Supplemental Information

Neutralizing positive charges at the surface of a protein lowers its rate of amide hydrogen exchange without altering its structure or increasing its thermostability.

Bryan F. Shaw, Haribabu Arthanari , Max Narovlyansky, Armando Durazo, Dominique P. Frueh, Michael P. Pollastri, Andrew Lee, Basar Bilgicer, Steven P. Gygi, Gerhard Wagner, and George M. Whitesides

Supplemental Materials and Methods

H/D exchange measured by ESI-MS.

Immediately prior to hydrogen exchange, protein aliquots were thawed and concentrated to approximately 20 mg/mL in a Microcon centrifugal filtration device (MW 10,000; Millipore). Concentrated protein samples were transferred to PCR tubes and incubated at 15 °C for 5 minutes and then quickly diluted 1:10 (v/v) into a deuterated phosphate buffer (10 mM, pD 7.4, 15 °C; prepared as previously described $\,1$). The pD value is approximated by adding 0.4 units to the value that is measured by a pH electrode. 2° During isotopic exchange, the temperature was maintained at 15 ºC with a Peltier-effect device (MJ Research). At various time points, aliquots (5µL) were removed and isotopic exchange was quenched for later analysis with ESI-MS. Quenching was done by diluting deuterated protein samples (1:10; v/v ; 5 μ L:50 μ L) into cold, low pH buffer (100 mM

phosphate; pH 2.4, 0 °C) followed by immediate flash freezing in N_2 (1) and storage for \leq 24 hr at -80 ºC.

Immediately prior to mass spectral analysis, each protein sample was thawed and loaded into an ice-chilled Rheodyne injector that was attached to an ice-chilled column for desalting (MacroTrap, Michrom Inc.). The injector was connected to the electrospray ion source and HPLC pump. The desalting column was washed with ice-chilled 0.3 % formic acid/ H_2O (v/v) prior to sample loading. For the purpose of desalting an additional 500 µL of 0.3 % formic acid/H₂O (v/v) was passed through the MacroTrap immediately after sample loading. The injector was then immediately turned to the "inject" position and proteins were eluted from the trap with an ionization solution consisting of formic acid/H₂O/acetonitrile (0.3/49.8/49.8 % v/v). This solution was delivered to the electrospray ion source by an HPLC pump (running at $80 \mu L/min$). Mass spectra were consistently acquired less than 2 minutes after thawing of protein samples. Mass spectrometric data was collected with a Q-STAR hybrid quadrupole-time of flight mass spectrometer (Applied Biosystems). Calibrations were performed daily using the +2 and +3 charges state of angiotensin.

The rate of deuterium incorporation was monitored as a function of time. The value "number of unexchanged hydrogen" (H_{unex}) was calculated by subtracting the mass of the protein after a given time, t, in deuterated buffer and under conditions that favor the native state (denoted $M[D]_{\text{Native},t}$) from the mass of the thermally unfolded (and therefore perdeuterated) protein (denoted $M[D]$ _{Unfolded}).

 $H_{\text{unex}} = M[D]_{\text{Unfolded}} - M[D]_{\text{Native, t}}$

The following sum of three exponentials was fit to the hydrogen exchange data for all 19 rungs of the charge ladder:

$$
M_{D,t} = M_{D,\infty} - Ae^{-k_1t} - Be^{-k_2t} - Ce^{-k_3t}
$$

Where $M_{D,t}$ = mass at each time point in deuterated buffer; $M_{D,\infty}$ = mass of completely exchanged protein; A, B and C denote the number of amide hydrogens undergoing exchange with fast $(k_1 > 1 \text{ min}^{-1})$, intermediate $(0.01 \text{ min}^{-1} < k_2 < 1 \text{ min}^{-1})$ and slow $(k_3 <$ 0.01 min^{-1}) rate constants. The presence of the slow-exchanging amide hydrogen atoms proved troublesome in the curve fitting. In order to make the fittings converge, an artificial endpoint time and mass value (t = 43,200 min, mass = $M_{D,\infty}$) was added to each data set.

Back exchange controls were performed as previously described.¹ The backexchange control samples are necessary because 1) it accounts for the loss of incorporated deuterium that occurs during quenching, desalting and mass spectral analysis and 2) it accounts for the increased number of exchangeable amide protons that can be expected as a result of lysine acetylation. Back-exchange controls were done by heating the same BCA II protein solutions that were used to collect mass-time points at 80 ºC for 10 min. These back-exchange control samples were then cooled and incubated at 15 ºC for 10 min before being quenched and frozen as described above.

HSQC and TROSY enhanced NMR analysis of carbonic anhydrase II.

The spectral widths were 16.033 ppm for proton (centered on the water resonance at 4.690 ppm) and 36 ppm for nitrogen (centered at 118 ppm). A recycling delay of 0.9 s was used and 16 scans were accumulated per increment; $512(^1H) \times 128(^{15}N)$ complex

points were recorded. The total measuring time was 1 hour for each experiment. The data was processed with the NMRPipe package. The ¹⁵N dimension was linear predicted and zero filled to a final size of 1024 (1 H) \times 1024 (15 N). The resulting NMR spectrum was analyzed using the software CARA (available from http://www.nmr.ch/) and intensities of the peaks were measured using the program Sparky.

 Separately, the spectral widths were 16.033 ppm for proton (centered on the resonance of water at 4.70 ppm) and 5 ppm for nitrogen (centered at 128 ppm). A recycling delay of 0.9 s was used and 16 scans were accumulated per increment; 512 (^1H) \times 24 (¹⁵N) complex points were recorded. The total measuring time was (15 min) for each experiment. The data was processed with the NMRPipe package. The ¹⁵N dimension was linearly predicted and zero filled to a final size of 1024 (1 H) \times 128 (1 ⁵N)

Preparation of lysine and -ε-NH3 acetylated lysine derivatives.

Reagent **1** (below) was purchased from Acros, Inc., and PS-DIEA and MP-TsOH were obtained from Biotage, Inc. All other reagents were purchased from Sigma-Aldrich and were used as received. To a solution of 30 mg Ac-Lys $(\epsilon\text{-}NH_3^+)$ -NHMe (0.15 mmol, 1.0 eq) in 1.5 mL anhydrous DMF was added 150 mg of PS-DIEA (4 mmol/g loading, 0.6 mmol, 4 eq) and 34 uL acetic anhydride (0.36 mmol, 2.4 eq). The reaction was stirred for 3 hours, and 133 mg MP-TsOH was added (4.5 mmol/g loading, 0.6 mmol, 4 eq) to scavenge any unreacted amine. The reaction mixture was filtered, and the solids were rinsed with DMF. The combined filtrate was evaporated to dryness to provide an analytically pure sample of 2 in 83% yield (30.3 mg) , a white solid (mp: 183-184). ¹H NMR (CD₃OD, 300 MHz): 4.22 (dd, J₁=8.8 Hz; J₂= 6.4 Hz, 1H), 3.15 (t, J= 6.9 Hz, 2H),

2.72 (s, 3H), 1.98 (s, 3H), 1.92 (s, 3H), 1.75 (m, 1H), 1.63 (m, 1H), 1.50 (m, 2H), 1.35 (m, 2H). ¹³C NMR (CD₃OD, 125 MHz): 173.8, 172.2, 172.0, 53.7, 39.0, 31.5, 28.8, 25.1, 23.1, 21.3, 21.2. HRMS (MH⁺): Calculated: 244.1661; Found: 244.1668.

Measuring the rate of amide H/D exchange of model lysine amino acids at pD 4.5 with 1 H NMR spectroscopy.

HPLC purified compounds 1 and 2 were dissolved separately in 200 mM acetate-d₃ buffer (pH 3.95) at a concentration of 10 mg/mL. After each compound was dissolved, the pH of each solution was measured again and adjusted to $pH = 3.95$. Hydrogen exchange was initiated by diluting an aliquot of each solution (45 μL) into 450 μL of $D₂O$. Analysis with NMR spectroscopy began immediately after dilution and the first data point was typically collected at 4 min 53 s after dilution into D_2O . The exchange of the backbone amide hydrogen with deuterium was monitored by the disappearance of amide signals at approximately 8.1 and 7.8 ppm. The final pH of each solution was measured after analysis and found to be pH=4.08 (pD \approx 4.5). Each time point was an average of 16 scans collected over 60 s. A zero time point was collected by diluting an identical aliquot into 95% H_2O , 5% D_2O ; 5% of the integrated signal intensity was added back to the measured integral of signal intensity in order to account for decreased signal that arises from 5 $\%$ D₂O.

Measuring the rate of amide hydrogen-hydrogen exchange of model lysine amino acids at pH 7.1 with water proton magnetization transfer 1 H NMR.

Water proton magnetization transfer NMR experiments: all experiments were performed using a Bruker Avance 900 MHz spectrometer that was equipped with a 5 mm cryogenic probe. Water proton magnetization was inverted 180° using a selective DANTE (Delay Alternating with Nutation for Tailored Excitation) pulse. After magnetization, lysine compounds were allowed to exchange with water (e.g., magnetized protons from water were allowed to exchange with unmagnetized NH protons) for a variable amount of time ranging from 0.1 s to 4s (referred to as t_{mix}), followed by a jump and return pulse sequence³ (τ = 168 μ s) in order to suppress water magnetization and maximize detection in the amide NH spectral region (7.8-8.2 ppm). These hydrogen exchange experiments were collected with 32768 points, a 16 ppm spectral width and a recycle delay of 10s. The data points that were collected at increasing values of t_{mix} were collected in a scrambled order. The data was processed and the integrated using the software Mestrenova.

Inversion recovery NMR experiments: A DANTE pulse sequence⁴ was used to selectively invert water proton magnetization. The sequence consisted of 6 hard 30° pulses separated by a delay of 60 μs. Z-field gradients of 80% and 3% were used to dephase transverse water proton magnetization that might have remained after the DANTE pulse and t_{mix} . This dephasing with a Z-field gradient was done in order to reduce the effects of radiation damping.⁵ A 10-second recycle delay was used to allow full recovery of water magnetization to equilibrium.

Supplemental Results

Kinetic analysis of hydrogen exchange data derived from LC-ESI-MS.

The global exchange of amide hydrogen in proteins, when measured by ESI-MS, is typically analyzed by fitting a tri-exponential function to the kinetic data; this analysis yields parameters in terms of "fast", "intermediate", and "slow" exchanging amide hydrogen. We fit the kinetic data for all 19 rungs of the charge ladder to the following triexponential function:

$$
M[D]_{\text{native}} = M[D]_{\text{unfolded}} - Ae^{-k_1t} - Be^{-k_2t} - Ce^{-k_3t} \quad (S1)
$$

In equation S1, $M[D]_{\text{native}}$ equals the measured mass of folded BCA II per time in deuterated buffer; $M[D]_{\text{unfolded}}$ equals the mass of the perdeuterated protein (i.e., the mass of the unfolded protein in deuterated buffer). The pre-exponential factors A, B and C denote the number of amide hydrogen that are undergoing exchange at a relatively fast $(k_1 > 1 \text{ min}^{-1})$, intermediate (0.01 min⁻¹ $\le k_2 < 1 \text{ min}^{-1}$) or slow $(k_3 < 0.01 \text{ min}^{-1})$ rate. For BCA-Ac(0), $k_1=5.8 \text{ min}^{-1}$, $k_2=1.9 \cdot 10^{-1} \text{ min}^{-1}$, and $k_3=1.1 \cdot 10^{-3} \text{ min}^{-1}$ (Supplemental Table 2). The pre-exponential coefficients for BCA-Ac(0) are: A =36.2, B =24.3 and C =96.4 (Supplemental Table 2). These three values are interpreted to mean that: approximately 36 hydrogen underwent exchange at a fast rate $(k_1 > 1 \text{ min}^{-1})$, approximately 24 hydrogen exchange at an intermediate rate $(0.01 \text{ min}^{-1} < k< 1 \text{ min}^{-1})$ and approximately 96 are undergoing slow exchange $(k_3 < 0.01 \text{ min}^{-1})$. We consider the

96 slowest exchanging hydrogen to be "protected" from solvent exchange. The sum of A, B, and C, for BCA-Ac(0) is 156.9 and this is the number of amide hydrogen whose exchange can be measured using our mass spectrometric apparatus and methods.⁶ These six kinetic parameters express the effects of lysine acetylation on the rate of H/D exchange for BCA II (Supplemental Table 2 lists these parameters for all 19 rungs of the charge ladder). For example, the three rate constants, k_1 , k_2 , and k_3 are equal for all rungs of the charge ladder. Two of the pre-exponential coefficients (A and B) are also similar, meaning that each rung of the charge ladder has a similar number of "fast" and "intermediate" exchanging hydrogen. The value of C, however, increases by approximately 1 for each higher rung of the charge ladder, meaning that the addition of 1 acetyl group results in the addition of 1 hydrogen that exchanges with a first order rate constant of $k \le 1 \cdot 10^{-3}$ min⁻¹.

Measuring the rate of amide hydrogen-hydrogen exchange of model lysine amino acids at pH 7.1 with water proton magnetization transfer 1 H NMR.

H/D exchange occurs too rapidly at pH 7.1 to be measured with the NMR methods we used at pD 4.5 (e.g., exchange is complete before the first data point can be collected with NMR, after the addition of D_2O). In order to measure the effect of acetylation on the rate of amide *hydrogen* exchange of Ac-Lys(ε -NH₃⁺)-NHMe at pH 7.1 we performed NMR experiments to measure magnetization transfer from water protons⁵ to amides of Ac-Lys(ε -NHCOCH₃)-NHMe and Ac-Lys(ε -NH₃⁺)-NHMe (Supplemental Figure 5). These experiments involve magnetically labeling water protons (using a radiofrequency pulse to selectively invert the magnetization of water protons), followed by the

measurement of the amide ¹H NMR signal *after* a variable time of solvent-amide exchange (ranging from 0.1s to 3s).⁵

The water proton magnetization experiments on Ac-Lys(ε -NH₃⁺)-NHMe at pH 7.1 demonstrated that the acetylation of ϵ -NH₃⁺ decreased the rate of amide hydrogen exchange by a similar magnitude at pH 7.1 as at pD 4.5. For example, we found that acetylation of Ac-Lys(ε -NH₃⁺)-NHMe reduced the rate of amide hydrogen exchange (at pH 7.1) by a factor of 3.0 for amide 1, and by a factor of 2.9 for amide 2 (Figure S5).

Magnetization experiments were done according to the methods pioneered by Geuron et al.⁵ Briefly, we first measured the efficiency of the inversion of water magnetization (*E*), defined as:

$$
E = 1 - \left[\frac{A_{\text{inverted}}}{A_{\text{equilibrium}}} \right] \tag{S2}
$$

In equation S2, A is the area of proton peak measured at equilibrium or immediately following an inversion pulse. The measured values of *E* of water inversion in the presence of two types of small molecules Ac-Lys(ε-NHCOCH3)-NHMe and Ac-Lys(ε- $NH₃⁺$)-NHMe were very similar, and equal to approximately 1.95; this calculated value of *E* is close to the maximum value of 2 and suggests that the inversion of the population with the DANTE pulse was very efficient.

We next determined the rate of relaxation of water protons (Rl_w) for each sample using the exponential decay:

$$
A = A_o[1 - E e^{(-Rl_w t_{mix})}]
$$
 (S3)

This equation is valid when radiation damping is negligible.⁵ Radiation damping is a phenomenon where a spin at a high concentration induces a current in the receiver coil due to imperfections in the inversion pulse. This current causes the further relaxation of the inverted spins, thus resulting in an apparent longitudinal relaxation rate⁵. Radiation damping is prominent in high-Q probes, especially in cryogenically cooled probes, such as the kind used in this experiment. Because radiation damping contributes significantly to the relaxation of experiments done at $t_{mix} > 800$ ms, we chose to use data points at t_{mix} \leq 800 ms to extract the exponential coefficient Rl_w (using equation S2; Figure S5B) When calculating the amide relaxation and the rate of amide exchange we also used data points at t_{mix} < 800 ms in order to be consistent. The rates that were calculated with this truncated set of data were very similar, however, to values calculated from fitting the full set of data (up to 4 s) to equation S3 and S4 (data not shown).

We measured relaxation properties of water in the presence of both acetylated and non-acetylated lysine derivatives (Figure S5). Small molecules present in the water have a negligible effect on the relaxation properties of water, as is apparent from panels (**B**) and (**F**) in Figure S5.

Finally, we calculated the rate of exchange of longitudinal magnetization from water to amides (k_{ex}) by fitting the areas of amide proton peaks as a function of exchange time, t_{mix} , using equation S4. The value k_{ex} is apparent rate of amide hydrogen exchange in Ac-Lys(ε -NHCOCH₃)-NHMe and Ac-Lys(ε -NH₃⁺)-NHMe.

$$
A = A_{o} - A_{o} k_{ex} E \left[\frac{e^{(-R l_{v} t_{mix})} - e^{(-R l_{vH} t_{mix})}}{R l_{vH} - R l_{w}} \right]
$$
(S4)

Supplemental Figure 1. Lysine acetylation decreases the rate of H/D exchange of BCA II as measured by ESI-MS. A) H/D exchange kinetics of the BCA II charge ladder (90 % D₂O, pD 7.4, 15 °C). Data for all 19 rungs of the charge ladder are shown here. BCA-Ac(0) retained ~85 unexchanged hydrogen after 100 minutes in D₂O. B) Lysine acetylation does not alter the mechanism of hydrogen exchange in BCA II. H/D exchange kinetics for BCA-Ac(0) and BCA-Ac(18) at pD 6.4 and 8.4 (15 °C). The rate of exchange of both proteins was similarly dependent upon the pD of solvent which suggests that both proteins exchange similarily via an EX2 mechanism. Error bars represent the standard deviation of average mass values calculated from seven charge states for each protein. C) A plot of the number of deuterons that can be measured with LC-ESI-MS to be incorporated into each rung of the BCA II charge ladder, as a function of the number of acetyl modification for each rung. The maximum number of deuterons that can be incorporated into each rung (y-axis) is calculated by subtracting the mass of the *perprotonated* rung (e.g., the measured molecular weight of the rung in water) from the mass of the perdeuterated rung (e.g., the measured mass of the thermally unfolded protein in D_2O). Both of these measurements are made after proteins are delivered to the ion source via LC pumping in H_2O , pH 2.4. As can be seen in the plot, the number of deuterons that can be incorporated into BCA II increases by approximately one (on average) with each acetylation, demonstrating that the deuterons on acetyl-amides of lysine-ε-NHCOCH₃ do not completely exchange with H₂O during their short delivery (< 10 min) to the electro-spray ion source. The data on the Y-axis of this plot are the values listed on the far right column in Supplemental Table 1; data on the X-axis are values from the far left column of the same table. Red dashed line indicates the trend line that would be expected if ε-amides of lysine-ε-NHCOCH₃ did undergo complete back-exchange with H_2O during LC-ESI-MS.

Supplemental Figure 2. The N-H plane(2D-HN-HNCO) of a TROSY-HNCO spectrum of peracetylated HCA II. Protein was acetylated with 13 C labeled acetic anhydride. Each signal in this spectrum corresponds to an ¹⁵N-¹H group that is coupled to a ¹³C=O.

Supplemental Figure 3. Amide H/D exchange of acetyl groups in peracetylated HCA II measured by TROSY (15 °C, pD 7.4, 90 % D_2O , 10 mM PO_4^3). The intensity of most peaks decreases to 5-10 % intensity after \leq 27 minutes in D₂O suggesting that the amide groups of acetyl functionalities are undergoing H/D exchange rapidly.

Supplemental Figure 4. H/D exchange of backbone amides in peracetylated HCA II measured by HSQC NMR. A. TROSY-HSQC spectra of unmodified and peracetylated HCA II. B) Three amide signals were chosen from A based upon their observed exchange (e.g., "fast", "medium", "slow").

Supplemental Figure 5. Amide hydrogen exchange of Ac-Lys(ε-NHCOCH3)-NHMe and Ac-Lys(ε -NH $_{3}^{+}$)-NHMe at pH 7.1 measured with water proton magnetization transfer NMR. The left-hand column $(A-D)$ shows results of experiments on $Ac-Lys(\epsilon-NH_3^+)$ -NHMe; and the right-hand column (**E-H**) shows the same experiments on Ac-Lys(ε-NHCOCH3)- NHMe, with a third data set for the additional amide (ε-NHCOCH3). (**A**) Plot of relaxation longitudinal magnetization of water as a function of mixing time (t_{mix}) between magnetization of water protons and measurement of amide ¹H resonance. We used the data points < 800 ms to extract the exponential coefficient RI_w from relaxation data using equation S2, because radiation damping makes a significant contribution to the relaxation for $t_{mix} > 800$ ms. **(B)** Expanded view of graph in **A**. Panels **(C)** and **(G)** are plots of the relaxation of the longitudinal magnetization of the amides, from which we calculate the intrinsic rates of relaxation of each amide in the two lysine derivatives. Panels (**D**) and (**H**) illustrate the exchange of magnetization between water and amide moieties. The resulting data were fitted to a double exponential decay (eqn. S4) to calculate the rates of exchange, denoted k_{ex} . For Ac-Lys(ε -NH₃⁺)-NHMe k_{ex} (NH1) = 10.98 sec⁻¹, k_{ex}(NH2) = 3.65 sec⁻¹; for Ac-Lys(ε-NHCOCH₃)-NHMe, k_{ex}(NH1) = 3.72 sec⁻¹, k_{ex}(NH2) = 1.24 sec⁻¹, k_{ex}(NH3) = 0.32 sec⁻¹.

REFERENCES

- (1) Shaw, B. F.; Durazo, A.; Nersissian, A. M.; Whitelegge, J. P.; Faull, K. F.; Valentine, J. S. *J Biol Chem* **2006**, *281*, 18167-76.
- (2) Connelly, G. P.; Bai, Y.; Jeng, M. F.; Englander, S. W. *Proteins* **1993**, *17*, 87-92.
- (3) Plateau, P.; Geuron, M. *J. Am. Chem. Soc.* **1982**, *104*, 7310-7311.
- (4) Morris, G. A.; Freeman, R. *J. Magn. Reson.* **1978**, *29*, 433-462.
- (5) Gueron, M.; Leroy, J. L. *Methods Enzymol* **1995**, *261*, 383-413.
- (6) The BCA II polypeptide has 239 backbone amide hydrogen atoms (after cleavage of N-terminal methionine and N-terminal acetylation; proline residues do not have an amide hydrogen). In 90 % D_2O , BCA-Ac(0) can, therefore, incorporate 215 amide deuterons. With our methods, however, only 157 deuterons are measured to be incorporated into the unfolded polypeptide—the other 58 undergo back exchange with solvent during quenching and analysis with LC-ESI-MS (Supplemental Table 1).

Supplemental Table 1. Experimentally determined mass values. "M[H]" is the measured mass of BCA-Ac(N) in H_2O . "M[D]_{Native}" is the mass of folded BCA-Ac(N) after 80 min in deuterated buffer (pD 7.4, 15 °C). The difference between "M[D] $_{\text{Native}}$ " and "M[H]" yields the number of deuterons that are incorporated into BCA-Ac(N) at 80 min, pD 7.4, 15 °C. "M[D]Unfolded" is the mass of each rung in deuterated buffer under thermally denaturing conditions and expresses the highest possible mass of each rung under the experimental conditions. The difference between "M[D]Unfolded" and "M[D] $_{\text{Native}}$ " yields the number of unexchanged protons in BCA-Ac(N) at 80 min (i.e. the number of hydrogens that are protected from exchange with solvent after 80 min). The difference between "M[H]" and "M[D]_{Unfolded}" yields the maximum number of deuterons that can be incorporated into each rung, as determined by our methods and apparatus. All mass values are in Daltons (Da). Mass values were determined from the average of at least 7 different charge states in each mass spectrum, resulting in errors that are typically < 1 Da.

Supplemental Table 2. Kinetic parameters of hydrogen exchange data for the BCA II charge ladder generated from a least squares fit of the tri-exponential function: $M[D]_{\text{native}}$ $=$ M[D]_{unfolded} - Ae^{-k}₁^t - Be^{-k}₂^t - Ce^{-k}₃^t where M[D]_{folded} = mass of protein at each time point in deuterated buffer. $M[D]_{unfolded}$ = the mass of the perdeuterated protein (i.e. the mass of the unfolded protein in 90 % D₂O). The pre-exponential factors A, B, and C denote the number of amide hydrogens that are undergoing exchange at fast $(k_1 > 1 \text{ min}^{-1})$ ¹), intermediate (0.01 min⁻¹ \leq k₂ \leq 1 min⁻¹) and slow (k₃ \leq 0.01 min⁻¹) rates. Standard errors for all values are listed in parentheses.