Supporting Information

Kurnik et al. 10.1073/pnas.1118640109

SI Text

SI Data Analysis and Methods. ***NMR data analysis.*** To measure the level of secondary structure in folded $S6^{+1-17}$, we used the SSP scale proposed by Forman-Kay (1), where SSP = 1 corresponds to fully formed α-helical structure and SSP = −1 corresponds to a fully formed β structure. SSP values for $S6^{+1-17}$ were calculated from the combined use of C^{α} and C^{β} secondary chemical shifts. As reference we used the SSP values of wild-type $S6^{+17-17}$ calculated from published H^{α} chemical shifts (2). Comparison of the SSP values for wild-type $S6^{+17-17}$ and $S6^{+1-17}$ indicates that the secondary-structure elements are maintained upon supercharging the protein (Fig. S2). The data indicates that both α1 and α2, as well as the strands β1 and β4, remain wild-type like in $S6^{+1-17}$. Although the missing assignments in parts of β2 and β3 in $S6^{+1-17}$ make blind spots in the comparison with the wild-type protein, the positions where data overlap show folded structure. The overall integrity of the supercharged structure was further tested by correlating the SSP values of $S6^{+1-17}$ with those of wild-type $S6^{+17-17}$, using a running average over 5 residues to filter local scatter (Fig. S2). The data shows a linear correlation with a slope of 1.0 and r = 0.95. This correlation is very similar to that of wild-type $S6^{+17-17}$ and the structurally ordered circular permutant $S6^{54-55}$ (2) (Fig. S2). Taken together, these data provide strong evidence that one-sided removal of the positive charges of S6 to obtain the supercharged species $S6^{+1-17}$ has negligible effect on the folded structure. This, in turn, suggests that neutralization of the remaining negative charges by protonation to obtain $S6^{+1}$ should also not have any effect on the folded structure: if anything this stability-promoting relief of intramolecular repulsion should be structurally consolidating, consistent with the highly dispersed HSQC spectrum of $S6^{+1}$ at pH 1 in manuscript Fig. 2.

At low ionic strength the NMR spectrum of $S6^{+1-17}$ shows two distinct sets of peaks. One well-dispersed set that corresponds to the folded state (N), and one characteristically collapsed set that corresponds to the unfolded state (D) (Fig. S4). The result is characteristic for a two-state protein poised at equilibrium between D and N. Upon relieving the intramolecular charge repulsion with 100 mM NaCl, the folding equilibrium of $S6^{+1-17}$ is shifted entirely to the native state, producing a single set of well-dispersed HSQC cross-peaks (manuscript Fig. 2).

Kinetics data analysis. The data in Fig. S5 and Table S1 show that the stability change of mutating R and K to S at pH 6.3 ($\Delta\Delta G_{D-N}^{pH6.3} = 2.89 - 8.24 = -5.35$ kcal/mol) is larger than at pH 2.3 ($\Delta\Delta G_{D-N}^{pH2.3} = 3.99 - 5.72 = -1.73$ kcal/mol). The reason for this difference is at least twofold. First, mutation of $S6^{+17-17}$ at pH 6.3 includes the breakage of surface salt bridges, whereas mutation of $S6^{+17}$ at pH 2.3 does not. Second, mutation at pH 6.3 leaves an unfavorable species with unmatched negative charges ($S6^{+17-17} \rightarrow S6^{+1-17}$), whereas, reversely, mutation at pH 2.3 eliminates an unfavorable species with unmatched positive charges ($S6^{+17} \rightarrow S6^{+1}$). Deconvolution of these effects would require determination of the stability change upon deprotonation of the K and R moieties at high pH. This cannot be achieved without complete substitution of, not only D and E, but also the S and Y side chains, as these deprotonate to become negative in an overlapping pH interval. Construction of such S6 constructs involving 23 point mutations has so far been precluded for stability reasons. The same arguments hold for the kinetics. The negligible effect on k_f upon mutation of the R and K side chains at pH 2.3, does not necessarily mean that the kinetic effect of removing positive charges is smaller than that of removing negative charges. It can also be the result of a simultaneous retardation of k_f by the destabilizing mutations.

Mutagenesis, expression and purification of $S6^{+1-17}$. $S6^{+1-17}$ was based on the wild-type DNA sequence of $S6^{+17-17}$ from *Thermus thermophilus* (3): ATGCGTCGTTACGAAGTAAACATCGTACTGAACCCGAACCTGGACCAGAGCCAGCTGGCTCTGGAAAAAGAAATCATCCAGCGTGCTCTGGAAAACTACGGCGCTCGTGTTGAAAAAGTTGAAGAACTGGGCCTGCGTCGTCTGGCTTACCCGATCGCTAAAGACCCGCAGGGCTACTTCCTGTGGTACCAGGTTGAAATGCCGGAAGACCGTGTTAACGACCTGGCTCGTGAACTGCGTATCCGTGACAACGTTCGTCGTGTTATGGTTGTTAAAAGCCAGGAACCGTTCCTGGCTAACGCTTGA. In $S6^{+1-17}$, all 16 basic residues were mutated to serines (R2S, R3S, K23S, R28S, R36S, K39S, R46S, R47S, K54S, R71S, R77S, R80S, R82S, R86S, R87S, K92S):

ATGAGCTCTTATGAAGTGAACATTGTTCTGAATCCGAACCTGGATCAGAGTCAACTGGCGCTGGAATCCGAAATCATTCAGTCAGCCCTGGAAAATTACGGCGCATCGGTCGAAAGCGTGGAAGAACTGGGTCTGTCTAGTCTGGCTTATCCGATCGCGTCCGACCCGCAAGGCTACTTTCTGTGGTATCAGGTTGAAATGCCGGAAGATTCAGTCAACGACCTGGCCTCGGAACTGAGCATTTCTGATAATGTTAGTAGCGTTATGGTCGTTTCTAGCCAAGAACCGTTCCTGGCAAACGCCTAA. Transformants harbouring the $S6^{+1-17}$ construct were grown at 37 °C in LB medium containing 100 μg/mL carbenicillin. Overexpression was induced by 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG) at $OD_{600} = 0.5$. Growth was continued for 4 h before harvesting by centrifugation at 5000 rpm. in a Beckman Avanti J-20 centrifuge, JLA 8.1000 rotor. Isotope-labeled protein was grown in M9 minimal medium at 23 °C, supplemented with 100 μg/mL carbenicillin, 1 g/liter $^{15}NH_4Cl$, and 4 g/liter $^{13}C^6$ glucose, and induced using 0.5 mM IPTG overnight.

Harvested cells were resuspended in 50 mM Tris-HCl, pH 7.5. All purification steps were performed at 4 °C unless otherwise stated. Cell suspensions were treated with DNase and RNase (Sigma-Aldrich), lysed by ultrasonication, and centrifuged at 15,000 rpm. in a Beckman Avanti J-25 centrifuge, JA-25.50 rotor. The supernatant was dialysed at room temperature against 10 mM Tris-HCl, pH 7.5. The precipitate formed during dialysis was removed by centrifugation at 18,000 rpm. in a Beckman Avanti J-25 centrifuge, JA-25.50 rotor. The supernatant was purified on a Q-sepharose anion exchange chromatography column (GE Healthcare) using a linear gradient of 0-1 M NaCl in 10 mM Tris, pH 7.5. Fractions containing $S6^{+1-17}$ were further purified on a Sephacryl S-100 HR gel filtration column (GE Healthcare). Purity was analyzed using the Ready Gel SDS-PAGE system (Bio-Rad) and by electrospray ionization mass spectrometry and Edman degradation performed at the Protein Analysis Center (Karolinska Institute) after dialysis against milliQ H_2O and centrifugation at 18,000 rpm. in a Beckman Avanti J-25 centrifuge, JA-25.50 rotor.

1. Marsh JA, Singh VK, Jia Z, Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: Implications for fibrillation. *Protein Sci* 15:2795–2804.
2. Ohman A, Oman T, Oliveberg M (2010) Solution structures and backbone dynamics of the ribosomal protein S6 and its permutant P(54–55). *Protein Sci* 19:183–189.
3. Otzen DE, Kristensen O, Proctor M, Oliveberg M (1999) Structural changes in the transition state of protein folding: Alternative interpretations of curved chevron plots. *Biochemistry* 38:6499–6511.
4. Tanford C (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv Protein Chem* 24:1–95.
5. Oliveberg M, Arcus VL, Fersht AR (1995) pKA values of carboxyl groups in the native and denatured states of barnase: The pKA values of the denatured state are on average 0.4 units lower than those of model compounds. *Biochemistry* 34:9424–9433.
6. Haglund E, et al. (2009) The HD-exchange motions of ribosomal protein S6 are insensitive to reversal of the protein-folding pathway. *Proc Natl Acad Sci USA* 106:21619–21624.
7. Lindberg MO, Haglund E, Hubner IA, Shakhnovich EI, Oliveberg M (2006) Identification of the minimal protein-folding nucleus through loop-entropy perturbations. *Proc Natl Acad Sci USA* 103:4083–4088.
8. Haglund E, Lindberg MO, Oliveberg M (2008) Changes of protein folding pathways by circular permutation. Overlapping nuclei promote global cooperativity. *J Biol Chem* 283:27904–27915.
9. Lindberg MO & Oliveberg M (2007) Malleability of protein folding pathways: A simple reason for complex behaviour. *Curr Opin Struct Biol* 17:21–29.
10. Silow M, Oliveberg M (1997) Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc Natl Acad Sci USA* 94:6084–6086.
11. Silow M, Tan YJ, Fersht AR, Oliveberg M (1999) Formation of short-lived protein aggregates directly from the coil in two-state folding. *Biochemistry* 38:13006–13012.
12. Cho SS, Weinkam P, Wolynes PG (2008) Origins of barriers and barrierless folding in BBL. *Proc Natl Acad Sci USA* 105:118–123.
13. Yang WY, Gruebele M (2003) Folding at the speed limit. *Nature* 423:193–197.

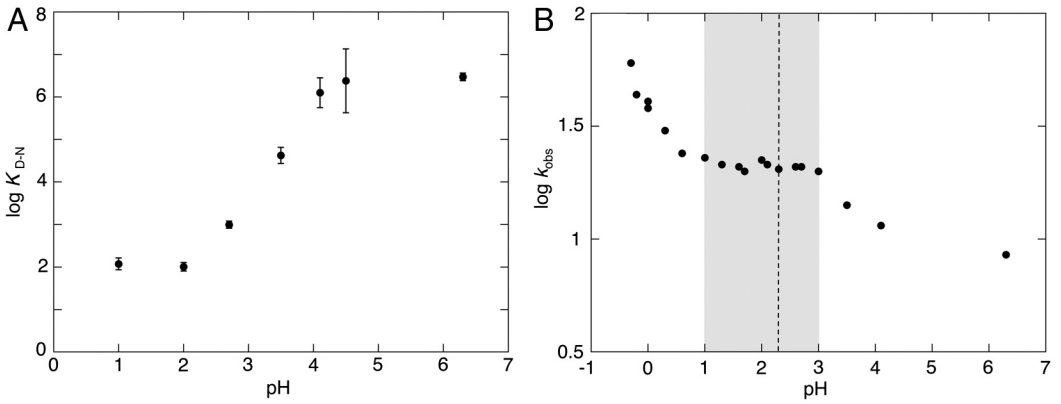
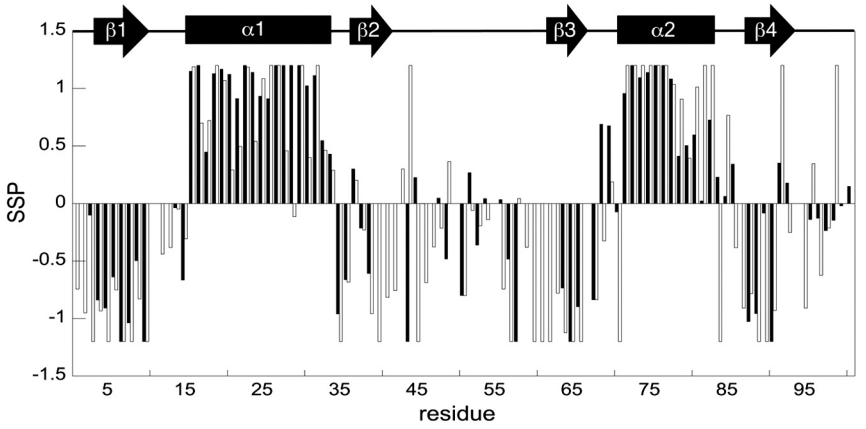


Fig. S1. Protonation of wild-type $S6^{+17-17}$ and $S6^{+1-17}$ measured by the pH dependence of protein stability at 0 M denaturant and the unfolding rate constant at 3.6 M GdmCl, respectively. *A*. The stability of wild-type $S6^{+17-17}$, measured here by $\log K_{D-N} = \log[D]/[N]$ from urea-chevron data, shows a characteristic decrease between pH 5 and pH 3, where $\partial \log K_{D-N}/\partial pH = \Delta H^{+}$ is the number of protons taken up in the unfolding transition (4) (5). Accordingly, $\partial \log K_{D-N}/\partial pH$ becomes zero below pH 2.3 where both D and N are fully protonated. The pK_a shifts of the carboxylate groups in the native state of wild-type $S6^{+17-17}$ are thus not sufficiently large for unfolding the protein at low pH: $\log K_{D-N}$ levels out at about 2. *B*. The carboxylate side chains of $S6^{+1-17}$, which have no positive charges to interact with, are expected to have pK_a values similar to those of model compounds; i.e., above 4. Consistently, the unfolding rate constant ($\log k_u$), which measures the stability of N relative to the more expanded transition-state ensemble, shows a pronounced change around pH 4 where the protein becomes protonated. Between pH 3 and pH 1, $\partial \log k_u/\partial pH$ is approximately zero (gray interval) indicating that all side chains are fully protonated. In this interval, the protein is fully depleted of side-chain charges ($S6^{+1}$), carrying only a single positive charge at the solvent protruding n-terminus. The dashed line marks pH 2.3, where $S6^{+1}$ was characterized in this study. Below pH 1, protonation of the backbone carbonyl groups with $pK'_a s < 0$ leads to a second increase of $\log k_u$, presumably by destabilization of the secondary-structure elements in the native state.



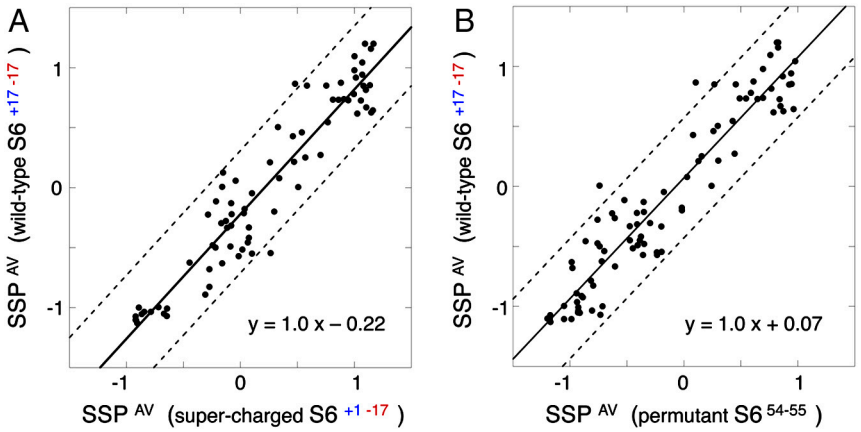


Fig. S2. Secondary-structure propensity (SSP) of $S6^{+1-17}$ (filled bars) and wild-type $S6^{+17-17}$ (open bars). The crystallographic secondary-structure of wild-type $S6^{+17-17}$ is outlined above. *A*. Correlation between the secondary-structure propensity averaged over five residues (SSP^{AV}) of wild-type $S6^{+17-17}$ and the super-charged $S6^{+1-17}$. *B*. Corresponding plot of data from wild-type $S6^{+17-17}$ and the circular permutant $S6^{54-55}$. Dashed lines indicate $SSP^{AV} \pm 0.5$ from the linear fits.

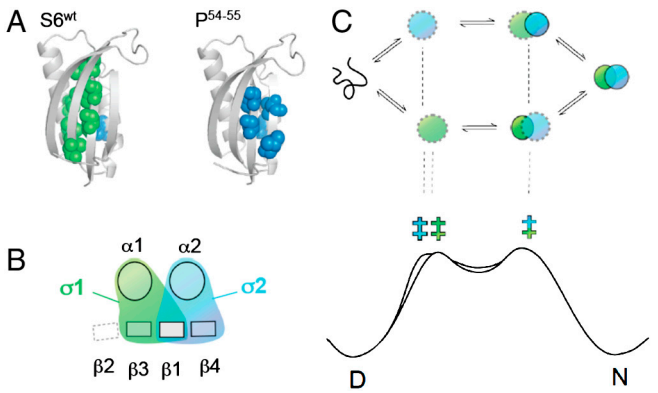


Fig. S3. Schematic outline of how the two competing nuclei of S6 produce parallel pathways of opposed folding order. Figure adapted from (6). *A*. The positions of residues with ϕ-values above 0.4 show that the transition states of wild-type $S6^{+17-17}$ and the circular permutant P^{54-55} are structurally different. The transition state of $S6^{+17-17}$ involves the nucleus $\sigma 1$ (green) and P^{54-55} the nucleus $\sigma 2$ (blue) (7). *B*. Simplistic top view of the S6 topology illustrating that $\sigma 1$ and $\sigma 2$ are partly overlapping by sharing $\beta 1$ (8). The role of this overlap could be to allosterically optimize the global cooperativity: the formation of one nucleus drives the formation of the other (9). *C*. The two ways of nucleating the folding reaction create two competing pathways with opposite folding order. $S6^{+17-17}$ is biased mainly to the $\sigma 1$ pathway, whereas P^{54-55} folds predominantly by the $\sigma 2$ pathway. Below is a schematic outline of the free-energy profiles for the two competing pathways, showing how the folding order of $\sigma 1$ and $\sigma 2$ can be altered by just subtle changes of the local barrier heights (‡). Consistent with this model, the destabilized mutant $S6^{+1-17}$ displays a downward curvature in the chevron plot at high [GdmCl], indicating a Hammond shift between the early and late transition-state features (manuscript Fig. 3).

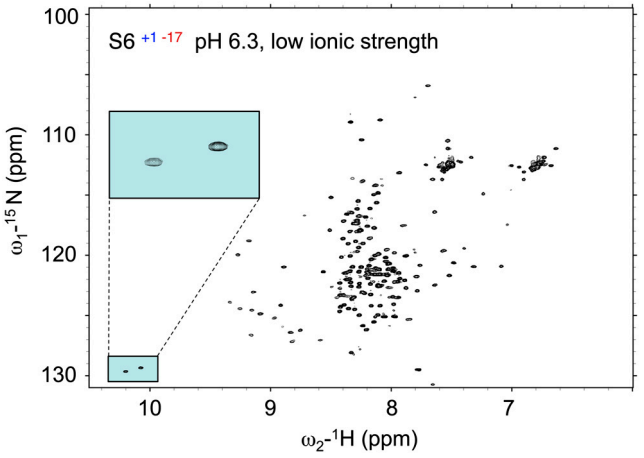


Fig. S4. HSQC spectrum of $S6^{+1-17}$ at low ion strength shows two populations in slow exchange on the NMR time scale. The *Inset* shows the splitting of the tryptophan side chain into two distinct peaks, one corresponding to the denatured ensemble and the other to the native state.

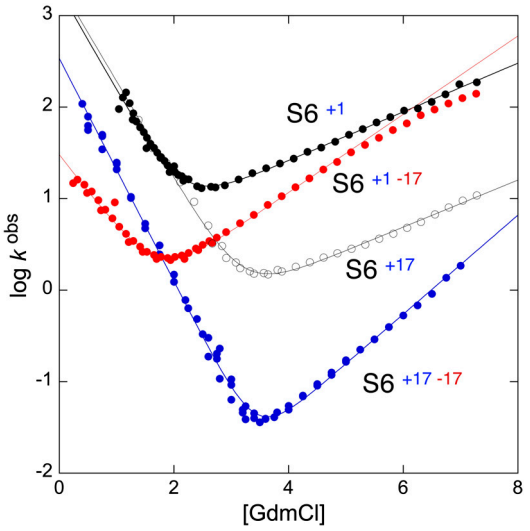


Fig. S5. Chevron plots of wild-type $S6^{+17-17}$ and $S6^{+1-17}$ showing the effect of complete R/K mutation at pH 6.3 and pH 2.3. Protonation of the wild-type protein at pH 2.3 yields the positively supercharged species $S6^{+17}$.

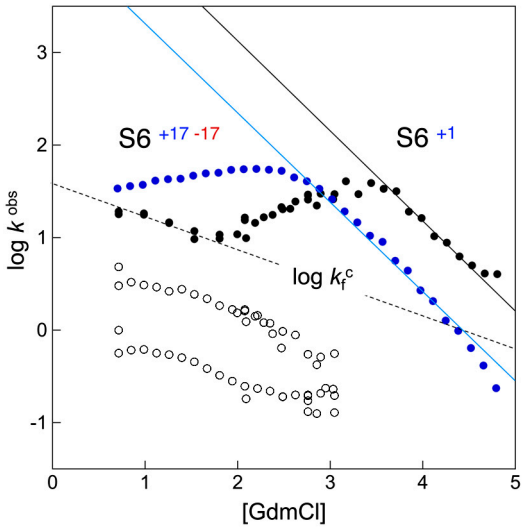


Fig. S6. The refolding regime of the chevron plots of $S6^{+1}$ (black) and wild-type $S6^{+17-17}$ (blue) at a background of $[Na_2SO_4] = 1$ M. The lines indicate log k_f in two-state folding, the rollovers below $[GdmCl] \approx 3$ M indicate the accumulation of a kinetic trap, and the dotted line indicates the kinetic detour log k_f^c according to manuscript Scheme 2. Accordingly, increasing protein stability by decreasing [GdmCl] produces the same kinetic pattern as increasing protein stability by $[Na_2SO_4]$ titration (manuscript Fig. 3 and Fig. S3). Following the accumulation of the kinetic trap of $S6^{+1}$, is the appearance of several slower phases (open circles), the amplitudes of which are higher than for the usual proline cis-trans isomerisation phases (see Fig. S7).

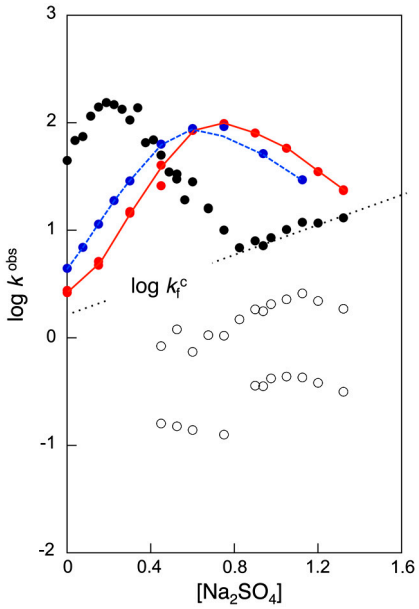


Fig. S7. [Na_2SO_4] titration of log k_f at [GdmCl] = 1.6 M showing slower phases of $S6^{+1}$: $S6^{+1}$ (black), $S6^{+17-17}$ (blue), and $S6^{+1-17}$ (red). The slow phases of $S6^{+1}$ (open circles) appears in connection to the kinetic trap in the rollover regime and has larger amplitudes than typical proline cis-trans isomerisation phases. The [Na_2SO_4] dependence of the slow phases follows that of log k_f^c, cf. manuscript Scheme 2. Matching phases, including one superimposing with log k_f^c, can be resolved for the supercharged $S6^{+1-17}$ (but not for wild-type $S6^{+17-17}$) albeit with significantly lower amplitudes. It is thus possible that the refolding kinetics of $S6^{+1-17}$ will eventually become rate-limited by log k_f^c, as observed for $S6^{+1}$, upon boosting the stabilizing conditions even further. The origin of log k_f^c and the slow phases could be transient aggregation (10, 11) or transitions between multiple intermediates in the collapsed refolding regime. It is also conceivable that the multiphase pattern indicates stretched-exponential kinetics (12, 13).

Table S1. Kinetic parameters and protein stabilities

	$S6^{+17-17}$	$S6^{+17}$	$S6^{+1-17}$‖	$S6^{+1}$
log $k_f^{H_2O}$*	2.53 ± 0.03	3.34 ± 0.06	1.48 ± 0.03	3.28 ± 0.06
m_f*(M^{-1})	−1.22 ± 0.01	−1.06 ± 0.03	−0.82 ± 0.04	−1.10 ± 0.04
log $k_u^{H_2O}$*	−3.51 ± 0.08	−0.86 ± 0.05	−0.64 ± 0.05	0.35 ± 0.04
m_u* (M^{-1})	0.54 ± 0.02	0.26 ± 0.01	0.43 ± 0.01	0.27 ± 0.01
MP † (M)	3.44 ± 0.03	3.18 ± 0.04	1.71 ± 0.05	2.15 ± 0.04
m_{D-N} ‡ M^{-1})	1.76 ± 0.02	1.32 ± 0.03	1.24 ± 0.04	1.36 ± 0.04
log $K_{D-N}^{H_2O}$§	6.04 ± 0.09	4.20 ± 0.08	2.12 ± 0.06	2.93 ± 0.07
$\Delta G_{D-N}^{H_2O}$ ¶ (kcal/mol)	8.24 ± 0.12	5.72 ± 0.10	2.89 ± 0.08	3.99 ± 0.09

*Derived from chevron data according to Eq. 3.
†Transition midpoint derived from the intersect between log k_f and log k_u.
‡$m_{D-N} = m_u$-m_f according to Eq. **2**.
§Calculated from Eq. **2**.
¶Calculated from $\Delta G_{D-N}^{H_2O} = -2.3RT(\log k_u^{H_2O} - \log k_f^{H_2O})$ (Eq. **4**).
‖Due to curvature in the unfolding limb, parameters were obtained from fitting of Eq. **3** to the linear part of the chevron plot at 0-5.1 M GdmCl.