strong NOE interactions. Bar graphs of CSI [5] and ^{1}H-^{15}N heteronuclear NOE data are shown in blue. The secondary structural elements in the final Alr2454 NMR structure (2LJW) are also shown. In general, the secondary structural elements in the final structure are well-defined by the CSI and NOESY patterns. Protocols for this analysis are available on the NESG Wiki protocol site (http://www.nmr2.buffalo.edu/nesg.wiki/Main_Page).

Supplementary References

1. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay, LE (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. Biochemistry 33:5984-6003.
2. Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry 28:8972-8979.
3. Fushman D, Weisemann R, Thüring H, Rüterjans H (1994) Backbone dynamics of ribonuclease T1 and its complex with 2´-GMP studied by two-dimensional heteronuclear NMR spectroscopy. J Biomol NMR 4:61-78.
4. Moseley HNB, Monleon D, Montelione GT (2001) Automatic determination of protein backbone resonance assignments from triple resonance nuclear magnetic resonance data. Methods Enzymol 339:91-108.
5. Wishart DS, Sykes BD (1994) The ^{13}C Chemical-Shift Index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. J Biomol NMR 4:171-180.