

Visualization of Transient Ultra-Weak Protein Self-Association in Solution Using Paramagnetic Relaxation Enhancement

Chun Tang,[†] Rodolfo Ghirlando,[‡] and G. Marius Clore*[†]

Laboratories of Chemical Physics and Molecular Biology, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received November 20, 2007; E-mail: mariusc@mail.nih.gov

Abstract: Ultra-weak macromolecular self-association is exceptionally difficult to both detect and visualize using conventional biophysical techniques owing to the very low population of the associated species, yet such weak intermolecular interactions coupled with nucleation events play an important role in driving spontaneous self-assembly to form higher-order architectures (such as crystals, viral capsids, and amyloid fibrils). In this article, we detect and characterize transient, ultra-weak self-association ($K_D \geq 15$ mM) involving the histidine-containing protein HPr by means of paramagnetic relaxation enhancement (PRE), using EDTA–Mn²⁺ conjugated at three separate sites (E5C, E25C, and E32C, one at a time). Large intermolecular PRE effects, above the background observed with hydroxylamine–EDTA–Mn²⁺ as a control, are observed for two of the three paramagnetically labeled sites, E5C and E32C. The extent of self-association can be modulated (significantly reduced) by increasing the ionic strength or by the introduction of a negative charge (S46D mutation) within a positively charged surface patch, and abolished upon the addition of the N-terminal domain of enzyme I (EIN) to form a specific EIN–HPr complex. The PRE profiles observed for E5C and E32C can be fitted simultaneously and accounted for quantitatively using conjoined rigid body/torsion angle dynamics-simulated annealing with an ensemble of states to represent the distribution of one molecule of HPr relative to its partner.

Introduction

Macromolecules can spontaneously form high-order architectures consisting of identical subunits that are either morphologically homogeneous (e.g., crystals, viral capsids, amyloid fibrils) or amorphous (e.g., aggregates).^{1–3} These processes can generally be viewed as a phase transition, whereby molecules gradually shift from bulk solvent to a macromolecular cluster. From a thermodynamic perspective, the molecules at the earliest stages of this process are in rapid equilibrium between the monomer and metastable oligomeric species (generally referred to as a cluster), either dissolving back into solution or being assimilated into the cluster. Once the cluster attains a critical nuclear mass, each additional molecule is efficiently buried into the cluster and self-assembly into higher-order structures becomes an energetically downhill process.¹ The formation of a critical nucleus implies the pre-existence of an array of ultra-weak intermolecular interactions in solution from which a particular set of interactions is selected dependent on the solution conditions. Experimentally, however, visualization, of highly transient, macromolecular oligomers formed by ultra-weak interactions is exceptionally difficult, using conventional

biophysical techniques because the population of these weakly associated species is extremely low.

It has recently been shown that paramagnetic relaxation enhancement (PRE) provides a very sensitive tool for detecting the presence of low (<5%) population, highly transient species in solution.^{4–9} This includes the demonstration of sliding and intersegment transfer during the course of specific protein–DNA recognition,^{4,5} the detection of pre-encounter protein–protein complexes in rapid exchange with the stereospecific complex,^{6,7,9} and the structural characterization of a minor state involving large-scale domain motions in a multidomain protein.⁸ The PRE has also been used to probe the existence of residual transient structures in both unfolded and intrinsically disordered proteins.¹⁰ In the fast exchange regime, the observed transverse PRE rate is a weighted population average of the Γ_2 rates of

- (4) Iwahara, J.; Clore, G. M. *Nature* **2006**, *440*, 1227–1230.
- (5) Iwahara, J.; Zweckstetter, M.; Clore, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15062–15067.
- (6) Tang, C.; Iwahara, J.; Clore, G. M. *Nature* **2006**, *444*, 383–386.
- (7) Suh, J. Y.; Tang, C.; Clore, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 12954–12955.
- (8) Tang, C.; Schwieters, C. D.; Clore, G. M. *Nature* **2007**, *449*, 1078–1082.
- (9) (a) Volkov, A. N.; Worall, J. A.; Holtzmann, E.; Ubbink, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18945–18950. (b) Hulsker, R.; Baranova, M. V.; Bullejahn, G. S.; Ubbink, M. *J. Am. Chem. Soc.* **2008**, ASAP, doi: 10.1021/ja077453p.
- (10) (a) Shortle, D.; Ackerman, M. S. *Science* **2001**, *293*, 487–489. (b) Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M. *J. Am. Chem. Soc.* **2005**, *127*, 476–477. (c) Bertocini, C. W.; Jung, Y. -S.; Fernandez, C. D.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 1430–1435. (d) Bertocini, C. W.; Rasia, R. M.; Lamberto, G. R.; Binolfi, A.; Zweckstetter, M.; Griesinger, C.; Fernandez, C. O. *J. Mol. Biol.* **2007**, *372*, 708–722.

[†] Laboratory of Chemical Physics.

[‡] Laboratory of Molecular Biology.

- (1) García-Ruiz, J. M. *J. Struct. Biol.* **2003**, *142*, 22–31.
- (2) Zlotnick, A. *J. Mol. Recognit.* **2005**, *18*, 470–490.
- (3) Fawzi, N. L.; Okabe, Y.; Yap, E. H.; Head-Gordon, T. *J. Mol. Biol.* **2007**, *365*, 535–550.

the exchanging species in solution. The PRE is very sensitive to the presence of minor species in this regime because the magnitude of the PRE is proportional to the $\langle r^{-6} \rangle$ average distance between a proton and a paramagnetic center, and the effects are large, owing to the large magnetic moment of an unpaired electron. In this article, we extend PRE to the detection, characterization, and visualization, under dilute (sub-millimolar) conditions, of ultra-weak protein self-association that is effectively invisible to other biophysical techniques. The system we have chosen to study is the histidine-containing phosphocarrier protein (HPr) of *Escherichia coli*, a component of the bacterial phosphotransferase system (PTS, ref 11) that has been the subject of extensive NMR^{12,13} and crystallographic investigation.¹⁴

Results and Discussion

Sedimentation Velocity. Wild-type HPr is monomeric in solution as assessed by sedimentation velocity measurements. Using three different loading concentrations (0.02, 0.1, and 1.0 mM), the sedimentation velocity data yield excellent fits to a single species with rms deviations between observed data and best-fit curves of 0.006–0.007 absorbance units (part A of Figure 1). The sedimentation coefficient $s_{20,w}$ (1.31 ± 0.05 S) is independent of the loading concentration, and the mass estimate of 8.8 ± 0.4 kDa is fully consistent with monomeric HPr (which has a calculated molecular weight of 9121 Da). Analysis of the data in terms of a continuous $c(s)$ distribution further confirms the presence of a single species with a weight average $s_{20,w}$ of 1.33 ± 0.05 S (part B of Figure 1). Sedimentation velocity measurements performed on E5C–EDTA–Mn²⁺ HPr yields an $s_{20,w}$ value of 1.38 ± 0.05 S that is essentially identical to that of wild-type HPr (part B of Figure 1). The small difference in average $s_{20,w}$ values between derivatized and wild-type HPr (1.38 versus 1.33 S) can be attributed to the additional molecular weight of the EDTA–Mn²⁺ derivative (378 Da, corresponding to $\sim 4\%$ of the total molecular weight of wild-type HPr). These data indicate that if any soluble higher-order HPr species are present in solution, such species represent less than 1–2% of the total soluble material.

Intermolecular PRE. PRE measurements were carried out on 1:1 mixtures (~ 0.3 mM each) comprising wild-type U-[¹⁵N]-labeled HPr and unlabeled HPr bearing the paramagnetic label, EDTA–Mn²⁺. EDTA was conjugated to HPr via a disulfide linkage to an engineered surface cysteine residue at three different sites individually (E5C, E25C, and E32C; inset in part A of Figure 2) to yield a (cysteaminyl–EDTA)–Cys adduct.

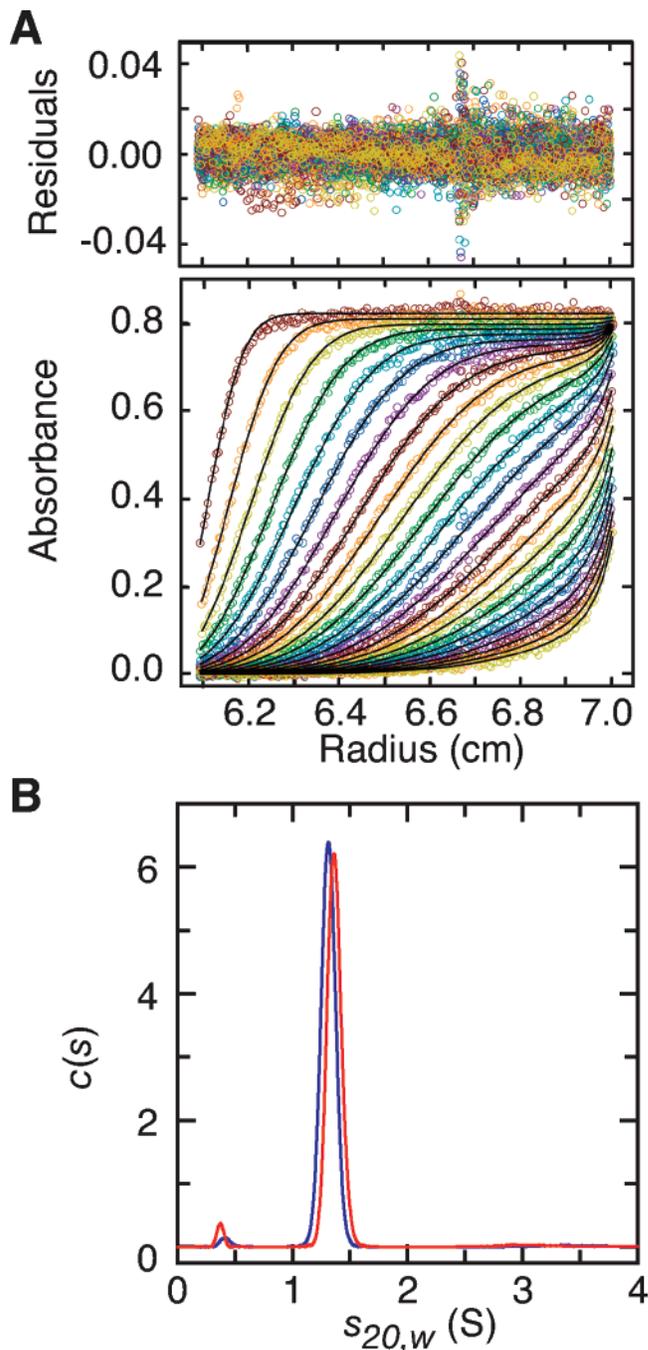


Figure 1. Sedimentation velocity experiments on HPr. (A) Sedimentation velocity profiles at 60 krpm and 20 °C obtained for wild-type HPr at a loading concentration of 0.1 mM showing the absorbance at 234 nm versus the radius. For clarity, scans are shown at 25 min intervals or every fifth scan. The solid lines are best-fit curves to a single species (monomeric HPr). The pooled residuals to this fit are also shown (top panel). (B) Analysis of the data in terms of a continuous $c(s)$ distribution confirms the presence of a single species with an averaged $s_{20,w}$ of 1.33 ± 0.05 S and a molecular mass of 9.4 ± 0.6 kDa for wild-type HPr (colored in blue) and an $s_{20,w}$ of 1.38 ± 0.05 S and a molecular mass of 10.1 ± 0.2 kDa for E5C–EDTA–Mn²⁺ HPr (colored in red).

- (11) Robillard, G. T.; Broos, J. *Biochem. Biophys. Acta* **1999**, *1422*, 73–104.
- (12) (a) Klevit, R. E.; Waygood, E. B. *Biochemistry* **1986**, *25*, 7774–7781. (b) Hammer, P. K.; Waygood, E. B.; Klevit, R. E. *Biochemistry* **1991**, *30*, 18842–11850. (c) Rajagopal, P.; Waygood, E. B.; Klevit, R. E. *Biochemistry* **1994**, *33*, 15271–15282. (d) van Nuland, N. A.; Grötzinger, J.; Dijkstra, K.; Scheek, R. M.; Robillard, G. T. *Eur. J. Biochem.* **1992**, *210*, 881–891. (e) van Nuland, N. A.; Hangyi, I. W.; van Schaik, R. C.; Berendsen, H. J.; van Gunsteren, W. F.; Scheek, R. M.; Robillard, G. T. *J. Mol. Biol.* **1994**, *237*, 544–559. (f) Hess, B.; Scheek, R. M. *J. Magn. Reson.* **2003**, *164*, 19–27.
- (13) (a) Garrett, D. S.; Seok, Y. J.; Peterkofsky, A.; Gronenborn, A. M.; Clore, G. M. *Nature Struct. Biol.* **1999**, *6*, 166–173. (b) Wang, G.; Louis, J. M.; Sondej, M.; Seok, Y. -J.; Peterkofsky, A.; Clore, G. M. *EMBO J.* **2000**, *19*, 5635–5649. (c) Cornilescu, J.; Lee, B. R.; Cornilescu, C.; Wang, G.; Peterkofsky, A.; Clore, G. M. *J. Biol. Chem.* **2002**, *277*, 42289–42298. (d) Williams, D. C.; Cai, M.; Suh, Y. -J.; Peterkofsky, A.; Clore, G. M. *J. Biol. Chem.* **2005**, *280*, 20775–20784.
- (14) (a) Jia, Z.; Quail, J. W.; Waygood, E. B.; Delbaere, L. T. *J. Biol. Chem.* **1993**, *268*, 22490–22501. (b) Napper, S.; Anderson, J. W.; Georges, F.; Quail, J. W.; Delbaere, L. T.; Waygood, E. B. *Biochemistry* **1996**, *35*, 11260–11267.

The latter was chelated to either Mn²⁺ (paramagnetic state) or Ca²⁺ (diamagnetic state). The charge state of HPr is unperturbed by this modification. PRE ¹H_N–Γ₂ rates are given by the difference in R₂ relaxation rates between the paramagnetic and diamagnetic states. Since the ¹H_N–R₂ relaxation rates are measured using a ¹H–¹⁵N correlation-based experiment,¹⁵ and

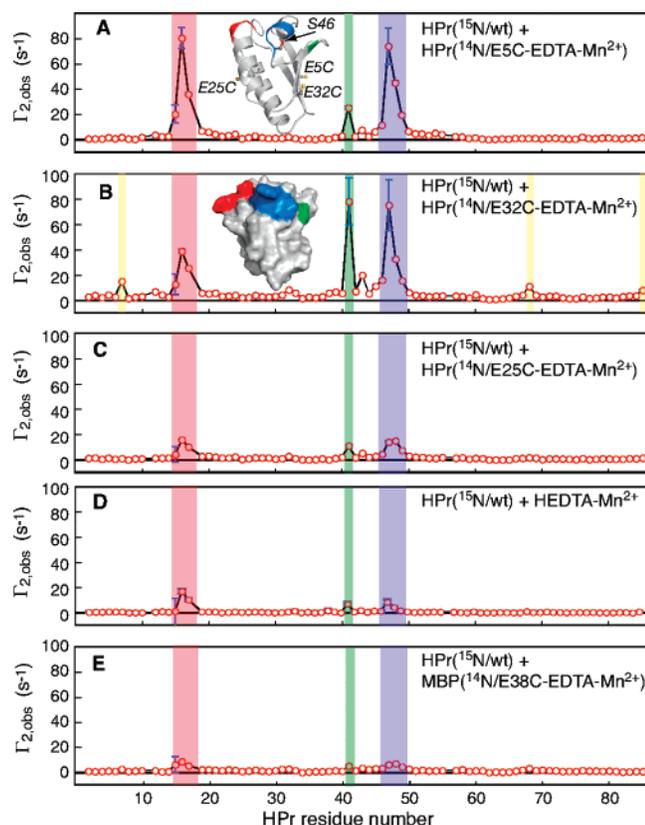


Figure 2. Intermolecular PRE profiles observed between ^{15}N -labeled wild-type HPr and EDTA- Mn^{2+} -conjugated (A) E5C-HPr, (B) E32C-HPr, and (C) E25C-HPr. Negative controls are provided by the intermolecular PRE profiles observed between (D) ^{15}N -labeled HPr, and hydroxylamine-EDTA- Mn^{2+} , and between (E) ^{15}N -labeled HPr and EDTA- Mn^{2+} conjugated to the unrelated protein MBP at E38C. The small PRE effects seen in (D) and (E) arise from very weak background intermolecular interactions involving the EDTA- Mn^{2+} group. The inset in (A) indicates the location of the three conjugation sites (E5C, E25C, and E32C) and of Ser-46. Three groups of residues that experience large intermolecular PREs are colored in red (group 1, residues 15–17), green (group 2, residue 41), and blue (group 3, residues 46–49). A fourth group, comprising residues 7, 67, and 85, displaying small PRE effects only with paramagnetically labeled E32C-HPr, is colored in yellow (B). (Error bars, ± 1 SD).

the observed $^1\text{H}_\text{N}$ - Γ_2 rates arise solely from intermolecular interactions between paramagnetically labeled HPr and ^{15}N -labeled HPr.

Intermolecular PRE profiles observed between paramagnetically labeled HPr and ^{15}N -labeled HPr are displayed in Figure 2. Significant intermolecular PREs, involving three groups of residues (15–17, 41, and 46–49) that form a contiguous surface patch (inset in part B of Figure 2) are observed for EDTA- Mn^{2+} conjugated to the E5C (part A of Figure 2) and E32C (part B of Figure 2) sites, but only minimal effects are observed for the E25C site (part C of Figure 2). A solvent PRE effect arising from diffusion and random elastic collisions is excluded because: (a) solvent PRE effects are negligible (< 1 s $^{-1}$) at the protein concentrations (0.3 mM) employed in the current experiment,^{6,16} and (b) solvent PRE effects would impact all surface residues approximately equally, and all three conjugation sites would display a similar PRE pattern.^{17,18} A control

experiment using hydroxylamine-EDTA- Mn^{2+} (which bears the same charge as EDTA- Mn^{2+} conjugated to a protein) reveals the presence of very small PREs for the same group of residues that can be attributed to very weak background interactions involving the isolated EDTA- Mn^{2+} group (part D of Figure 2). Indeed, the PRE profile for the E25C site is virtually identical to that obtained with hydroxylamine-EDTA- Mn^{2+} , and the PREs for residues 47–49 are barely above the background (cf. parts C and D of Figure 2). Thus, the PRE profile originating from the E25C site also predominantly reflects low-level background interactions with the EDTA- Mn^{2+} group. An almost identical profile to that seen with hydroxylamine-EDTA- Mn^{2+} is also observed using a completely unrelated paramagnetically labeled protein, apo maltose binding protein (MBP) conjugated at E38C, as a control with the paramagnetic label conjugated to a protein surface (part E of Figure 2). In this regard, we note that the secondary β -sheet structure and topology around the E38C conjugation site on MBP¹⁹ is very similar to that around the E5C site on HPr.^{12–14} The magnitudes of the PREs observed for the E5C and E32C conjugation sites are much larger than those observed in the control experiments and can therefore be attributed to ultra-weak self-association between HPr molecules in solution.

The magnitude of the intermolecular PREs between HPr molecules originating from the E5C and E32C sites can be modulated and reduced in several ways. First, the intermolecular PREs can be completely eliminated (parts A and B of Figure 3) upon the addition of the N-terminal domain of enzyme I (EIN), the upstream PTS partner of HPr,^{8,13a} to form an EIN-HPr complex ($K_D = \sim 6$ μM ; ref 7). Because residues 15–17 (group 1 in the PRE profile) and residues 46–49 (group 3 in the PRE profile) comprise a portion of the specific binding interface between EIN and HPr,^{13a} these data indicate that the EIN-HPr interface overlaps with that of the HPr self-association interface. Second, introduction of the S46D mutation, that mimics phosphorylated Ser46 *in vivo*,^{14b} effectively eliminates the intermolecular PREs observed on residues 46–49 (group 3 in the PRE profile) and results in an approximately 2-fold reduction in the intermolecular PREs observed on residues 15–17 (group 1 in the PRE profile, parts C and D of Figure 3). Third, the magnitude of the intermolecular PREs decreases as a function of increasing salt concentration (over a range of 0.1 to 0.5 M NaCl, parts E and F of Figure 3). The salt dependencies observed for the two paramagnetically labeled HPr derivatives, E5C (part E of Figure 7) and E32C (part F of Figure 7), are very similar. Interestingly, the sum of the PRE Γ_2 rates for residues 46–49 display a significantly larger salt dependence than that for residues 15–17 or 41 (slopes in log-log plots of -3.1 to -2.7 versus -0.8 to -0.3), indicating that self-association in the vicinity of residues 46–49 has a larger electrostatic component. These observations suggest that there are multiple self-association configurations, probed by both the E5C and E32C paramagnetically labeled derivatives, that give rise to the observed PRE profiles.

Given the location of the paramagnetic labels and the distinct sites for which intermolecular PREs are observed on ^{15}N -labeled HPr, one can conclude that self-association is asymmetric.

(15) Iwahara, J.; Tang, C.; Clore, G. M. *J. Magn. Reson.* **2007**, *184*, 185–195.

(16) Donaldson, L. W.; Skrynnikov, N. R.; Choy, W. Y.; Muhandiram, D. R.; Sarkar, B.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 9843–9857.

(17) Pintacuda, G.; Otting, G. *J. Am. Chem. Soc.* **2002**, *124*, 372–373.

(18) Hernandez, G.; Teng, C. L.; Bryant, R. G.; LeMaster, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 4463–4472.

(19) Sharff, A. J.; Rodseth, L. E.; Spurlino, J. C.; Quijcho, F. A. *Biochemistry* **1992**, *31*, 10657–10663.

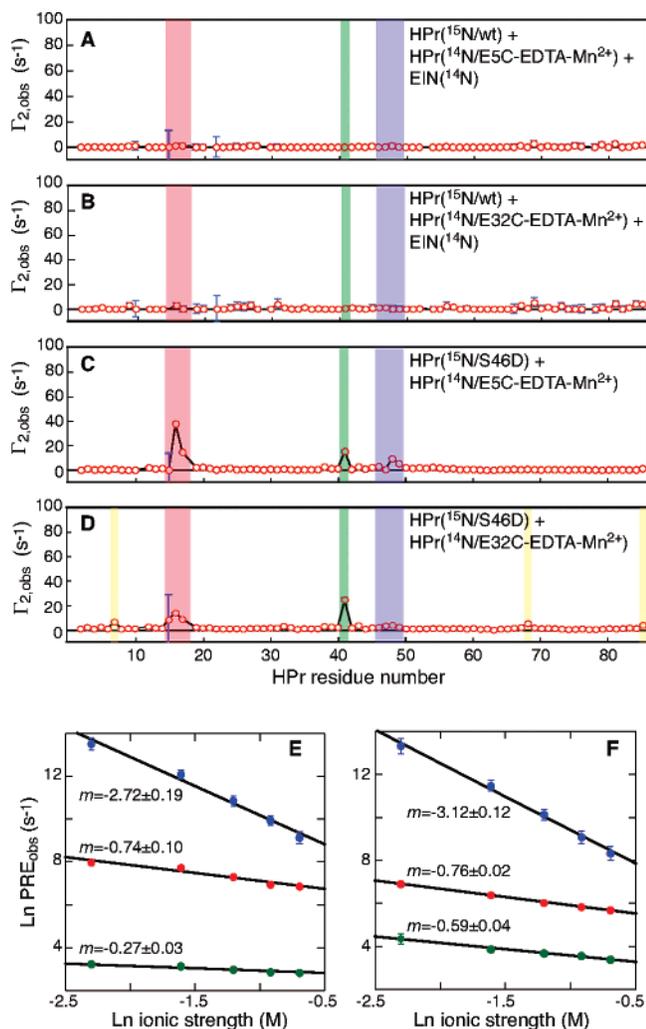


Figure 3. Modulation of intermolecular PREs arising from transient, ultra-weak self-association of HPr. Intermolecular PRE profiles between ^{15}N -labeled wild-type HPr and unlabeled EDTA- Mn^{2+} -conjugated (A) E5C-HPr and (B) E32C-HPr in the presence of 2 equiv of EIN (to ensure saturation of the EIN binding sites on both ^{15}N -labeled and paramagnetically labeled HPr). Intermolecular PRE profiles between ^{15}N -labeled S46D HPr and EDTA- Mn^{2+} -conjugated (C) E5C-HPr and (D) E32C-HPr. The intermolecular PREs for group 3, residues 46–49 (blue), are essentially eliminated by the S46D mutation. Ionic strength dependence of the intermolecular PREs observed between ^{15}N -labeled wild-type HPr and EDTA- Mn^{2+} -conjugated (E) E5C-HPr and (F) E32C-HPr. The PREs are plotted as the sum of the PRE values of the residues within each of the three groups (red, residues 15–17; green, residue 41; and blue, residues 46–49). The slopes (m) of the linear log–log correlations are displayed on the plots (error bars, ± 1 SD).

Quantitative Analysis of PRE Data. To analyze the PRE data quantitatively, we made use of conjoined rigid body/torsion angle simulated annealing²⁰ to optimize the agreement between observed (after subtraction of the control data obtained with hydroxylamine-EDTA- Mn^{2+}) and calculated Γ_2 rates arising from both the E5C and E32C paramagnetic labels *simultaneously*, as assessed by the PRE Q-factor (Methods section for definition).²¹ The function minimized includes a term for the PREs, a simple van der Waals repulsion term to prevent atomic overlap, and a very weak radius of gyration term to ensure, in the case of ensemble calculations, that each member makes at

least some intermolecular contacts.⁶ The HPr molecule bearing the paramagnetic label is held fixed; the paramagnetic labels at E5C and E32C are represented by three-member ensembles to account for the conformational space sampled by the paramagnetic label;²¹ and the second HPr molecule, corresponding to ^{15}N -labeled HPr, is given translational and rotational degrees of freedom. The second HPr molecule is represented by either a single configuration ($N_e = 1$) or a multiple-member ($N_e = 2, 3, \text{ or } 4$) ensemble with equal occupancies for the constituting members. In the case of the multiple-member ensemble calculations, a separate three-conformer ensemble for the paramagnetic probe is used to describe the interaction of each member of the ensemble with a given paramagnetic label (i.e., there are $3 \times N_e$ conformers for each paramagnetic label). This eliminates site-specific stereochemical restrictions imposed on the conformational space sampled by the paramagnetic labels for one set of intermolecular interactions being transferred to another set of interactions. Atomic overlap is permitted between the ensemble members of ^{15}N -labeled HPr molecules and between the paramagnetic probe and HPr conformers that do *not* interact with one another because these represent separate but rapidly interconverting HPr–HPr interactions that do not coexist simultaneously in solution. Calculations were carried out at populations of 0.5, 1, and 2% for the self-associated states, and a large number of independent calculations were carried out to fully sample the configurational space that is consistent simultaneously with the intermolecular PRE data probed by the E5C and E32C paramagnetic labels.

Single Species ($N_e = 1$) Representation. Reasonable agreement between observed and calculated intermolecular PRE data is obtained with the single-species ($N_e = 1$) representation of the HPr–HPr interaction (Figures 4 and 5). Even at a population of 0.5%, the overall average PRE Q-factor, $\langle Q \rangle$, is 0.33 ± 0.01 (which can be decomposed into Q-factors of 0.29 ± 0.01 and 0.35 ± 0.00 for the E5C and E32C data, respectively). Insignificant improvements in $\langle Q \rangle$ values, accompanied by lower van der Waals repulsion energies, are observed as the population of the self-associated species is increased from 0.5 to 1%, with no further changes as the population is increased to 2% (Figure 4). Correlation plots for the single-species ($N_e = 1$) representation at a population of 1% for the self-associated state are shown in part A of Figure 5. Several outliers are clearly apparent, including the PRE data for His-15, Ser-41, and Lys-49 in the case of E5C–EDTA- Mn^{2+} HPr (left-hand panel, part A of Figure 5) and Thr-7, Thr-16, and Arg-17 for E32C–EDTA- Mn^{2+} HPr (right-hand panel, part A of Figure 5): in each instance the calculated PRE values are much smaller than the calculated values. We also note that very high convergence is obtained with a coordinate precision of 0.35 \AA (given by the $C\alpha$ rms difference between the individual ^{15}N -labeled HPr molecules and the mean coordinate positions with the paramagnetically labeled HPr superposed). This converged structure could potentially represent the predominant self-associated species in solution, as has been assumed for several weak electron-transfer complexes derived from PRE data that are not fully satisfied.^{9,22} However, such an interpretation, in the absence of functional data to support a single major species, may be erroneous and does not appear to be justified in the present case.

(20) Schwieters, C. D.; Clore, G. M. *J. Magn. Reson.* **2001**, *152*, 288–302.

(21) Iwahara, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 5879–5896.

(22) Vlasie, M. D.; Fernandez-Busnadiego, R.; Prudencio, M.; Ubbink, M. J. *Mol. Biol.* **2008**, *375*, 1404–1415.

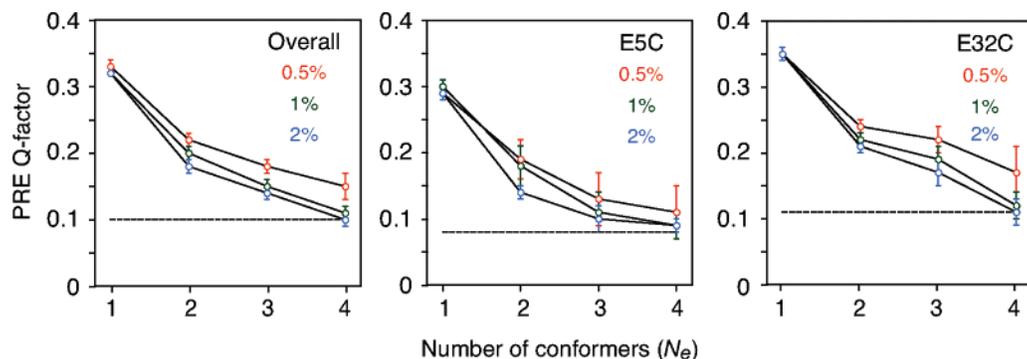


Figure 4. PRE Q-factor as a function of number of conformers (N_e) and the population used to represent the self-associated species. The left-hand panel shows the average overall Q-factor, whereas the middle and right-hand panels display the average Q-factors for the E5C and E32C data individually (error bars, ± 1 SD). Note that the E5C and 32C data are fit simultaneously in the conjoined rigid body/torsion angle dynamics-simulated annealing calculations. Calculations were carried out for three different populations of self-associated species: 0.5% (red), 1% (green), and 2% (blue). The dashed lines denote the expected PRE Q-factors when agreement between observed and calculated Γ_2 rates is comparable to the experimental error in the measurements. To obtain the expected PRE Q-factor, the experimental errors for the individual Γ_2 rates are multiplied by a random number between -1 and 1 , and the resulting values are added to the corresponding experimental Γ_2 rates; a Q-factor is then calculated from the calculated and experimental Γ_2 rates, and the process is repeated 100 times to obtain an average expected Q-value. The average expected Q-factors are 0.10 ± 0.02 for all of the PRE data together, and 0.08 ± 0.02 and 0.11 ± 0.04 for the E5C and E32C PRE data, respectively.

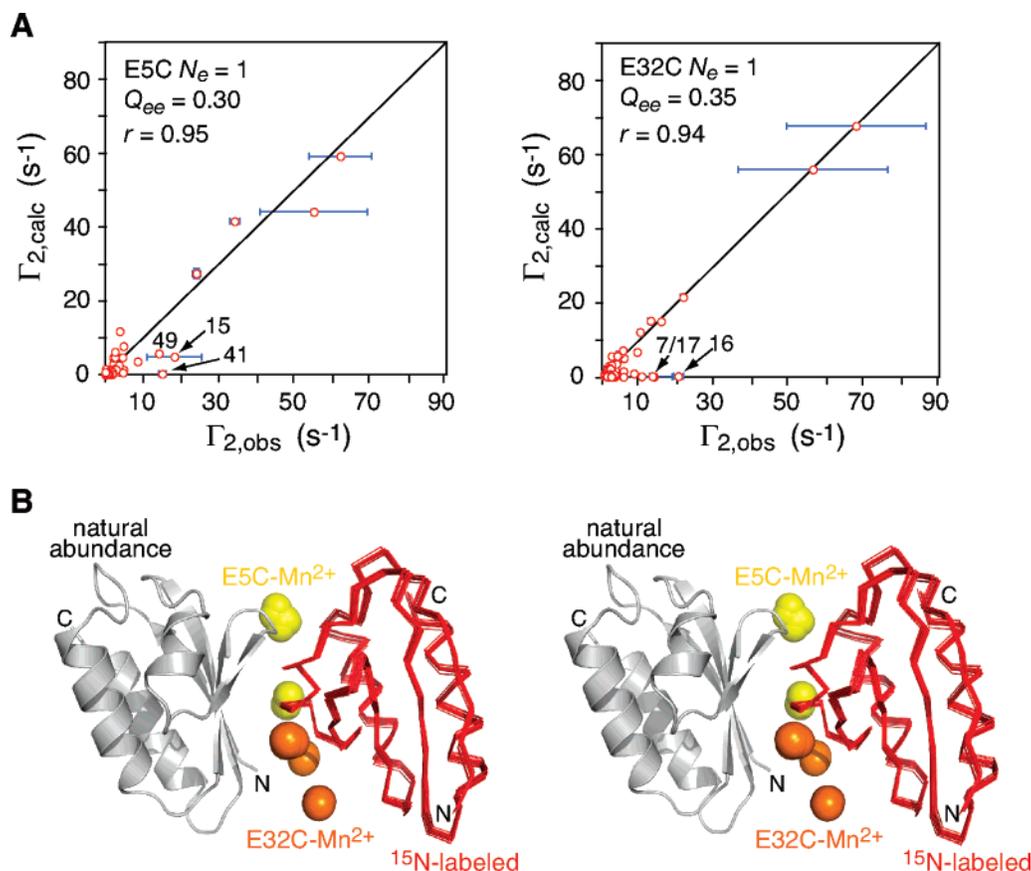


Figure 5. Single-species representation ($N_e = 1$) for the self-associated state. (A) Correlation between observed and calculated intermolecular PREs between ^{15}N -labeled HPr and paramagnetically labeled HPr with the label conjugated at E5C (left-hand panel) and E32C (right-hand panel) for a single-species ($N_e = 1$) representation at a population of 1% for the self-associated state. The values of the ensemble average Q-factor (Q_{ee} , see Methods section for definition), and correlation coefficient (r) are indicated in the figures (error bars, ± 1 SD). (B) Stereoview of the 20 calculated structures with the paramagnetically labeled HPr molecule (ribbon diagram in gray) superimposed and the ^{15}N -labeled HPr molecule (C α atom trace) shown in red. The conformational space sampled by each paramagnetic label is represented by a three-conformer ensemble, and the Mn^{2+} atoms conjugated at E5C and E32C are colored in yellow and orange, respectively. The average C α rms difference between the ^{15}N -labeled HPr molecules and the mean coordinate positions is 0.35 \AA .

Multiple Species ($N_e = 2$ to 4) Ensemble Representations.

The above results suggest that an ensemble of self-association interactions is required to fully account for the experimental PRE data at the level of the errors in the measurement. Calculations were carried out for $N_e = 2, 3$, and 4 . A substantial decrease in PRE Q-factors is seen for $N_e = 2$, leveling off at N_e

$= 4$ (Figure 4). At a population of 1% for the self-associated states, the overall average PRE Q-factor for $N_e = 4$ is 0.11 ± 0.02 , with Q-factors of 0.09 ± 0.02 and 0.12 ± 0.02 for the E5C and E32C data, respectively. These values are very close to the expected values when agreement between observed and calculated PRE data is comparable to the experimental error

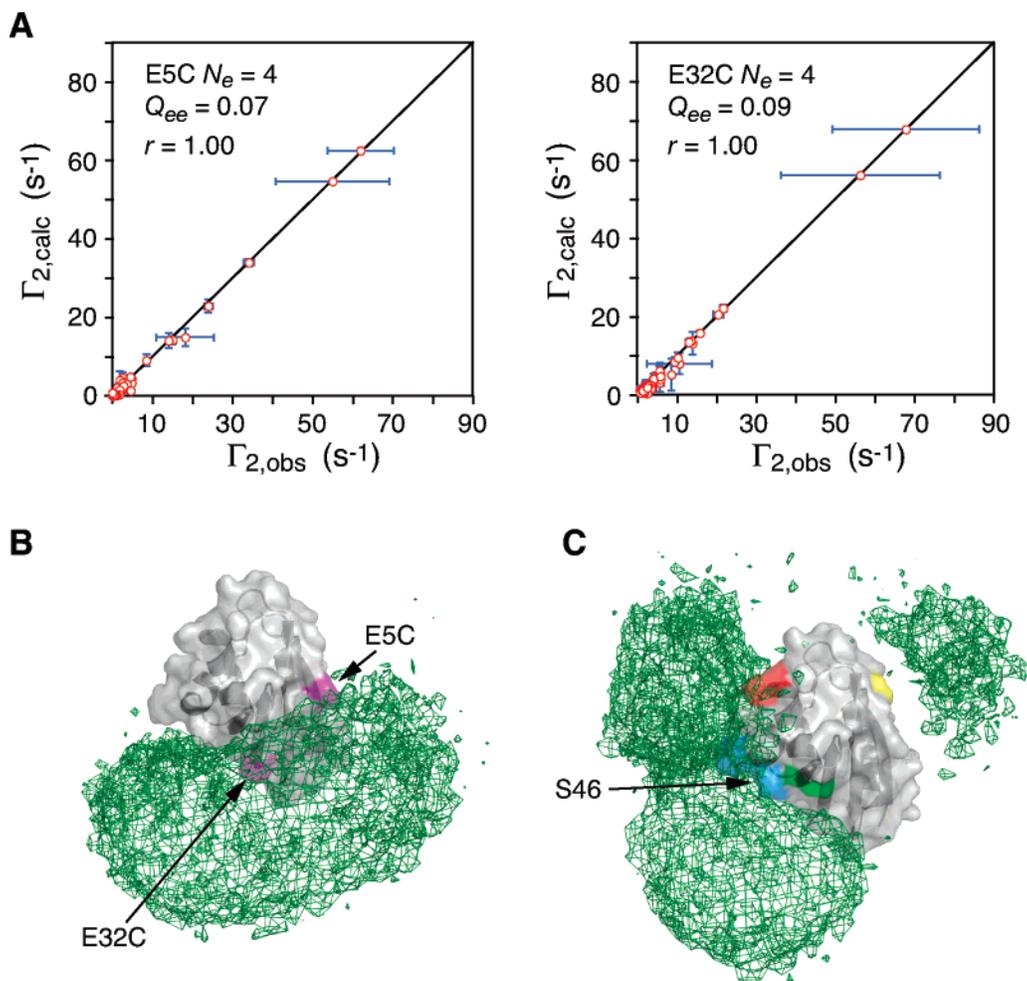


Figure 6. $N_e = 4$ ensemble representation for the self-associated state. (A) Correlation between observed and calculated intermolecular Γ_2 rates between ^{15}N -labeled HPr and paramagnetically labeled HPr with the label conjugated at E5C (left-hand panel) and E32C (right-hand panel) for an $N_e = 4$ ensemble representation at a population of 1% for the self-associated state. Q_{ee} is the ensemble of ensembles average Q-factor (error bars, ± 1 SD). (B) Reweighted atomic probability density map²³ (green mesh), plotted at 20% threshold, showing the distribution of ^{15}N -labeled HPr relative to paramagnetically labeled HPr (gray transparent surface and ribbon, with the location of the E5C and E32C sites indicated in magenta). (C) Reweighted atomic probability density map (green mesh), plotted at 20% threshold, showing the distribution of the paramagnetically labeled HPr molecule relative to the ^{15}N -labeled HPr molecule (gray surface and ribbon, with the four groups of residues that experience large intermolecular PREs colored in red, green, blue, and yellow using the same color scheme as in Figure 2).

(Figure 4). With the available PRE data, the use of an ensemble size $N_e > 4$ is unjustified and would result in over-fitting of the data. Correlation plots between observed and calculated ($N_e = 4$, population 1%) Γ_2 rates are shown for the E5C and E32C data in part A of Figure 6. The agreement with the experimental PRE data is significantly worse when the population is reduced from 1 to 0.5% (and, in addition, the van der Waals repulsion energies are significantly higher) but is essentially unchanged when the population is increased from 1 to 2% (Figure 4). Because the PRE experiments were carried out at a concentration of 0.3 mM of paramagnetically labeled and ^{15}N -labeled HPr, one can conclude on the basis of a population of $\sim 1\%$ for the self-associated heterodimeric species that the apparent equilibrium dissociation constant, K_D , for HPr–HPr self-association observed by the PRE measurements is ≥ 15 mM (given the binomial distribution of homo- and hetero-associated states).

For the ensemble calculations, the PRE data do not yield unique solutions but a range of solutions. The configurational space consistent with the PRE data is sampled by carrying out a large number of independent calculations, and the distributions are best represented by a weighted atomic probability density

map²³ derived from these calculations. The distribution of ^{15}N -labeled HPr about the paramagnetically labeled HPr for the $N_e = 4$ ensemble is shown in part B of Figure 6. The configurational space sampled by ^{15}N -labeled HPr is distributed about the E5C and E32C conjugation sites, where the structure of HPr is mostly β -sheet. (Note that the distribution observed for the $N_e = 3$ calculations is very similar to that for the $N_e = 4$ calculations, except that the $N_e = 3$ distribution is slightly narrower and does not extend quite as far up the sides, in the view shown in part B of Figure 6, as the $N_e = 4$ distribution). The complementary picture showing the distribution of paramagnetically labeled HPr about ^{15}N -labeled HPr is shown in part C of Figure 6. (This is obtained by simply superimposing all molecules within the ^{15}N -labeled HPr ensembles, and calculating the atomic probability density map for paramagnetically labeled HPr.) The distribution of paramagnetically labeled HPr relative to ^{15}N -labeled HPr is largely bimodal, centered about Ser-46 (part C of Figure 6), thus accounting for the impact of the S46D point mutation on HPr self-association.

(23) Schwieters, C. D.; Clore, G. M. *J. Biomol. NMR* **2002**, *23*, 221–225.

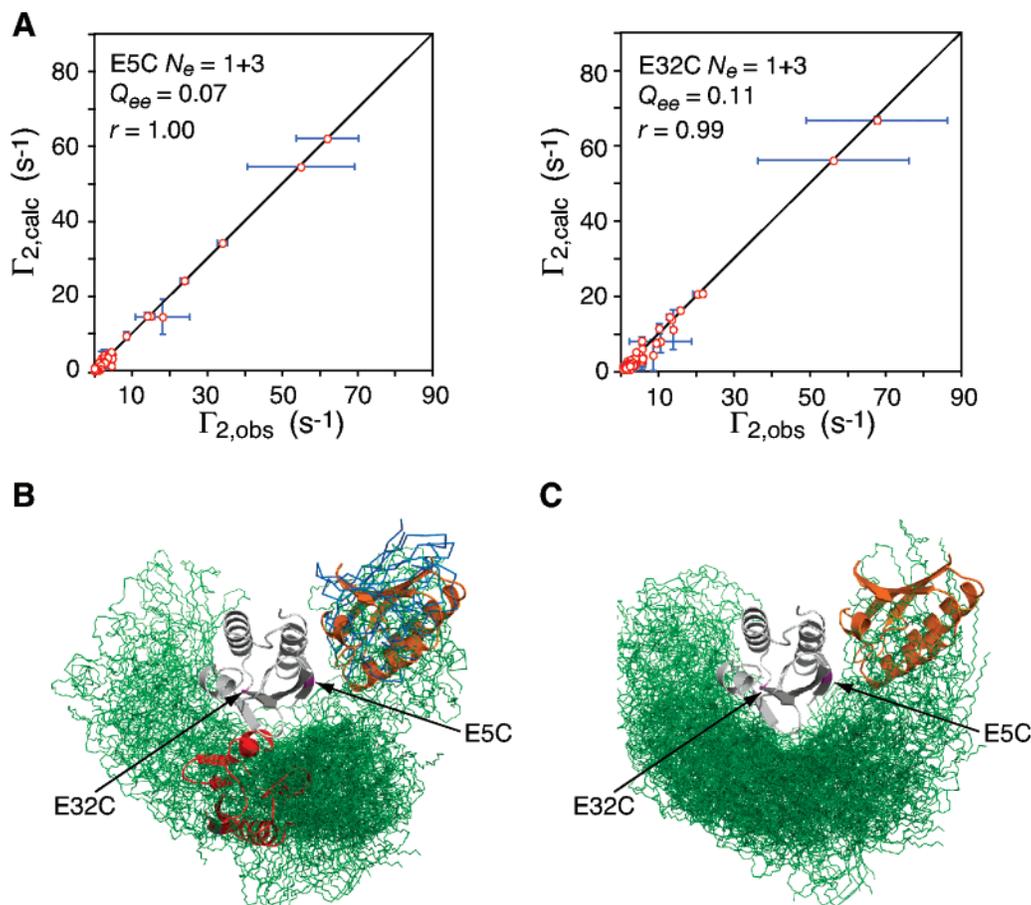


Figure 7. $N_e = (1 + 3)$ ensemble representation for the self-associated state. (A) Correlation between observed and calculated intermolecular Γ_2 rates between ^{15}N -labeled HPr and paramagnetically labeled HPr with the label conjugated at E5C (left-hand panel) and E32C (right-hand panel) for the $N_e = (1 + 3)$ ensemble representation at a population of 1% for the self-associated state partitioned between a population of 0.25% for the $N_e = 1$ converged structure and 0.75% for a $N_e = 3$ ensemble. Q_{ee} is the ensemble of ensembles average Q-factor (error bars, ± 1 SD). (B) Structures (backbone atoms) of the top 20 ensembles (20×3 structures, green) illustrating the distribution of the $N_e = 3$ ensemble of ^{15}N -labeled HPr molecules relative to the paramagnetically labeled HPr molecule (gray ribbon), with the $N_e = 1$ converged structure (held fixed in the calculations) shown as a red ribbon, and the X-ray symmetry partner from the $P2_1$ crystal lattice shown as an orange ribbon. Members of the $N_e = 3$ ensembles that are in close proximity to the location of the symmetry partner in the $P2_1$ crystal lattice, differing by only small translations, are colored in blue. (C) For comparison, the conformers of the top 20 ensembles (20×4 structures) obtained with the $N_e = 4$ representation are shown using the same view and color scheme as in panel (B).

A small cluster of paramagnetically labeled HPr molecules is also observed on the opposite site of the ^{15}N -labeled HPr molecule and accounts for the small intermolecular PREs for two of the residues (Thr-7 and Glu-85) comprising the fourth group of residues (colored in yellow in part B of Figure 2) probed by EDTA– Mn^{2+} conjugated at the E32C site. (Note the third residue of this group, Thr-68, is located at the bottom tip of HPr in the view shown in part C of Figure 6.)

$N_e = (1 + 3)$ ensemble representation. In the $N_e = 4$ calculations, there are structures that are close to the converged structure obtained with the single-species ($N_e = 1$) representation. We therefore carried out another set of calculations in which the converged $N_e = 1$ structure was held fixed and supplemented by a $N_e = 3$ ensemble. The total population of the self-associated state was set to 1% and partitioned between the fixed $N_e = 1$ structure (0.25%) and the $N_e = 3$ ensemble (0.75%). The agreement with the experimental PRE data is comparable to that obtained with the $N_e = 4$ ensemble, with an average overall PRE Q-factor of 0.13 ± 0.02 , and Q-factors of 0.09 ± 0.02 and 0.15 ± 0.03 for the E5C and E32C data, respectively (part A of Figure 7). A comparison of the distribution of ^{15}N -labeled HPr relative to paramagnetically

labeled HPr for the $N_e = (1 + 3)$ and $N_e = 4$ representations is afforded by parts B and C of Figure 7. The overall configurational space sampled is similar but the conformers within the $N_e = (1 + 3)$ distribution are more clustered, whereas those in the $N_e = 4$ distribution are more evenly spread out. In particular, there is a greater concentration of ensemble members in the immediate vicinity of the $N_e = 1$ species in the $N_e = (1 + 3)$ representation than in the $N_e = 4$ representation, whereas the density of ensemble members on the far left side of E32C (in the view shown in parts B and C of Figure 7) is reduced in the $N_e = (1 + 3)$ representation. The latter probably accounts for the marginally worse agreement with the E32C data in the case of the $N_e = (1 + 3)$ representation ($\langle Q \rangle = 0.15 \pm 0.03$, $Q_{ee} = 0.11$) versus the $N_e = 4$ representation ($\langle Q \rangle = 0.12 \pm 0.02$ and $Q_{ee} = 0.09$). In addition, in the $N_e = (1 + 3)$ representation there is a distinct, but small, cluster of structures in close proximity to the location of the symmetry partner in the $P2_1$ lattice of the crystal structure of wild-type HPr,^{13a} including several (colored in blue in part B of Figure 7) that differ from the position of the symmetry partner by only small translations. This observation could be taken to suggest that the interaction seen in the crystal lattice between symmetry partners is sampled

in solution and represents a minor subset of the ensemble of self-associated states. In this regard, it is interesting to note that the X-ray symmetry partner can account in part for the E5C data by simply optimizing the ensemble space occupied by the paramagnetic label with a Q-factor of 0.22 ± 0.01 ($r = 0.98$) and only a few outliers (including Ser-41 and Lys-49) with observed values that are much larger than the calculated ones, but does not account at all for the E32C data (Q-factor = 0.99 ± 0).

Interactions Involved in HPr Self-Association. The average buried accessible surface area observed for the ensemble of self-association states in the $N_e = 4$ calculations is $410 \pm 190 \text{ \AA}^2$ with a gap index (ratio of gap volume to buried accessible surface area) of 11 ± 8 , spanning a range of 2.5 to 37 and clustering between values of 4 to 8. The corresponding indices for the $N_e = 1$ converged structures (buried accessible surface area of 500 \AA^2 and gap index of 6.3) fall within the range observed in the ensemble calculations. These values indicate that the intermolecular interfaces for the self-associated states are in general rather loosely packed,²⁴ consistent with ultra-weak binding ($K_D \geq 15 \text{ mM}$). The observation of a salt dependence of the intermolecular PRE effects (parts E and F of Figure 3) clearly indicates that electrostatic interactions contribute, at least in part, to HPr self-association. Indeed, the E5C and E32C conjugation sites are located within largely negative patches on the surface of HPr: Glu-32 is part of a cluster comprising Glu-66 and Glu-68, and Glu-5 is separated from Glu-32 by a short contiguous patch of polar residues. The complementary regions on which intermolecular PRE effects are observed have positive character arising from Arg-17 (region 1), Lys-40 (region 2), and Lys-45 and Lys-49 (region 3). Ser-46 is located between Lys-45 and Lys-49, and its mutation to Asp almost eliminates the intermolecular PREs observed in region 3 and significantly reduces those on the neighboring region 2 (parts C and D of Figure 3). By way of contrast, the E25C conjugation site is sandwiched between two positively charged residues, Lys-24 and Lys-79, which may account for the observation that the intermolecular PREs observed with E25C-EDTA-Mn²⁺ are minimal and essentially at the level of background. However, electrostatic interactions are not the only interactions driving self-association, as evidenced by the differential effects of salt on the intermolecular PRE effects, which are 3–8-fold larger for region 3 (parts E and F of Figure 3) than regions 1 and 2 (parts E and F of Figure 3). This suggests that a significant contribution from hydrophobic interactions is also present, including intermolecular contacts involving the aromatic ring of Phe-48, which protrudes from the surface of region 3. The importance of hydrophobic interactions is highlighted by the observation that the majority of atoms at the interfaces of the calculated ensembles are nonpolar ($62 \pm 15\%$ nonpolar atoms versus $38 \pm 15\%$ polar atoms for the $N_e = 4$ ensemble, and 62% and 38%, respectively, for the $N_e = 1$ structures).

Concluding Remarks. In this article, we have succeeded by means of PRE measurements in detecting and visualizing transient, ultra-weak self-association states of a protein in solution that are invisible to other biophysical techniques. The weak intermolecular interactions observed for HPr encompass an ensemble of states and are driven by both electrostatic and

hydrophobic interactions. The experimental approach presented in this article, based on the observation of intermolecular PREs, may potentially be useful for characterizing the early stages of biologically significant oligomerization events, including the assembly of viral capsids and the formation of amyloid fibrils.

Materials and Methods

Sample Preparation and NMR Spectroscopy. The HPr (E5C, E25C, and E32C) and MBP (E38C) mutants were constructed using QuikChange (Stratagene). Conjugation to a cysteamine-EDTA moiety via a disulfide linkage was carried out as described previously,⁶ with the EDTA group preincubated with either Ca²⁺ (for diamagnetic control) or Mn²⁺ (paramagnetic state). The proteins employed in the current study, HPr, MBP, and EIN, were purified according to previously published protocols.^{6,8} NMR samples at a concentration of $\sim 0.3 \text{ mM}$ ¹⁵N-labeled and paramagnetically labeled HPr were prepared in buffer comprising 20 mM sodium phosphate, pH 6.5, and 100 mM NaCl. (The ionic strength dependence of the intermolecular PREs was investigated by increasing the concentration of NaCl up to 0.5 M.) For negative controls, 0.3 mM hydroxylamine-EDTA-Mn²⁺ or 0.3 mM MBP conjugated to EDTA-Mn²⁺ at E38C were mixed with 0.6 mM ¹⁵N-labeled HPr. PRE ¹H_N-Γ₂ data were collected at 35 °C on a Bruker DMX500 spectrometer equipped with a z-gradient cryoprobe using a ¹H-¹⁵N correlation-based pulse scheme with a two-time-point measurement as described in ref 15.

Sedimentation Velocity. Sedimentation velocity experiments were conducted at 20 °C on a Beckman Optima XL-A analytical ultracentrifuge. HPr samples (loading volume of 360 μL in the NMR buffer) were studied at various loading concentrations and at a rotor speed of 60 krpm. Aluminum double-sector cells were used and data were acquired as single absorbance measurements at wavelengths such that the loading absorbance was approximately 1.0 (i.e., $\lambda = 256, 238, 234,$ or 230 nm depending on the protein concentration). A radial spacing of 0.003 cm was used, and 120–150 scans were collected at 5 min intervals. Data were analyzed in *SEDFIT*β2d both in terms of a single, non-interacting discrete species or a continuous *c(s)* distribution.^{25,26} The solvent density ρ , viscosity η , and protein partial specific volumes v were calculated using *SEDNTERP 1.1* (Hayes DB, Laue T and Philo J, <http://www.jphilo.mailway.com>). Sedimentation coefficients s were corrected to $s_{20,w}$. Single-species analysis was implemented using time-independent noise corrections; continuous *c(s)* distributions cover a $s_{20,w}$ range of 0.005–5.0 S with a resolution of 100.

Simulated Annealing Calculations. All conjoined rigid body/torsion angle dynamics-simulated annealing calculations²⁰ based on the PRE data were carried out with *Xplor-NIH*²⁷ as described previously.⁶ To account for the flexibility of the linker between the paramagnetic center and the protein backbone, the paramagnetic probe was represented by a three-conformer ensemble.^{6,21} The PRE correlation time τ_c is given by $(\tau_r^{-1} + \tau_s^{-1})^{-1}$, where τ_r is the rotational correlation time and τ_s is the electron relaxation time. For an HPr dimer, $\tau_c = 4.0 \text{ ns}$, with $\tau_r = 6.8 \text{ ns}$ (obtained from the experimentally measured value of 3.4 ns for monomeric HPr, ref 6) and $\tau_s = 9.6 \text{ ns}$.^{15,21} The PRE restraints were obtained by subtracting the control Γ₂ rates obtained with hydroxylamine-EDTA-Mn²⁺ from the Γ₂ rates obtained with EDTA-Mn²⁺ conjugated to E5C and E32C HPr. For a given structure calculation, the PRE Q-factor is given by $[\sum_i \{\Gamma_{2,i}^{\text{obs}} - p\Gamma_{2,i}^{\text{calc}}\}^2 / \sum_i (\Gamma_{2,i}^{\text{obs}})^2]^{1/2}$, where $\Gamma_{2,i}^{\text{obs}}$ is the observed transverse PRE rate for residue i , $\Gamma_{2,i}^{\text{calc}}$ the calculated PRE rate, and p is the population of the self-associated state(s).^{6,21} In the case of calculations employing an ensemble size of greater than 1 to represent the self-associated states, $\Gamma_{2,i}^{\text{calc}}$ is the average of the PRE rates of the members of the ensemble. The ensemble of the ensemble average Q-factor, Q_{ee} , over all calculated structures or ensembles is given by $[\sum_i \{\Gamma_{2,i}^{\text{obs}} - \langle p\Gamma_{2,i}^{\text{calc}} \rangle\}^2 / (\sum_i \Gamma_{2,i}^{\text{obs}})^2]^{1/2}$.^{6,8}

(25) Schuck, P. *Anal. Biochem.* **2003**, *320*, 104–124.

(26) Lebowitz, J.; Lewis, M. S.; Schuck, P. *Protein Sci.* **2002**, *11*, 2067–2079.

(24) Jones, S.; Thornton, J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13–20.

Calculations were performed starting from randomized positions of the HPr molecules and random conformations of the paramagnetic probe. Refinement was carried out simultaneously against the PRE data obtained for EDTA–Mn²⁺ conjugated to the E5C and E32C sites. Structures were ranked according to PRE Q-factors²¹ and van der Waals repulsion energies,^{27,28} and the top 20 ensembles with the smallest PRE Q-factors were used for subsequent analysis.

-
- (27) Schwieters, C. D.; Kuszewski, J.; Clore, G. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 47–62.
(28) Clore, G. M.; Gronenborn, A. M. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 29–63.
(29) Delano, W. L. *The PyMol Molecular Graphics System*, <http://www.pymol.org>, 2002.

Figures were rendered with *PyMol*,²⁹ and reweighted atomic probability maps were computed using *Xplor-NIH*²⁷ as described.²³

Acknowledgment. We thank Charles Schwieters for many stimulating discussions. This work was supported by the Intramural program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, and by the AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G.M.C.)

JA710493M