Journal of the Korean Magnetic Resonance Society **2018**, *22*, 125-131
DOI 10.6564/JKMRS.2018.22.4.125

Protein-ligand interaction investigated by HSQC titration study

Joon-Hwa Lee*

Department of Chemistry and RINS, Gyeongsang National University, Jinju, Gyeongnam 52828, Republic of Korea

Received Nov 16, 2018; Revised Dec 13, 2018; Accepted Dec 15, 2018

Abstract Chemical shift perturbation (CSP) is a simple NMR technique for studying binding of a protein to various ligands. CSP is the only technique that can directly provide both a value for the dissociation constant and a binding site from the same set of measurements. To accurately analyze the CSP data, the exact binding mode such as multiple binding, should be carefully considered. In this review, we analyzed systematically the CSP data with multiple modes. This analysis might provide insight into the mechanism on how proteins selectively recognize their target ligands to achieve the biological function.

Keywords NMR, chemical shift perturbation, protein-ligand interaction, HSQC, titration

Introduction

CSP is a simple NMR technique for studying binding of a protein to various ligands.[1] The unlabeled ligand, which can be a small molecule or another macromolecule (protein and nucleic acid), is titrated into the ^{15}N-labeled protein, monitored at each stage of the titration by acquiring a two-dimensional (2D) HSQC spectrum.[2] The chemical shift change, that can be measured accurately, is very sensitive to structural change and provides the information about almost genuine binding interaction.[1] The peaks that experience significant shift are likely to map the binding site for the ligand.[1] Moreover, the titration curve could be fitted to obtain a value for the dissociation constant of the ligand, K_d.[3] Thus, CSP is thought to be the unique technique that can directly provide both a K_d value and a binding site from the single same set of NMR measurement.[3] The chemical shift change can be detected, only if a ligand binds to proteins, unlike other techniques such as spectrophotometry, calorimetry, or enzyme assay.[3] For this reason, CSP is widely used in drug discovery.[3] CSPs are experimentally and conceptually simple and thus it requires little analytical effort. However, to accurately analyze the CSP application, there are a range of issues that should be considered carefully; most importantly with multiple binding modes. Thus, in this review, we analyzed systematically the CSP data with multiple binding modes. This analysis might provide insight into the mechanism on how proteins selectively recognize their target ligands to achieve the biological function.

One-site Binding Model

We think of a one-site binding model, where a protein (P) reversibly binds to a ligand (L) at a single site, given by:

$$\mathrm{P} + \mathrm{L} \overset{K_d}{\leftrightarrow} \mathrm{PL}$$

characterized by the dissociation constant K_d:

* Address correspondence to: **Joon-Hwa Lee,** Department of Chemistry and RINS, Gyeongsang National University, Jinju, Gyeongnam 52828, Republic of Korea, Tel: 82-55-772-1490; E-mail: joonhwa@gnu.ac.kr

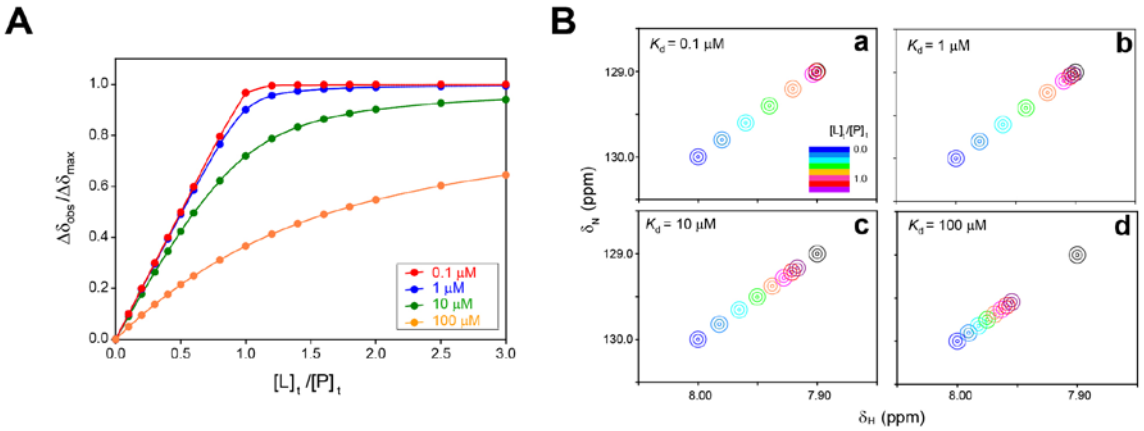

Figure 1. One-site binding model. (A) The simulated titration curves for P with L as a function of $[L]_t/[P]_t$ ratio using Eq. 14. The colors used to illustrate the K_d are: red, 0.1 μM; blue, 1 μM; green, 10 μM; orange, 100 μM. The initial value of $[P]_t$ is 0.1 mM. (B) The calculated $^1H/^{15}N$-HSQC cross-peaks of P during titration with L using K_d of (a) 0.1 μM, (b) 1 μM, (c) 10 μM, or (d) 100 μM. The cross-peak color changes gradually from blue to purple according to the $[L]_t/[P]_t$ ratio. The black cross-peak indicates the bound state (PL).

$$K_d = \frac{[P][L]}{[PL]} \quad (1)$$

where [P], [L], and [PL] represent the concentrations of free protein, free ligand, and protein-ligand complex (PL), respectively.

The total concentration of added proteins ($[P]_t$) and ligands ($[L]_t$) are given by:

$$[P]_t = [P] + [PL] \quad (2)$$

$$[L]_t = [L] + [PL] \quad (3)$$

Eqs. 1 and 3 give rise to:

$$[L]_t = [L] + \frac{[P]}{K_d}[L] \quad (4)$$

Eq. 4 becomes Eq. 5:

$$\frac{[L]}{[L]_t} = \frac{K_d}{K_d+[P]} \quad (5)$$

Becuase $[PL] = [L]_t - [L]$, $[PL]/[L]_t$ is described as:

$$\frac{[PL]}{[L]_t} = \frac{[P]}{K_d+[P]} \quad (6)$$

Eqs. 2 and 6 give rise to

$$[P]_t = [P] + [L]_t\frac{[P]}{K_d+[P]} \quad (7)$$

Eq. 7 becomes Eq. 8:

$$[P]^2 + [P]([L]_t - [P]_t + K_d) - K_d[P]_t = 0 \quad (8)$$

The solution of Eq. 8 is given by:

$$[P] = \frac{\sqrt{([L]_t+[P]_t+K_d)^2-4[L]_t[P]_t}-([L]_t-[P]_t+K_d)}{2} \quad (9)$$

The observed chemical shift (δ_{obs}) in fast exchange is the weighted average of the shifts in the free (δ_{free}) and bound (δ_{bound}) states:

$$\delta_{obs} = \delta_{free}\frac{[P]}{[P]_t} + \delta_{bound}\frac{[PL]}{[P]_t} \quad (10)$$

The change in the observed chemical shift from the free state ($\Delta\delta_{obs} = \delta_{obs} - \delta_{free}$) is given by:

$$\Delta\delta_{obs} = \delta_{bound}\frac{[PL]}{[P]_t} - \delta_{free}\left(1 - \frac{[P]}{[P]_t}\right) \quad (11)$$

Because $[PL] = [P]_t - [P]$, the $\Delta\delta_{obs}$ value becomes:

$$\Delta\delta_{obs} = \Delta\delta_{max}\frac{[PL]}{[P]_t} = \Delta\delta_{max}\left(1 - \frac{[P]}{[P]_t}\right) \quad (12)$$

where $\Delta\delta_{max}$ is the maximum chemical shift change upon saturation ($\Delta\delta_{max} = \delta_{bound} - \delta_{free}$). Using Eq. 9, the $\Delta\delta_{obs}/\Delta\delta_{max}$ ratio can be described as:

$$\frac{\Delta\delta_{obs}}{\Delta\delta_{max}} = \frac{[L]_t+[P]_t+K_d-\sqrt{([L]_t+[P]_t+K_d)^2-4[L]_t[P]_t}}{2[P]_t} \quad (13)$$

$$\frac{\Delta\delta_{obs}}{\Delta\delta_{max}} = \frac{\left(1+\frac{[L]_t}{[P]_t}+\frac{K_d}{[P]_t}\right)-\sqrt{\left(1+\frac{[L]_t}{[P]_t}+\frac{K_d}{[P]_t}\right)^2-4\frac{[L]_t}{[P]_t}}}{2} \quad (14)$$

This equation allow us to fit K_d from the chemical shift perturbation data as a function of $[L]_t/[P]_t$ ratio (Fig. 1). During titration, the concentrrations of protein and ligand should be treated carefully to achieve imporved fitting.[4] In addition, in order to more accurately estimate the K_d value, the titration data for many amide signals should be globally fitted; that is, using the same value of K_d but different $\Delta\delta_{obs}$.[4] It has been reported that, based on Eq. 14, a good estimate of K_d can only be obtained if $[P]_t$ and $[L]_t$ are similar to K_d.[5,6] These studies showed that the optimal value of $[P]_t$ is $0.5 \times K_d$.[5,6]

Two-site Binding Model: One Protein and Two Ligands

Next, we think of a two-site binding model, where a protein (P) reversibly binds to two ligands (L) at two binding sites, given by:

$$P + L \overset{K_{d,1}}{\longleftrightarrow} PL$$

$$PL + L \overset{K_{d,2}}{\longleftrightarrow} PL_2$$

characterized by the dissociation constants, $K_{d,1}$ and $K_{d,2}$, of the PL and PL_2 complexes, respectively:

$$K_{d,1} = \frac{[P][L]}{[PL]} \quad (15)$$

$$K_{d,2} = \frac{[PL][L]}{[PL_2]} \quad (16)$$

where $[PL_2]$ represents the concentration of the PL_2 complex.

In this system, the $[P]_t$ and $[L]_t$ are given by:

$$[P]_t = [P] + [PL] + [PL_2] \quad (17)$$

$$[L]_t = [L] + [PL] + 2[PL_2] \quad (18)$$

Eqs. 15 – 17 give rise to:

$$[P]_t = [P] + \frac{[L]}{K_{d,1}}[P] + \frac{[L]^2}{K_{d,1}K_{d,2}}[P] \quad (19)$$

Eq. 19 becomes Eq. 20:

$$[P]_t = [P]\frac{K_{d,1}K_{d,2}+K_{d,2}[L]+[L]^2}{K_{d,1}K_{d,2}} \quad (20)$$

From Eq. 20, $[P]/[P]_t$ is expressed by:

$$\frac{[P]}{[P]_t} = \frac{K_{d,1}K_{d,2}}{K_{d,1}K_{d,2}+K_{d,2}[L]+[L]^2} \quad (21)$$

From Eqs. 15, 16 and 21, $[PL]/[P]_t$ and $[PL_2]/[P]_t$ can be expressed by:

$$\frac{[PL]}{[P]_t} = \frac{[P]}{[P]_t}\frac{[L]}{K_{d,1}} = \frac{K_{d,2}[L]}{K_{d,1}K_{d,2}+K_{d,2}[L]+[L]^2} \quad (22)$$

$$\frac{[PL_2]}{[P]_t} = \frac{[PL]}{[P]_t}\frac{[L]}{K_{d,2}} = \frac{[L]^2}{K_{d,1}K_{d,2}+K_{d,2}[L]+[L]^2} \quad (23)$$

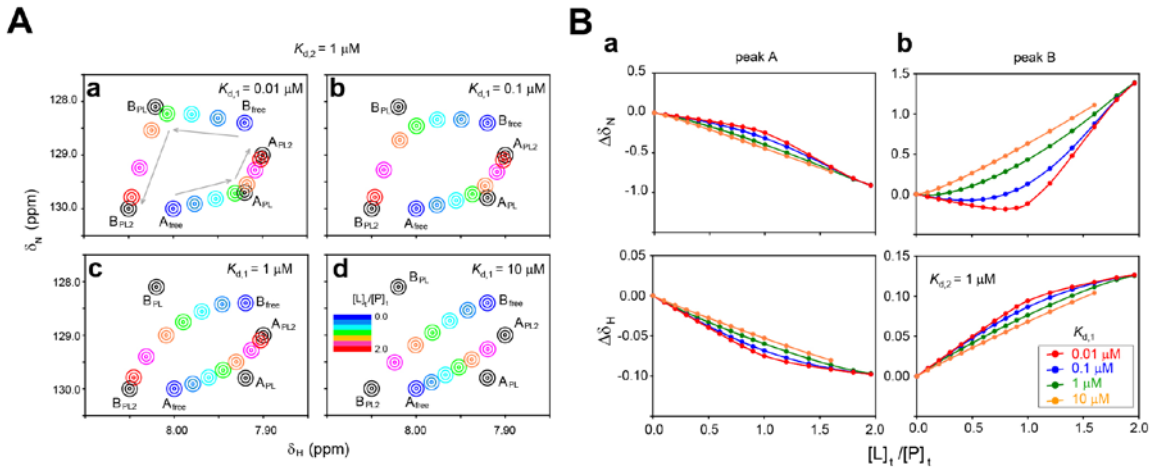

Figure 2. Two-site binding model: one protein and two ligands. (A) The calculated $^{1}H/^{15}N$-HSQC cross-peaks of P during titration with L using $K_{d,2}$ of 1 μM and $K_{d,1}$ of (a) 0.01 μM, (b) 0.1 μM, (c) 1 μM, or (d) 10 μM. The cross-peak color changes gradually from blue to purple according to the $[L]_t/[P]_t$ ratio. The black cross-peak indicates the two bound states (PL and PL_2). (B) The simulated titration curves for P (a: peak A; b: peak B in Figure 2A) with L as a function of $[L]_t/[P]_t$ ratio using Eq. 29 ($K_{d,2}$ = 1 μM). The colors used to illustrate the $K_{d,1}$ are: red, 0.01 μM; blue, 0.1 μM; green, 1 μM; orange, 10 μM. The initial value of $[P]_t$ is 0.1 mM.

Eqs. 18, 22, 23 give rise to

$$[L]_t = [L] + [P]_t \frac{K_{d,2}[L]+2[L]^2}{K_{d,1}K_{d,2}+K_{d,2}[L]+[L]^2} \quad (24)$$

Eq. 24 becomes the following cubic equation:

$$[L]^3 + a[L]^2 + b[L] + c = 0 \quad (25)$$

where,

$$a = 2[P]_t - [L]_t + K_{d,2}$$

$$b = K_{d,2}([P]_t - [L]_t + K_{d,1})$$

$$c = -K_{d,1}K_{d,2}[L]_t$$

The closed-form solution of Eq. 25 is given by:[4]

$$[L] = -\frac{a}{3} + \frac{2}{3}\sqrt{a^2 - 3b}\cos\left(\frac{\theta}{3}\right) \quad (26)$$

where,

$$\theta = \arccos\left(\frac{-2a^3+9ab-27c}{2\sqrt{(a^2-3b)^3}}\right)$$

The observed chemical shift change is the weight average of the shifts in the free form (δ_{free}) and the PL (δ_{PL}) and PL_2 (δ_{PL2}) complexes:

$$\delta_{obs} = \delta_{free}\frac{[P]}{[P]_t} + \delta_{PL}\frac{[PL]}{[P]_t} + \delta_{PL2}\frac{[PL_2]}{[P]_t} \quad (27)$$

The change in the observed chemical shift from the free state ($\Delta\delta_{obs}$) is given by:

$$\Delta\delta_{obs} = \Delta\delta_{PL}\frac{[PL]}{[P]_t} + \Delta\delta_{PL2}\frac{[PL_2]}{[P]_t} \quad (28)$$

where $\Delta\delta_{PL}$ and $\Delta\delta_{PL2}$ are the maximum chemical shift changes upon saturation by the PL and PL_2 complex states, respectively. Using Eqs. 22 and 23, $\Delta\delta_{obs}$ can be described as:

$$\Delta\delta_{obs} = \Delta\delta_{PL2}\frac{K_{d,2}[L]\frac{\Delta\delta_{PL}}{\Delta\delta_{PL2}}+[L]^2}{K_{d,1}K_{d,2}+K_{d,2}[L]+[L]^2} \quad (29)$$

where $\Delta\delta_{max} = \Delta\delta_{PL2}$ and [L] is expressed as Eq. 26. This equation allow us to fit $K_{d,1}$ and $K_{d,2}$ from the

chemical shift perturbation data as a function of $[L]_t/[P]_t$ ratio (Fig. 2). As shown in Figure 2, the observed amide signals changed the direction of their movement after achieving a certain position, when two amide signals for the PL and PL_2 states exhibit significant chemical shift difference. In the case that $K_{d,1} >> K_{d,2}$, the movement of the amide signals look like a straight line (Fig. 2A(d)).

Two-site Binding Model: Two Proteins and One Ligand

We can also think of a two-site binding model, where two proteins (P) reversibly binds to a ligand (L) at two binding sites of large-sized ligand such as DNA duplex, given by:

$$P + L \overset{K_{d,1}}{\longleftrightarrow} PL$$

$$P + PL \overset{K_{d,2}}{\longleftrightarrow} P_2L$$

characterized by the dissociation constants, $K_{d,1}$ and $K_{d,2}$, of the PL and P_2L complexes, respectively:

$$K_{d,1} = \frac{[P][L]}{[PL]} \tag{30}$$

$$K_{d,2} = \frac{[P][PL]}{[P_2L]} \tag{31}$$

where where $[P_2L]$ represents the concentration of the P_2L complex.
In this system, like Eqs. 17 and 18, the $[P]_t$ and $[L]_t$ are given by:

$$[P]_t = [P] + [PL] + 2[P_2L] \tag{32}$$

$$[L]_t = [L] + [PL] + [P_2L] \tag{33}$$

where $[P_2L]$ is concentration of P_2L state. Eqs. 30, 31, and 33 give rise to:

$$[L]_t = [L] + \frac{[P]}{K_{d,1}}[L] + \frac{[P]^2}{K_{d,1}K_{d,2}}[L] \tag{34}$$

Eq. 34 becomes Eq. 35:

$$[L]_t = [L]\frac{K_{d,1}K_{d,2} + K_{d,2}[P] + [P]^2}{K_{d,1}K_{d,2}} \tag{35}$$

From Eq. 35, $[L]/[L]_t$ is expressed by:

$$\frac{[L]}{[L]_t} = \frac{K_{d,1}K_{d,2}}{K_{d,1}K_{d,2} + K_{d,2}[P] + [P]^2} \tag{36}$$

From Eqs. 30, 31 and 36, $[PL]/[L]_t$ and $[P_2L]/[L]_t$ can be expressed by:

$$\frac{[PL]}{[L]_t} = \frac{K_{d,2}[P]}{K_{d,1}K_{d,2} + K_{d,2}[P] + [P]^2} \tag{37}$$

$$\frac{[P_2L]}{[L]_t} = \frac{[P]^2}{K_{d,1}K_{d,2} + K_{d,2}[P] + [P]^2} \tag{38}$$

Eqs. 32, 37, and 38 give rise to:

$$[P]_t = [P] + [L]_t\frac{K_{d,2}[P] + 2[P]^2}{K_{d,1}K_{d,2} + K_{d,2}[P] + [P]^2} \tag{39}$$

Eq. 39 becomes the following cubic equation:

$$[P]^3 + a[P]^2 + b[P] + c = 0 \tag{40}$$

where,

$$a = 2[L]_t - [P]_t + K_{d,2}$$

$$b = K_{d,2}\left([L]_t - [P]_t + K_{d,1}\right)$$

$$c = -K_{d,1}K_{d,2}[P]_t$$

As shown in Eq. 26, the closed-form solution of Eq. 40 is given by:

$$[P] = -\frac{a}{3} + \frac{2}{3}\sqrt{a^2 - 3b}\cos\left(\frac{\theta}{3}\right) \tag{41}$$

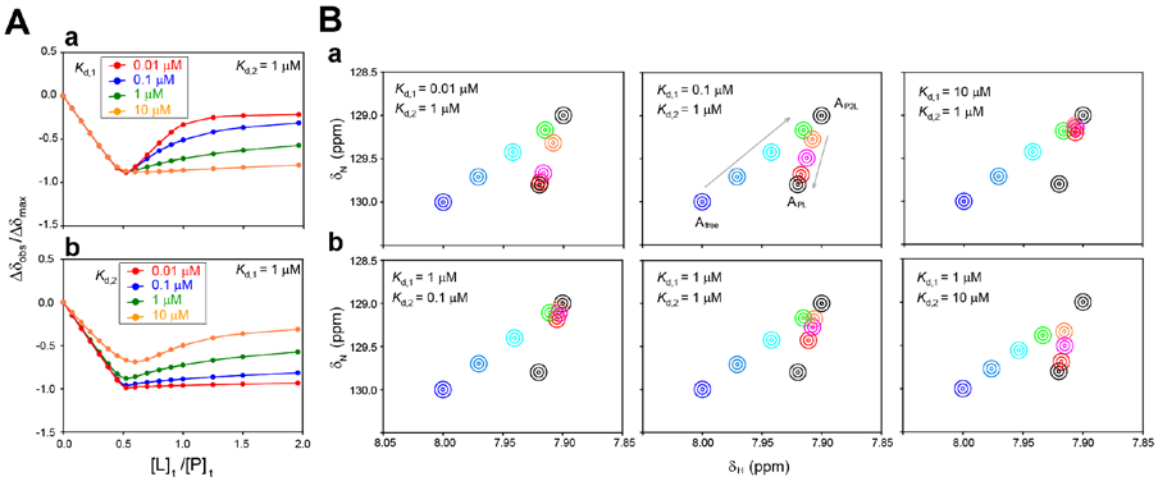

Figure 3. Two-site binding model: two proteins and one ligand. (A) The simulated titration curves for P with L as a function of $[L]_t/[P]_t$ ratio using Eq. 44 (a: $K_{d,2}$ = 1 μM and various $K_{d,1}$; b: $K_{d,1}$ = 1 μM and various $K_{d,2}$). The initial value of $[P]_t$ is 0.1 mM. (B) The calculated $^1H/^{15}N$-HSQC cross-peaks of P during titration with L using (a) $K_{d,2}$ of 1 μM and various $K_{d,1}$ values and (a) $K_{d,1}$ of 1 μM and various $K_{d,2}$ values. The cross-peak color changes gradually from blue to purple according to the $[L]_t/[P]_t$ ratio. The black cross-peak indicates the two bound states (PL and P_2L).

The observed chemical shift change is the weight average of the shifts in the free form (δ_{free}) and the PL (δ_{PL}) and P_2L (δ_{P2L}) complexes:

$$\delta_{obs} = \delta_{free}\frac{[P]}{[P]_t} + \delta_{PL}\frac{[PL]}{[P]_t} + \delta_{P2L}\frac{2[P_2L]}{[P]_t} \quad (42)$$

The change in the observed chemical shift from the free state ($\Delta\delta_{obs}$) is given by:

$$\Delta\delta_{obs} = \Delta\delta_{PL}\frac{[PL]}{[P]_t} + \Delta\delta_{P2L}\frac{2[P_2L]}{[P]_t} \quad (43)$$

where $\Delta\delta_{PL}$ and $\Delta\delta_{PL2}$ are the maximum chemical shift changes upon saturation by the PL and P_2L complexes, respectively. Using Eqs. 37 and 38, $\Delta\delta_{obs}$ can be described as:

$$\frac{\Delta\delta_{obs}}{\Delta\delta_{max}} = \frac{[[L]_t}{[P]_t}\frac{K_{d,2}[P]\frac{\Delta\delta_{PL}}{\Delta\delta_{max}} + 2[P]^2}{K_{d,1}K_{d,2} + K_{d,2}[P] + [P]^2} \quad (44)$$

where $\Delta\delta_{max} = \Delta\delta_{PL2}$ and [P] is expressed by Eq. 41. This equation allow us to fit $K_{d,1}$ and $K_{d,2}$ from the chemical shift perturbation data as a function of $[L]_t/[P]_t$ ratio (Fig. 3). When saturating amount of L is added, the observed amide signal moved toward the peak position for not the PL but P_2L states (Fig. 3), because the P_2L complex bigins to form rather than the PL complex. When $[L]_t > 0.5\times[P]_t$, the observed amide signals changed the direction of their movement toward the peak position for the PL state (Fig. 3).

In summary, CSP is a simple NMR technique to study binding of a protein to various ligands. In this review, we derived the equations to accurately analyze the CSP data using one-site binding model as well as two kinds of two-site binding model. This analysis might provide insight into the mechanism on how proteins selectively recognize their target ligands to achieve the biological function.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) Grants funded by the Korean Government (MSIP) [2017R1A2B2001832] and the Samsung Science and Technology Foundation [SSTF-BA1701-10]. We thank the GNU Central Instrument Facility for performing the NMR experiments.

References

1. M. P. Williamson, *Prog. Nuclear Magn. Reson. Spec*. **73,** 1 (2013)
2. E. R. P. Zuiderweg, *Biochemistry* **41**, 1 (2002)
3. M. A. McCoy and D. F. Wyss, *J. Am. Chem. Soc.* **124**, 11758 (2002)
4. M. Arai, J. C. Ferreon, and P. E. Wright, *J. Am. Chem. Soc.* **134**, 3792 (2012)
5. J. Granot, *J. Magn. Reson.* **55**, 216 (1983)
6. L. Fielding, *Prog. Nucl. Magn. Reson. Spectrosc.* **51**, 219 (2007)