JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

# Quinary Interactions Weaken the Electric Field Generated by Protein Side-Chain Charges in the Cell-like Environment

Ning Zhang,[†,‡,§,⊥] Liaoyuan An,[†,‡,⊥] Jingwen Li,[†,‡,§] Zhijun Liu,[∥] and Lishan Yao*[,†,‡]

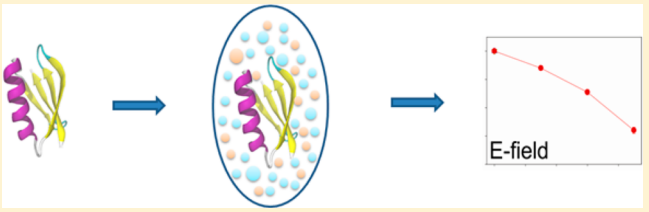
[†]Shandong Provincial Key Laboratory of Synthetic Biology and [‡]Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266061, China

[§]University of Chinese Academy of Sciences, Beijing 100049, China

[∥]National Center for Protein Science Shanghai, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 201210, China

Supporting Information

**ABSTRACT:** The intramolecular electric field (e-field) generated by protein GB3 side-chain charges K/E10, K/E19, and D/K40 was measured in the absence or presence of macromolecular crowding. The e-field responds differently to different crowding agents—dextran, Ficoll, BSA, and *E. coli* cell lysate. Dextran and Ficoll have no effect on the e-field. The lysate generally weakens the e-field but the amplitude of weakening varies greatly. For example, the e-field by K19 is reduced by 67% in the presence of 90 g/L lysate, corresponding to a charge change from 0.9 to 0.3 e for K19, whereas the e-fields by D/K40 are weakened only by ∼7% under the same lysate concentration. The extent of the e-field weakening by BSA is in between that by Ficoll (dextran) and lysate. Further investigations suggest that the e-field weakening mechanism by lysate is similar to that by NaCl. That is, the e-field generated by a protein surface charge affects the distribution of lysate which creates a reaction field and weakens the protein e-field. Our study indicates that the protein electrostatic property can be changed significantly due to quinary interaction with the cell environment.




## INTRODUCTION

The interior of a cell is crowded with macromolecules. For example, the macromolecular concentration in *Escherichia coli* can reach 300−400 g/L.[1] Although in-cell NMR studies suggest that the cell environment does not affect the overall structure of well-folded proteins,[2,3] the perturbation of the protein conformational exchange in cell lysate has been detected.[4,5] It has also been demonstrated that protein folding,[6−11] protein stability,[12−18] and protein ligand binding[19,20] can be affected by the cell environment. The change of protein properties in macromolecular crowding is attributed to the excluded volume effect[21−23] and quinary interactions, i.e., the weak interactions between the protein and the cell environment.[19,24−33] Although the excluded volume effect is well understood, the nature of quinary interactions remains elusive.

It was shown that the D40K mutation of GB1, the first immunoglobulin binding domain of protein G, significantly decreases the protein's stability in-cell but not in a buffered solution.[30] It was suggested that the generally negatively charged *E. coli* proteins are attracted to K40 (preferentially of the unfolded state), thereby destabilizing the protein.[30] A recent work demonstrates that biologically relevant crowders can interact particularly strongly with the unfolded ensemble.[31] The electrostatic nature of quinary interactions was also demonstrated in the unfolding measurement of SH3 in solutions of bovine serum albumin (BSA) and lysozyme.[31] The thermostability change of SH3 due to crowding of BSA or lysozyme varies with pH, suggesting charge−charge interactions between SH3 and the crowder. The thermostability measurement in crowders provides important information about the quinary interaction with the protein. However, quinary interactions with the folded and unfolded states both can contribute to protein stability. Separating the two from the thermostability measurement is impossible. How quinary interactions affect electrostatics of well-folded proteins is largely unknown.

In this work, we focus on the macromolecular crowding effect on the protein intramolecular electric field (e-field) by using the third immunoglobulin binding domain of protein G (56 residues), or GB3,[34−37] a well-folded protein as a model system. The electric field (e-field) is monitored by the chemical shift perturbation (CSP) method[38−40] which is briefly described as follows. Taking a lysine residue (K) in GB3 as an example, a mutation K→A is introduced to eliminate the e-field created by the positive charge of K, wherein the chemical shift change is recorded (CSP, $\Delta\delta = \delta_K - \delta_A$) for each backbone amide $^1$H whose chemical shift is sensitive to the e-field.[41] For the amides near the lysine residue, the CSP has the contribution from the e-field of the lysine, as well as the structural change[41] caused by the mutation. For the amides far

**Received:** May 25, 2016
**Revised:** October 23, 2016
**Published:** January 6, 2017





DOI: 10.1021/jacs.6b11058
J. Am. Chem. Soc. 2017, 139, 647−654

away, the contribution from the structural change diminishes, whereas the e-field effect persists because of the long-range nature of electrostatics. The CSPs are proportional to the e-field through the Buckingham equation.[42] As a control experiment, a K→E mutation, which doubles the e-field, is made where a 2-fold increase of CSPs is expected for the remote amides provided that these CSPs are solely caused by the e-field. Using the CSP value to quantify e-field at each amide is desired but practically difficult because e-field is a vector. But CSP values from all the remote amides together provide a sensitive way to monitor the e-field change generated by the lysine side chain due to its interaction with environment (e.g., ions or cell lysate).

The intramolecular e-field was monitored in the absence or presence of different macromolecular crowders, including ficoll 400 (Ficoll), dextran 500 (dextran), bovine serum albumin (BSA), and *E. coli* lysate. The results show that the e-field is generally weakened in the presence of lysate but not Ficoll or dextran and the extent of weakening varies greatly from one charged side chain to another. Further analysis suggests that the transient quinary interaction with lysate or BSA weakens the e-field in a similar way as NaCl, but with much higher specificity. Direct quinary interactions between charged side chains and lysate are also probed by the backbone amide $^{15}$N relaxation rates.

## METHODS AND MATERIALS

**Sample Preparation.** GB3 and its mutants K10A, K10E, K19A, K19E, D40A, and D40K were made by expression in *E. coli* BL21 (DE3*) cells, transformed with a pET-11 vector containing the GB3 or mutant gene. Details of the preparation and purification procedure have been described previously.[43] Mixed $^{15}$N-labeled and $^{15}$N/$^{13}$C-labeled protein samples were prepared, containing 10 mM sodium phosphate, pH 6.4, 5% $D_2O$, in 500 $\mu$L volume. Dextran, Ficoll and BSA were purchased from Solarbio (Beijing, China). BSA was dissolved, dialyzed against the distilled and deionized (DD) water and lyophilized before usage. *E. coli* cell lysate was prepared in the same way as described in the literature.[4] Briefly, *E. coli* BL21 (DE3*) cells were grown in LB media at 37 °C until $OD_{600} \approx 1.0$. The cells were harvested, sonicated, and centrifuged to remove the insoluble material. The lysate was dialyzed against the DD water using a dialysis membrane with a molecular weight cutoff of 3.5 kDa and then lyophilized for further usages. As the buffer, 10 mM sodium phosphate at a pH of 6.4 was selected because addition of 90 g/L lysate or BSA essentially maintains the pH value (decreases the pH value by ~0.05), which eliminates any pH effect on CSP values.

**NMR Measurements of CSPs.** All NMR experiments were carried out at 298 K on a Bruker Avance 600 MHz spectrometer, equipped with a *z*-axis gradient, triple-resonance, cryogenic probe. 2D constant-time $^{15}$N−$^{1}$H HSQC[40] interleaved spectra were recorded for six mixed-labeled samples, including K10($^{15}$N/$^{13}$C)/K10A($^{15}$N), K10($^{15}$N/$^{13}$C)/K10E($^{15}$N), K19($^{15}$N/$^{13}$C)/K19A($^{15}$N), K19($^{15}$N/$^{13}$C)/K19E($^{15}$N), D40($^{15}$N/$^{13}$C)/D10A($^{15}$N), and D40($^{15}$N/$^{13}$C)/D40K($^{15}$N), which permit us to extract the chemical shifts for both proteins with different isotope labeling in the same NMR tube. For both $^{15}$N/$^{13}$C- and $^{15}$N-labeled proteins, the concentration is ~0.5 mM. The backbone amide $^{1}$H chemical shift difference, e.g., $\Delta\delta_{A10K} = \delta_{K10(15N/13C)} - \delta_{A10(15N)} - (\delta_{WT(15N/13C)} - \delta_{WT(15N)})$, corresponds to the CSP caused by the mutation (A10K). The term in the parentheses is a small isotope effect correction (~1−2 ppb) due to the different isotope labeling of the two proteins.[40] This term was measured only for the wild-type (WT) GB3 and used for corrections of all the CSPs. Acquisition times were 63 ($^{15}$N) and 83 ms ($^{1}$H) with the data matrices consisting of 127*×1024* data points, where *N** indicates *N* complex points. In the titration experiments, a small amount of dextran, Ficoll, BSA or lysate powder was added to the sample to reach final crowder concentrations of 0, 30, 60, and 90 g/L. The signal-to-noise ratio decreases considerably for certain protein samples at the highest lysate concentration (90 g/L). Thus, the titration was stopped at 90 g/L of macromolecular crowders to permit the accurate measurement of CSPs. The data were processed and analyzed using the NMRPipe software.[44]

**Backbone Amide $^{15}$N Relaxation.** $^{15}$N relaxation rates, including longitudinal ($R_1$) and spin-locked transverse ($R_{1\rho}$), were measured at 600 MHz $^{1}$H frequency, for the six mixed-labeled samples listed above. The mixed protein samples have a concentration of 0.8 mM for both $^{15}$N/$^{13}$C- and $^{15}$N-labeled proteins, dissolved in 10 mM sodium phosphate (pH 6.4). An isotope filtering pulse sequence for $R_1$ and $R_{1\rho}$ measurements was used (Figure S1) which separated signals from the two proteins.[45] For the $R_{1\rho}$ measurement,[46] spin-lock durations were set to 4 and 60 ms. A $^{15}$N spin-lock field strength of 2.8 kHz was used with the carrier frequency positioned in the center of the $^{15}$N spectrum. The acquisition times were 36.6 ms ($^{15}$N) and 83 ms ($^{1}$H) with the data matrices consisting of 110*×1024* points. A 2-s delay was applied between scans, and $^{15}$N RF heating compensation pulses were utilized to ensure the temperature of the experiment to remain constant.[47] For the $R_1$ measurement, the relaxation delays were 100 and 600 ms. The acquisition times and the delay between scans were the same as in the $R_{1\rho}$ measurement. The $^{15}$N relaxation rates were derived from the intensity ratio of two spectra, for the mixed samples in the presence of 0, 60, and 90 g/L of Ficoll, BSA, or lysate. The relaxation rates for 48 out of 56 residues that have well-separated signals in all spectra were obtained (excluding M1, Q2, Y3, E15, A20, T25, E27, and N35). The error of the measured relaxation rates was propagated from the signal-to-noise ratio of the spectra.

**$R_1R_2$ Prediction.** The diffusion tensors of GB3 were fitted using the $R_2/R_1$ ratios obtained from the relaxation rates of mixed samples, with the NMR-refined GB3 structure (PDB entry 2OED).[48] An asymmetric diffusion tensor was fitted, as previously found for this protein.[49] Different $^{15}$N CSA tensors were utilized in the fitting, with ($\sigma_{ZZ}$, $\sigma_{XX}$, $\sigma_{YY}$) values of (−115, 71, 44 ppm) for residues in the $\alpha$-helix, (−108, 71, 37 ppm) for residues in the $\beta$-strands, and (−111, 69, 42 ppm) for the rest of the protein.[45] An effective N−H bond length of 1.04 Å was used for the dipole−dipole spectra density calculation.[50] After the diffusion tensors were obtained, the $R_1R_2$ values were predicted using a uniform order parameter 0.9[45] and a local motion time of 0 ps.

## RESULTS AND DISCUSSION

**Long-Range Backbone Amide $^{1}$H CSPs and Charges.** The structure of GB3 has a prolate shape with a length of 27 Å and an aspect ratio of 1.7.[48] Three mutational sites K10, K19, and D40 were selected, all of which are located close to the ends of the long axis of GB3 (Figure 1A). K19, located in the end of $\beta$2 (K13−K19), has its side chain inserted into solvent and does not form any salt bridge or H-bond with other residues.[48] The nearest charge is from the NH3 of N-terminal residue M1, with the Nζ(K19)−N(M1) distance of 6.7 Å (Figure S2, the second nearest charged residue is K4 which is ~13 Å away). These two positive charges create a positive electrostatic potential in the surroundings (Figure S2). The side chain of K10 is near that of D40, with the Nζ (K10)−Cγ(D40) distance of 8.2 Å. In between these two charged side chains, there are two carboxyl groups from the backbone and side chain of the C-terminal residue E56 (Figure S2). The distances from Nζ of K10 to the Cα atom of the E56 backbone and the side-chain carboxyl carbon atom are 6.6 and 5.7 Å, whereas the distances from the same two carbon atoms to Cγ of D40 are 8.2 and 6.6 Å, respectively. Although, the distances between the three carboxyl groups and Nζ of K10 are a bit too large for forming stable salt bridges, the presence of excessive negative charges creates a negative electrostatic potential in the surroundings (Figure S2). Mixed protein samples, including $^{15}$N/$^{13}$C-labeled K10/$^{15}$N-labeled K10A (same below), K10/


DOI: 10.1021/jacs.6b11058
J. Am. Chem. Soc. 2017, 139, 647−654

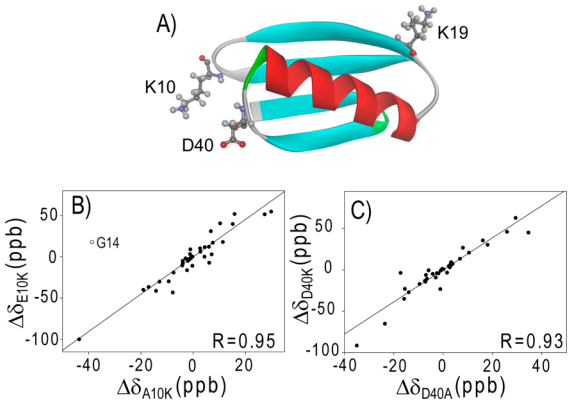



**Figure 1.** (A) Locations of the three mutated charged residues of GB3, K10, K19 and D40. (B,C) Correlation between remote backbone amide $^1$H CSP values, $\Delta\delta_{A10K}$ and $\Delta\delta_{E10K}$; $\Delta\delta_{D40A}$ and $\Delta\delta_{D40K}$. Only the CSPs of the residues with a backbone amide nitrogen >12 Å away from the C$\gamma$ of the mutated site (K10 or D40) were included. The best fitted line is $y = 2.25x$ for panel B (excluding the outlier G14 which is 12.05 Å away from K10) and $y = 1.94x$ for panel C. The correlation of $\Delta\delta_{E19K}$ and $\Delta\delta_{A19K}$ is shown in Figure S3, with a slope of 2.12. The CSPs were measured in 10 mM sodium phosphate (pH = 6.4). The error of CSP is ~0.1 ppb.

K10E, K19/K19A, K19/K19E, D40/D40A, and D40/D40K were used for the high accuracy backbone amide $^1$H CSP measurement.[40] The mutational effect on the protein structure is expected to be small, based on the similarity of their $^1$H–$^{15}$N spectra. Similar to K19 (Figure S3), a previously identified site,[40] the remote backbone amide $^1$H CSPs (with a distance $d$ between its backbone amide nitrogen and C$\gamma$ of the mutated site >12 Å) observed for the K10E ($\Delta\delta_{E10K}$) and K10A ($\Delta\delta_{A10K}$) mutants are well correlated (Figure 1B). A similar correlation is seen for $\Delta\delta_{D40K}$ and $\Delta\delta_{D40A}$, with a slope of 1.94 (Figure 1C). In the $^1$H CSP correlation plot, a total of 38, 31, and 37 remote $^1$H CSPs were included for sites 10, 19, and 40, respectively. To see whether such a correlation persists for long $d$ values, the backbone amide $^1$H CSPs were grouped to three clusters with the distance $d$ in the range of 12 Å < $d$ < 17 Å, 17 Å < $d$ < 22 Å, and $d$ > 22 Å, respectively. A good correlation was observed for the $^1$H CSPs in all three clusters, with generally smaller CSPs for longer $d$ values and a slightly varied slope probably due to the smaller amount (~10) of CSPs in each cluster (Figure 2), confirming that the e-field effect persists over the whole distance range. These positive correlations also indicate that the structural perturbation caused by the mutations in the remote sites is small. Assuming that the effective charges are −1 e for E10, E19, and D40, the corresponding charges are 0.8, 0.9, and 0.9 e for K10, K19, and K40 (Table 1), respectively, based on the slope of the best fitted lines (Figures 1 and S3). Up to 200 mM NaCl was titrated into the mixed protein samples. As can be seen, the backbone amide $^1$H CSPs are generally scaled down (scaling factor <1) as a result of electrostatic screening by the salt (Figure 3). The slightly different scaling factors are likely due to ion accessibility and/or affinity differences at the three sites. For example, for $\Delta\delta_{A10K}$ which has the highest CSP scaling factor (Figure 3A), the negative electrostatic potential around K10 may limit the $Cl^-$ accessibility to the side chain of K10. Thus, adding NaCl has no effect on the e-field created by K10.

**Macromolecular Crowding Effect on Long-Range $^1$H CSPs.** The powder of dextran, Ficoll, BSA, or *E. coli* lysate was

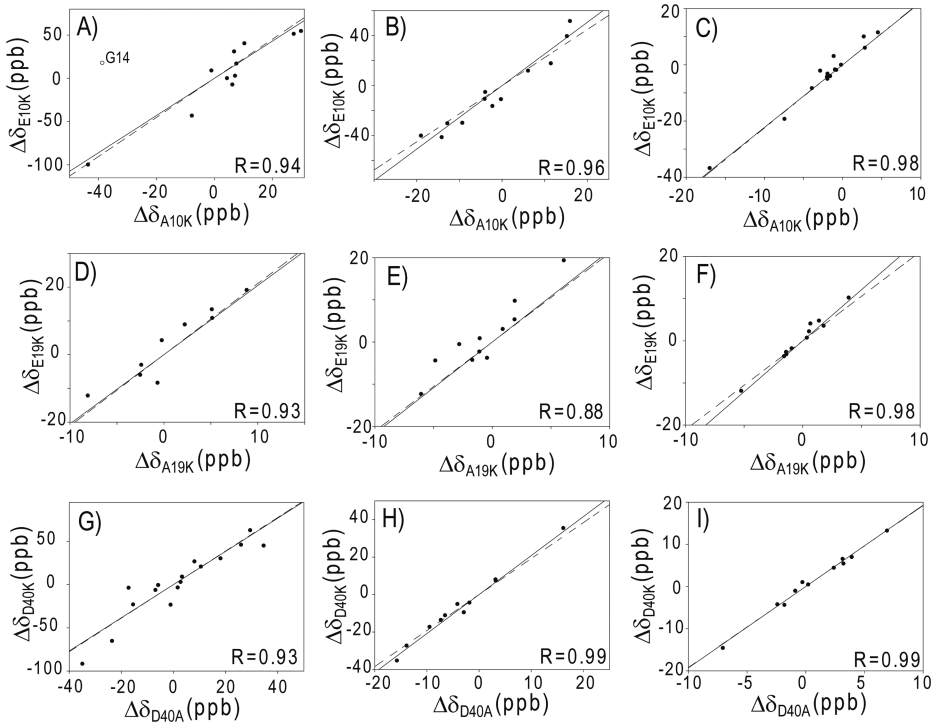



**Figure 2.** Correlation between backbone amide $^1$H CSP values of (A–C) $\Delta\delta_{A10K}$ and $\Delta\delta_{E10K}$, (D–F) $\Delta\delta_{A19K}$ and $\Delta\delta_{E19K}$, and (G–I) $\Delta\delta_{D40A}$ and $\Delta\delta_{D40K}$ with the distance $d$ to the mutated charge in different ranges, 12 Å < $d$ < 17 Å (A, D, and G), 17 Å < $d$ < 22 Å (B, E, and H), and $d$ > 22 Å (C, F, and I). The dashed line is the best-fitted line from the fitting of all the CSPs (Figures 1 and S3), whereas the solid line is from the fitting of the CSP values in the panel only. The persistence of the correlation at different distance ranges indicates that the CSPs are caused by the e-field of the mutated charge. The error of CSP is ~0.1 ppb.



DOI: 10.1021/jacs.6b11058
J. Am. Chem. Soc. 2017, 139, 647–654

**Table 1. Effective Charges of Different Side Chains in the Absence or Presence of 90 g/L Ficoll, Dextran, BSA, or *E. coli* Cell Lysate**

| | charge (e) | | | | |
|---|---|---|---|---|---|
| residue | buffer[a] | Ficoll | dextran | BSA | lysate |
| E10 | −1.0[b] | −1.0 ± 0.1 | −1.0 ± 0.1 | −0.6 ± 0.1 | −0.4 ± 0.1 |
| K10 | 0.8 ± 0.1 | 0.8 ± 0.1 | 0.8 ± 0.1 | 0.7 ± 0.1 | 0.7 ± 0.1 |
| E19 | −1.0[b] | −1.0 ± 0.2 | −1.0 ± 0.2 | −0.9 ± 0.2 | −0.5 ± 0.1 |
| K19 | 0.9 ± 0.1 | 0.9 ± 0.1 | 0.9 ± 0.1 | 0.8 ± 0.1 | 0.3 ± 0.1 |
| D40 | −1.0[b] | −1.0 ± 0.2 | −1.0 ± 0.2 | −1.0 ± 0.2 | −0.9 ± 0.2 |
| K40 | 0.9 ± 0.1 | 0.9 ± 0.1 | 0.9 ± 0.1 | 0.9 ± 0.1 | 0.9 ± 0.1 |

[a]10 mM sodium phosphate (pH 6.4). [b]The charges of E10, E19, and D40 in the buffer are set to −1.0 e and thus have no error.

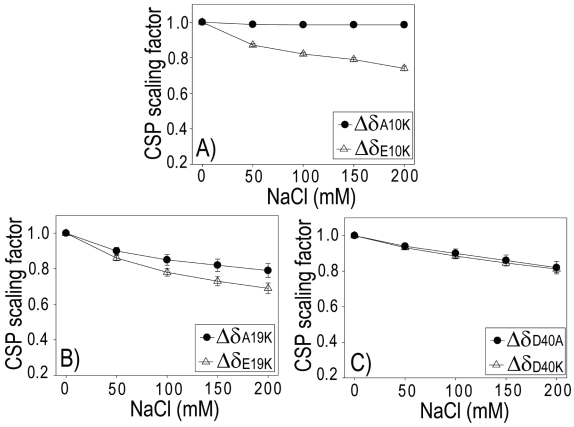



**Figure 3.** CSP scaling factor as a function of the NaCl concentration. The CSP values at different NaCl concentrations were correlated to those in 10 mM sodium phosphate alone and the slope of the best fitted line was defined as the CSP scaling factor. All the scaling factors are less than 1, due to electrostatic screening by NaCl. The variation of the scaling factor for CSPs originated from different charges at the same NaCl concentration suggests that the ion accessibility and affinity to the charges are different.

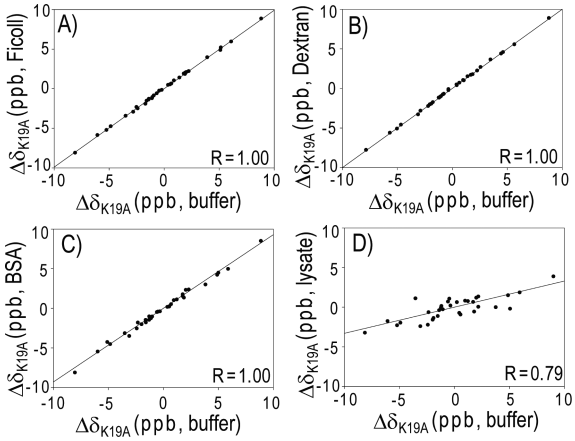



**Figure 4.** Correlation between $\Delta\delta_{K19A}$ values in the absence and presence of 90 g/L macromolecular crowding agents, (A) Ficoll, (B) dextran, (C) BSA, or (D) *E. coli* cell lysate. The best-fitted lines are $y = 0.99x$ (A), $y = 1.00x$ (B), $y = 0.93x$ (C), and $y = 0.33x$ (D). $\Delta\delta_{K19A}$ on the $x$-axis corresponds to the measurement in 10 mM sodium phosphate. The slope corresponds to the CSP scaling factor. The much smaller CSP scaling factor in panel D indicates that the side-chain charge of K19 is strongly reduced in lysate (Table 1). The error of CSP is 0.1 ppb in the buffer, Ficoll, and BSA, and 1.1 ppb in lysate.

titrated to the mixed protein samples, where the corresponding CSP changes were recorded. As can be seen, dextran and Ficoll do not change CSPs (Figures 4A,B and 5), indicating that the e-field is not detectably affected by the two crowders. Therefore, the six side-chain effective charges do not change in the presence of dextran or Ficoll (Table 1). By contrast, BSA scales down the CSPs (Figures 4C and 5). In the presence of 90 g/L BSA, the CSPs are reduced by ∼4−6% for $\Delta\delta_{A10K}$, $\Delta\delta_{D40A}$, and $\Delta\delta_{D40K}$, and ∼7−8% for $\Delta\delta_{A19K}$ and $\Delta\delta_{E19K}$, whereas for $\Delta\delta_{E10K}$ the CSPs are reduced by ∼26% (Figure 5). Since the change of $\Delta\delta_{A10K}$ is relatively small, the $\Delta\delta_{E10K}$ reduction mainly arises from loss of the effective charge of E10 (Table 1).

The titration of lysate into the samples reduces CSPs much stronger than BSA (Figures 4D and 5). In the presence of 90 g/L lysate, $\Delta\delta_{A19K}$ is reduced by 67% (Figure 4D), indicating that the effective charge of K19 decreases to $(1 - 0.67) \times 0.9$ e = 0.3 e (Table 1). Under the same condition, $\Delta\delta_{E19K}$ is reduced by 56%, corresponding to an effective charge change from −1 e to $(1 - 0.56) \times (1 + 0.9)$ e − 0.3 e = −0.5 e for E19 (Table 1). For residue 10, $\Delta\delta_{A10K}$ and $\Delta\delta_{E10K}$ are reduced by 16% and 43%, respectively, in 90 g/L lysate. The corresponding charges are 0.7 e for K10 and −0.4 e for E10, indicating that E10 is neutralized more than K10 by lysate, opposite to what is seen for residue 19. In contrast, for residue 40, the CSPs are reduced only slightly, similar to what is seen for the BSA titration (Figure 5E,F). In other words, the effective charge change is small for D40 and K40 in the presence of lysate (Table 1).

**Lysate Effect on Site-Specific $^1$H/$^{15}$N CSPs.** Chemical shifts are sensitive to local environmental change. If macromolecules in lysate bind to a specific region of GB3, chemical shifts of GB3 residues in the binding region should change accordingly. The backbone amide $^1$H and $^{15}$N chemical shift coefficient $d\delta/d$[lysate] was determined through fitting the chemical shift to the lysate concentration for the WT GB3 (Figure 6). As can be seen, the overall coefficient is small, with an average of −0.12 ± 0.06 and −0.12 ± 0.49 ppb·L/g for $^1$H and $^{15}$N, respectively, which corresponds to an average change of 10.8 ppb at the lysate concentration of 90 g/L, suggesting that the GB3 overall structure change is very small. The change of $^1$H chemical shift is comparable to the CSPs caused by mutation (Figure 1 and S3). From the WT GB3 $^1$H and $^{15}$N chemical shift coefficients, it is difficult to tell which residues bind to lysate (Figure 6A,B). This coefficient pattern is similar to that of the encounter protein complexes where many residues display comparable chemical shift changes.[51,52] The coefficient $d\Delta\delta/d$[lysate] was determined in a similar way for $^1$H and $^{15}$N of mutants where the CSP between the WT and the charge-changing mutant was used in the linear fit, so that we can focus on the role of a specific charged residue on the GB3−lysate transient interaction. $d\Delta\delta/d$[lysate] reflects the


DOI: 10.1021/jacs.6b11058
J. Am. Chem. Soc. 2017, 139, 647−654

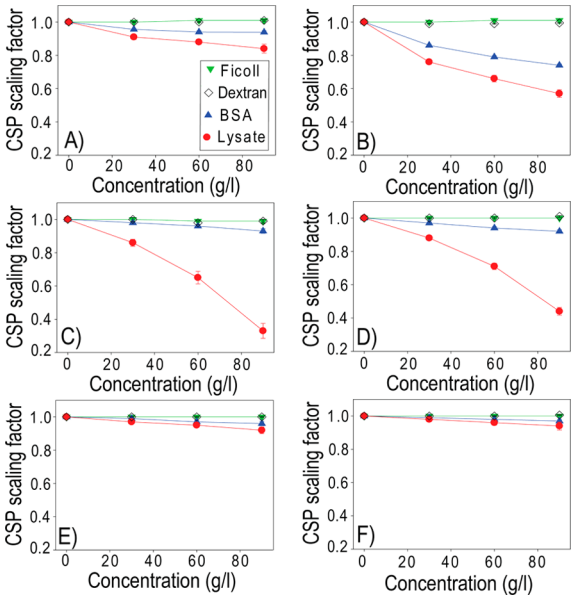



**Figure 5.** Macromolecular crowding effect on the CSP scaling factor of (A) $\Delta\delta_{A10K}$, (B) $\Delta\delta_{E10K}$, (C) $\Delta\delta_{A19K}$, (D) $\Delta\delta_{E19K}$, (E) $\Delta\delta_{D40A}$, and (F) $\Delta\delta_{D40K}$. The decrease of the CSP scaling factor in the presence of BSA or lysate reflects weakening of the e-field generated by the side-chain charge as a result of quinary interactions.

mutational effect on lysate distribution around the protein surface which affects the CSP. $d\Delta\delta/d[lysate]$ has two possible origins: (1) the lysate redistribution causes the transient GB3−lysate contact change, and (2) the lysate redistribution creates a reaction field which weakens the e-field generated by the mutational charge. The first origin, a direct effect, is larger for residues close to the mutational site, whereas the second origin, an indirect effect, becomes more important for residues far away from the mutational site. For K10A and K10E, large $d\Delta\delta/d[lysate]$ coefficients of residues 10, 11, 12, and 40 (near K10) were observed, suggesting that the mutation may change the GB3−lysate binding affinity or the binding mode (Figure 6C,D). For K19A and K19E, the chemical shift coefficients are much smaller. But residues near K19, such as L4, V5, A20, A23, D47−T51, have relatively larger coefficients (Figure 6E,F). For the D40A and D40K mutations, L12, D40, V42, and W43, close to the mutated charge, also have larger coefficients (Figure 6G,H). The relatively larger $d\Delta\delta/d[lysate]$ for the sites close to the mutational site suggests that the direct contact effect is generally larger than the indirect e-field weakening effect. But for K19 and D40, the difference between the two effects is smaller. In other words, it becomes less obvious whether these two charged residues involve direct interactions with lysate based solely on the $d\Delta\delta/d[lysate]$ values. However, one cannot exclude the possible direct interaction between these two sites and lysate either. One explanation is that more than one protein can bind transiently to K19 (or D40) and the binding perturbs the GB3 amide chemical shifts in different directions which on average cancel out each other. Thus, the $d\Delta\delta/d[lysate]$ values for residues near K19 (or D40) appear to be small, although the direct interaction exists.

**Crowding Effect on GB3 $^{15}$N Relaxation Rates.** Backbone amide $^{15}$N $R_1$ and $R_2$ relaxation rates were measured for the WT GB3 and all six mutants in the absence or presence of Ficoll, BSA, or lysate. The increase of sample viscosity by crowder is expected to increase the tumbling time $\tau_c$ which decreases $R_1$ but increases $R_2$. The product $R_1R_2$ is not sensitive to $\tau_c$ of the protein (if $\tau_c \gtrsim 5$ ns) provided that there is no

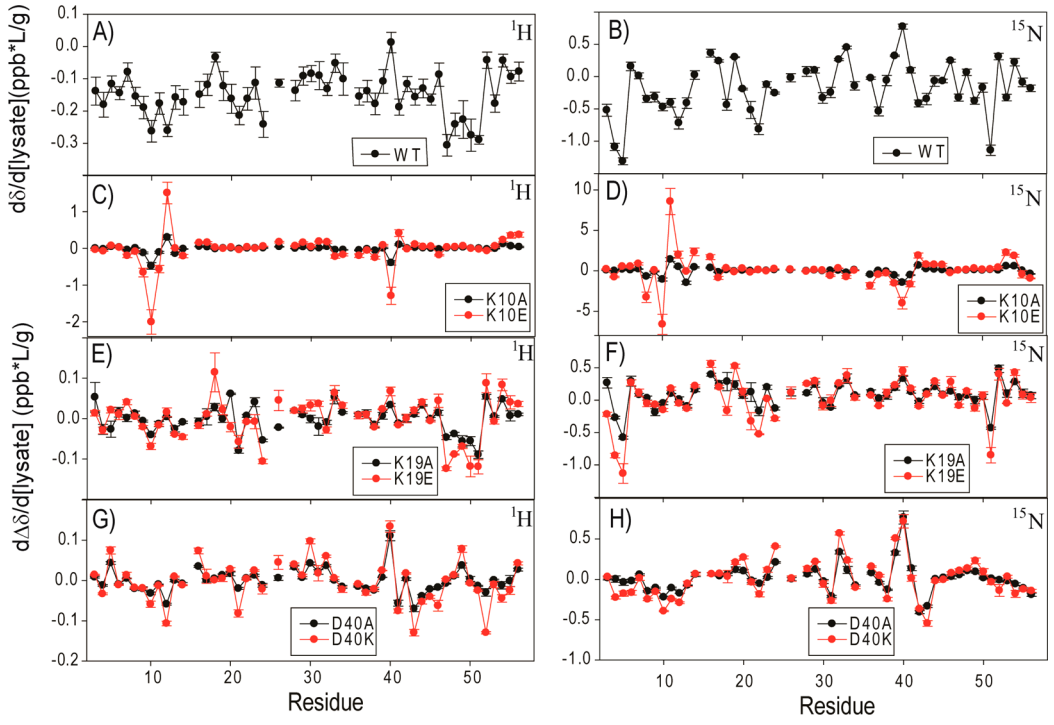



**Figure 6.** (A,B) Site-specific $d\delta/d[lysate]$, the linear fitting slope of $^1$H/$^{15}$N chemical shifts of the WT GB3 versus the lysate concentration. (C−F) Site-specific $d\Delta\delta/d[lysate]$, the linear fitting slope of $^1$H/$^{15}$N CSPs versus the lysate concentration. The lysate interaction with residue 10 and its vicinity is obvious, as suggested by the larger $d\Delta\delta/d[lysate]$ values for the residues nearby (C,D). In comparison, the lysate interaction with residues 19 and 40 is less obvious (E−H).



DOI: 10.1021/jacs.6b11058
J. Am. Chem. Soc. 2017, 139, 647−654

chemical exchange.[53] The contribution of chemical exchange to N15 $R_2$ relaxation rates of protein GB3 is very small.[49] Quinary interaction with proteins in lysate tends to dramatically increase $R_2$ because their sizes are generally much larger than GB3, whereas the effect on $R_1$ is smaller. Thus, the product $R_1R_2$ can be used as an indicator for the quinary interaction.[54] For protein GB3, due to its fast tumbling time $\tau_c$ of ∼3 ns,[49] increase of viscosity also increases $R_1R_2$ slightly. To eliminate this factor, we fitted the diffusion tensor of the WT GB3 using the $R_2/R_1$ ratio and then predicted the average $R_1R_2$. As shown in Figure S4, the predicted $R_1R_2$ agreed well with the experimental value in the buffer. Addition of Ficoll increases the tumbling time to 3.7 ns (60 g/L Ficoll) and 4.1 ns (90 g/L Ficoll). At these two Ficoll concentrations, the predicted and experimental $R_1R_2$ are still in good agreement (Figure S4). In comparison, the experimental $R_1R_2$ are 1.1 and 2.1 $s^{-2}$ higher than the predicted values in 60 g/L and 90 g/L BSA, respectively. In the presence of lysate, the differences between the experimental and predicted $R_1R_2$ are 20.7 $s^{-2}$ and 30.0 $s^{-2}$ in 60 g/L and 90 g/L lysate, respectively (Figures S4 and 7A,B). These results suggest that transient interactions with GB3 occur in BSA and lysate, but not in Ficoll.

The difference $\Delta R_1R_2$ ($\Delta R_1R_2 = R_1R_2(\text{mut}) - R_1R_2(\text{WT})$) between the WT and a mutant provides the information whether the mutated residue directly involves in the interaction with crowders. The effect of Ficoll concentration on $\Delta R_1R_2$ of all the mutants is very small, suggesting that they do not interact with Ficoll, consistent with the $R_1R_2$ analysis of the WT (Figure 7). $\Delta R_1R_2$ of K10E, K19A, and K19E change with the increase of BSA concentration. Specifically, K10E has an increased $\Delta R_1R_2$, whereas K19A and K19E have decreased $\Delta R_1R_2$. The increased $\Delta R_1R_2$ of K10E suggests that the mutant K10E has a higher affinity with BSA. For K19A and K19E, the affinity is lower. In the presence of lysate, all the mutants have decreased $\Delta R_1R_2$, suggesting that they all affect the quinary interaction with lysate. The largest $\Delta R_1R_2$ decrease is from K19A/E, with a value ∼16−22 $s^{-2}$ in 60 and 90 g/L lysate, which corroborates with the largest CSP scaling factor decrease of this residue.

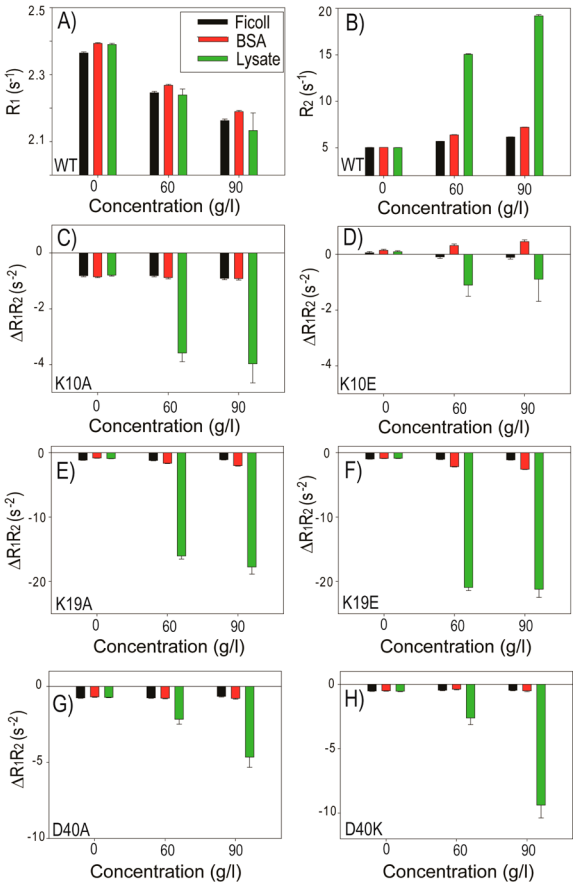



**Figure 7.** Relaxation rates $R_1$ and $R_2$ of the WT GB3 (A and B) and rate differences $\Delta R_1R_2 = R_1R_2(\text{mut}) - R_1R_2(\text{WT})$ of six mutants (C−H) in the absence or presence crowding agents, Ficoll, BSA or lysate. The dramatic increase of $R_2$ in lysate indicates that the WT GB3 interacts with this crowder (B). The large decrease of $\Delta R_1R_2$ in lysate for the mutants suggests that all six charges involve in the quinary interaction.

**E-Field Weakening Mechanisms in Lysate.** According to the Debye−Huckel theory, the electrostatic potential created by a point charge $q$ in a solution can be written as $q\exp(-r/\lambda)/\varepsilon r$, where $\lambda$ is the Debye screening length, $\varepsilon$ is the dielectric constant, and $r$ is the distance between the charge and the location of the potential. For a protein side-chain charge, its electrostatic potential is more complicated, but one can think of the protein as part of the solvent (with a larger radius and a smaller dielectric constant than water). One mechanism of the e-field weakening is from the $\varepsilon$ change, as the lysate concentration increases. A recent computational study shows that in the presence of 10% protein crowder, the solvent water dielectric decreases by ∼10% due to restraining of water molecules by the protein crowder.[55] The decrease of solvent dielectric constant is expected to increase the side-chain e-field, which is opposite to the experimental observation. Another mechanism of the e-field weakening is from the ionic strength increase which decreases the $\lambda$ value. The $^1$H 90° pulse length in NMR is sensitive to the ionic strength of solution and thus can be used as an ionic strength indicator. Addition of 90 mM NaCl to DD water increases the $^1$H 90° pulse length (at the power of 5.75 W) from 8 to 11 $\mu$s (Figure S5). In comparison, an addition of 90 g/L lysate increases the $^1$H 90° pulse length to 8.55 $\mu$s which corresponds to the NaCl concentration of ∼12 mM. This ionic strength increase is estimated to decrease the e-field by ∼5% (Figure 3), which is comparable to the e-field weakening of the D40/K40 charges in 90 g/L lysate solution (Figure 5E,F). In other words, the D40/K40 e-field weakening in lysate is likely caused by the ionic strength change. But for sites K10/E10 and K19/E19, the ionic strength effect is too small to account for the e-field weakening observed experimentally.

The third mechanism is that the transient quinary interaction between lysate and the charge of the protein decreases the long-range e-field created by a charge (Figure 8). For example, introducing a positive charge on the protein surface through mutation creates a positive electrostatic potential in space and causes the redistribution of lysate on the protein surface. This redistribution creates a negative electrostatic potential (a reaction field generated by charges and dipoles from lysate to form a favorable interaction with the positive protein charge) which effectively neutralizes the protein charge and thus weakens its long-range potential and e-field. The amplitude of e-field weakening depends on the amplitude of the reaction field. This mechanism explains the observation that the e-field generated by both positive and negative charges at the same site is weakened in the presence of lysate (Figure 5 and Table 1). This mechanism is supported by the positive correlation


DOI: 10.1021/jacs.6b11058
J. Am. Chem. Soc. 2017, 139, 647−654

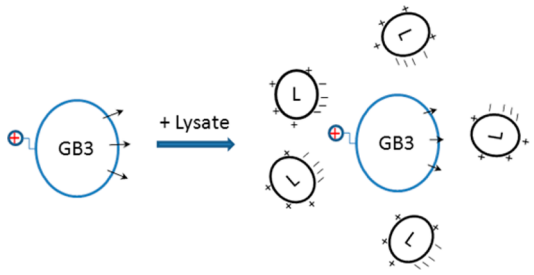



**Figure 8.** Mechanisms for the e-field (denoted by black arrows in the protein) weakening due to the presence of lysate (abbreviated as "L" in the figure). The positive GB3 charge attracts the negative surface charges of macromolecules in lysate which partially neutralizes the positive charge and thus decreases its e-field. (The same mechanism applies for the negative GB3 charge.) Lysate also slightly increases the ionic strength of the solution (Figure S5) which slightly weakens the e-field as well. Protein molecules (in lysate) also slightly decrease the dielectric constant of the solvent due to restraining of water molecules.[55] The decrease of dielectric constant is expected to increase the e-field, opposite to the experimental findings. See more discussion in the main text.

between $d\Delta\delta/d[lysate]$ and $d\Delta\delta/d[NaCl]$ (Figure 9), suggesting that the interaction governs the redistribution of

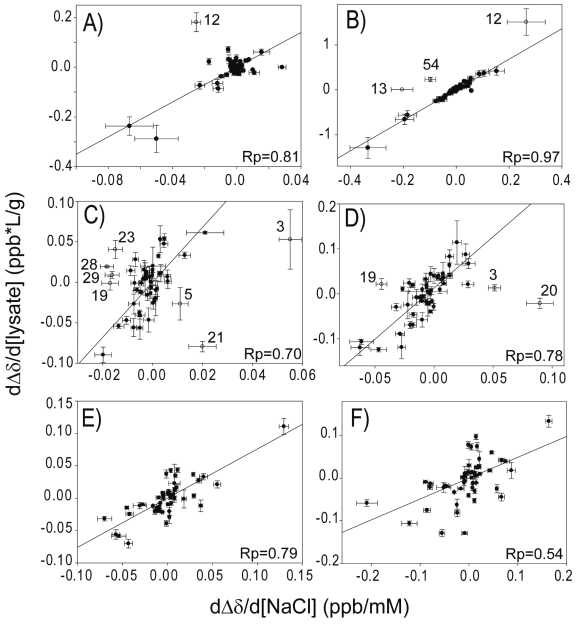



**Figure 9.** Correlation between $^1$H $d\Delta\delta/d[lysate]$ and $d\Delta\delta/d[NaCl]$ for $\Delta\delta_{A10K}$ (A), $\Delta\delta_{E10K}$ (B), $\Delta\delta_{A19K}$ (C), $\Delta\delta_{E19K}$ (D), $\Delta\delta_{D40A}$ (E), and $\Delta\delta_{D40K}$ (F). The slopes of the best fitted lines (excluding outliers) are 3.5 ± 0.4 (A), 3.4 ± 0.1 (B), 3.4 ± 0.5 (C), 2.3 ± 0.3 (D), 0.76 ± 0.08 (E), and 0.53 ± 0.12 (F) with the unit of mM(NaCl)/g/L(lysate). The outliers are usually close to the mutated charges. The positive correlation of the $\Delta\delta$ coefficients implies that the e-field weakening mechanism by lysate is similar to that by NaCl.

lysate is electrostatics, similar to that of NaCl. This mechanism also suggests that the direct and stable salt bridge between the protein charge and the lysate charge is not necessary for the e-field weakening. Long range electrostatics, which can accommodate more protein−lysate transient quinary interaction modes, may contribute more to e-field weakening. Thus, the $pK_a$ changes of K, D, and E caused by lysate, which cannot be measured experimentally because lysate becomes unstable when pH < 5 or pH > 11, may not be that significant. Several residues that deviate from the correlation (labeled in the correlation plots) are close to the mutational site (Figure 9). This is not surprising since the macromolecules in lysate are larger and more complex than NaCl ions and their redistribution can have a more complex effect on residues near the charge (the direct effect as discussed above). The slopes of $d\Delta\delta/d[NaCl]$ versus $d\Delta\delta/d[lysate]$ plots vary from 0.8 (0.5) of D40A (D40K) to 3.4 (2.3) of K19A (K19E) and 3.5 (3.4) of K10A (K10E). The variation suggests that the transient interaction with lysate has a higher specificity than that with NaCl. Combining CSP scaling factors, $d\Delta\delta/d[lysate]$ and the relaxation rates, one can see that all three sites appear to interact with lysate. K19A and K19E have the largest CSP scaling factors and $\Delta R_1R_2$ decreases in the presence of lysate, suggesting the redistribution of lysate caused by this residue mutation might be the largest. D40 and K10 respond differently to lysate (Figures 5−7) although the two residues are spatially close to each other. Lysate specificity to these three residues might be related to the different electrostatic potential around these three residues (Figure S2). In the yeast cytochrome *c* and cytochrome *c* peroxidase transient complex study, the mutant T12A decreases the transient complex population from 30% (WT) to 10%, whereas the mutant R13A increases the population to 80%.[56] Unlike the well-bound protein−protein complex, the transient complex has a shallow free energy profile, so that a single mutation can change both the complex population and the binding mode.[57] Thus, proximate residues can have rather different effect on the transient interaction. The difference between D40 and K10 highlights the difficulty in predicting the specificity of quinary interaction. In principle, if the change of lysate (or BSA) distribution on the GB3 surface caused by the charge changing mutation of GB3 can be determined with high accuracy, one can visualize how the e-field created by a single charge is screened by lysate (or BSA). With the help of paramagnetic relaxation enhancement (PRE) NMR,[58] determining the transient GB3−BSA complex distribution might be feasible.

## CONCLUSIONS

We have studied the e-field generated by protein GB3 side-chain charges in the absence and presence of macromolecular crowders. Whereas the e-field is not affected by Ficoll and dextran, it is generally weakened by lysate. The extent of the weakening depends on the location and the sign of the charge, highlighting the electrostatic nature of the quinary interaction. Similar to NaCl, the e-field weakening by lysate is caused by its redistribution. That is, the e-field generated by a protein surface charge causes the redistribution of lysate, which creates a reaction field and partially neutralizes the protein charge. The presence of lysate also increases the ionic strength of the solution, which plays a minor role in the e-field weakening. Although this study illustrates the importance of electrostatics in quinary interactions, how exactly the e-field weakening is related to the protein structure, especially its surface charge distribution, is not very clear. Understanding this relationship will provide more insights into transient quinary interactions and may eventually permit one to predict these interactions for any protein with a known structure.



DOI: 10.1021/jacs.6b11058
J. Am. Chem. Soc. 2017, 139, 647−654

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b11058.

Figure S1, pulse sequence used for $R_1$ and $R_2$ measurements; Figure S2, charged side chains around K10 of GB3; Figure S3, correlation between $\Delta\delta_{A19K}$ and $\Delta\delta_{E19K}$; Figure S4, average experimental and predicted $R_1R_2$ rate difference; Figure S5, $^1$H pulse length versus Ficoll, BSA, lysate, and NaCl concentration (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

*yaols@qibebt.ac.cn

### ORCID

Lishan Yao: 0000-0003-1797-922X


### Author Contributions

⊥N.Z. and L.A. contributed equally to the work.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant nos. 31270785 to L.Y., 31270905 to Z.L., and 31600690 to L.A.). A portion of this work was performed on the Bruker 600 MHz spectrometer at the National Center for Protein Science Shanghai, CAS, China.

## ■ REFERENCES


(1) Zimmerman, S. B.; Trach, S. O. *J. Mol. Biol.* **1991**, *222*, 599.
(2) Sakakibara, D.; Sasaki, A.; Ikeya, T.; Hamatsu, J.; Hanashima, T.; Mishima, M.; Yoshimasu, M.; Hayashi, N.; Mikawa, T.; Walchli, M.; Smith, B. O.; Shirakawa, M.; Guntert, P.; Ito, Y. *Nature* **2009**, *458*, 102.
(3) Bertrand, K.; Reverdatto, S.; Burz, D. S.; Zitomer, R.; Shekhtman, A. *J. Am. Chem. Soc.* **2012**, *134*, 12798.
(4) Latham, M. P.; Kay, L. E. *PLoS One* **2012**, *7*, e48226.
(5) Sekhar, A.; Latham, M. P.; Vallurupalli, P.; Kay, L. E. *J. Phys. Chem. B* **2014**, *118*, 4546.
(6) Eggers, D. K.; Valentine, J. S. *Protein Sci.* **2001**, *10*, 250.
(7) Stagg, L.; Christiansen, A.; Wittung-Stafshede, P. *J. Am. Chem. Soc.* **2011**, *133*, 646.
(8) Waudby, C. A.; Camilloni, C.; Fitzpatrick, A. W. P.; Cabrita, L. D.; Dobson, C. M.; Vendruscolo, M.; Christodoulou, J. *PLoS One* **2013**, *8*, e72286.
(9) Wirth, A. J.; Platkov, M.; Gruebele, M. *J. Am. Chem. Soc.* **2013**, *135*, 19215.
(10) Latham, M. P.; Kay, L. E. *J. Mol. Biol.* **2014**, *426*, 3214.
(11) Guo, M. H.; Gelman, H.; Gruebele, M. *PLoS One* **2014**, *9*, e113040.
(12) Miklos, A. C.; Sarkar, M.; Wang, Y. Q.; Pielak, G. J. *J. Am. Chem. Soc.* **2011**, *133*, 7116.
(13) Schlesinger, A. P.; Wang, Y. Q.; Tadeo, X.; Millet, O.; Pielak, G. *J. Am. Chem. Soc.* **2011**, *133*, 8082.
(14) Sarkar, M.; Smith, A. E.; Pielak, G. J. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 19342.
(15) Monteith, W. B.; Pielak, G. J. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 11335.
(16) Gershenson, A. *J. Mol. Biol.* **2014**, *426*, 4.
(17) Guzman, I.; Gelman, H.; Tai, J.; Gruebele, M. *J. Mol. Biol.* **2014**, *426*, 11.
(18) Danielsson, J.; Mu, X.; Lang, L.; Wang, H. B.; Binolfi, A.; Theillet, F. X.; Bekei, B.; Logan, D. T.; Selenko, P.; Wennerstrom, H.; Oliveberg, M. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 12402.
(19) Luh, L. M.; Hansel, R.; Lohr, F.; Kirchner, D. K.; Krauskopf, K.; Pitzius, S.; Schafer, B.; Tufar, P.; Corbeski, I.; Guntert, P.; Dotsch, V. *J. Am. Chem. Soc.* **2013**, *135*, 13796.
(20) Goodell, D. J.; Eliseeva, T. A.; Coultrap, S. J.; Bayer, K. U. *PLoS One* **2014**, *9*, e96522.
(21) Reiss, H.; Frisch, H. L.; Lebowitz, J. L. *J. Chem. Phys.* **1959**, *31*, 369.
(22) Zhou, H. X.; Rivas, G. N.; Minton, A. P. *Annu. Rev. Biophys.* **2008**, *37*, 375.
(23) Sharp, K. A. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 7990.
(24) Anfinsen, C. B. *Science* **1973**, *181*, 223.
(25) Mcconkey, E. H. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 3236.
(26) Li, C. G.; Pielak, G. J. *J. Am. Chem. Soc.* **2009**, *131*, 1368.
(27) Crowley, P. B.; Chow, E.; Papkovskaia, T. *ChemBioChem* **2011**, *12*, 1043.
(28) Waudby, C. A.; Mantle, M. D.; Cabrita, L. D.; Gladden, L. F.; Dobson, C. M.; Christodoulou, J. *J. Am. Chem. Soc.* **2012**, *134*, 11312.
(29) Latham, M. P.; Kay, L. E. *J. Biomol. NMR* **2013**, *55*, 239.
(30) Monteith, W. B.; Cohen, R. D.; Smith, A. E.; Guzman-Cisneros, E.; Pielak, G. J. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 1739.
(31) Smith, A. E.; Zhou, L. Z.; Gorensek, A. H.; Senske, M.; Pielak, G. J. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 1725.
(32) Sarkar, M.; Lu, J.; Pielak, G. J. *Biochemistry* **2014**, *53*, 1601.
(33) Wirth, A. J.; Gruebele, M. *BioEssays* **2013**, *35*, 984.
(34) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906.
(35) Markwick, P. R. L.; Bouvignies, G.; Blackledge, M. *J. Am. Chem. Soc.* **2007**, *129*, 4724.
(36) Granata, D.; Camilloni, C.; Vendruscolo, M.; Laio, A. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 6817.
(37) Li, F.; Grishaev, A.; Ying, J.; Bax, A. *J. Am. Chem. Soc.* **2015**, *137*, 14798.
(38) Hass, M. A. S.; Jensen, M. R.; Led, J. J. *Proteins: Struct., Funct., Genet.* **2008**, *72*, 333.
(39) Kukic, P.; Farrell, D.; McIntosh, L. P.; Garcia-Moreno, E. B.; Jensen, K. S.; Toleikis, Z.; Teilum, K.; Nielsen, J. E. *J. Am. Chem. Soc.* **2013**, *135*, 16968.
(40) An, L. Y.; Wang, Y. F.; Zhang, N.; Yan, S. H.; Bax, A.; Yao, L. S. *J. Am. Chem. Soc.* **2014**, *136*, 12816.
(41) Sitkoff, D.; Case, D. A. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 165.
(42) Buckingham, A. D. *Can. J. Chem.* **1960**, *38*, 300.
(43) Yao, L.; Ying, J.; Bax, A. *J. Biomol. NMR* **2009**, *43*, 161.
(44) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(45) Yao, L. S.; Grishaev, A.; Cornilescu, G.; Bax, A. *J. Am. Chem. Soc.* **2010**, *132*, 4295.
(46) Mulder, F. A. A.; de Graaf, R. A.; Kaptein, R.; Boelens, R. *J. Magn. Reson.* **1998**, *131*, 351.
(47) Wang, A. C.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 715.
(48) Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 9179.
(49) Hall, J. B.; Fushman, D. *J. Biomol. NMR* **2003**, *27*, 261.
(50) Yao, L. S.; Vogeli, B.; Ying, J. F.; Bax, A. *J. Am. Chem. Soc.* **2008**, *130*, 16518.
(51) Worrall, J. A. R.; Liu, Y. J.; Crowley, P. B.; Nocek, J. M.; Hoffman, B. M.; Ubbink, M. *Biochemistry* **2002**, *41*, 11721.
(52) Worrall, J. A. R.; Reinle, W.; Bernhardt, R.; Ubbink, M. *Biochemistry* **2003**, *42*, 7068.
(53) Kneller, J. M.; Lu, M.; Bracken, C. *J. Am. Chem. Soc.* **2002**, *124*, 1852.
(54) Wang, Y. Q.; Li, C. G.; Pielak, G. J. *J. Am. Chem. Soc.* **2010**, *132*, 9392.
(55) Harada, R.; Sugita, Y.; Feig, M. *J. Am. Chem. Soc.* **2012**, *134*, 4842.
(56) Volkov, A. N.; Bashir, Q.; Worrall, J. A. R.; Ullmann, G. M.; Ubbink, M. *J. Am. Chem. Soc.* **2010**, *132*, 11487.
(57) Schilder, J.; Ubbink, M. *Curr. Opin. Struct. Biol.* **2013**, *23*, 911.
(58) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108.




DOI: 10.1021/jacs.6b11058
J. Am. Chem. Soc. 2017, 139, 647−654