Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Ion-Neutral Clustering Alters Gas-Phase Hydrogen-Deuterium Exchange Rates

Authors: Haley M. Schramm[1], Tomoya Tamadate[2], Christopher J. Hogan[2], Brian H. Clowers[1]

[1]Department of Chemistry, Washington State University, Pullman, WA 99163, USA

[2]Department of Mechanical Engineering, University of Minnesota, Minneapolis, MN 55455, USA

Abstract

The rates and mechanisms of chemical reactions that occur at the phase boundary often differ considerably from chemical behavior in bulk solution, but remain difficult to quantify. Ion-neutral interactions are one such class of chemical reactions whose behavior during the nascent stages of solvation differs from bulk solution while occupying critical roles in aerosol formation, atmospheric chemistry, and gas-phase ion separations. Using a gas-phase ion separation technique coupled with a counter-current flow of deuterated vapor, we quantify the degree of hydrogen-deuterium exchange (HDX) and ion-neutral clustering on a series of model chemical systems (i.e. amino acids). By simultaneously quantifying the degree of vapor association and HDX, the effects of cluster formation on reaction kinetics are realized. The results imply that cluster formation cannot be ignored when modeling complex nucleation processes of atmospheric aerosols and biopolymer structural dynamics.

Keywords: Molecular Solvation; Ion Chemistry; Hydrogen-Deuterium Exchange; Gas-Phase Ions; Ion-Molecular Reactions; Ion Mobility; Mass Spectrometry

Supporting Information

Table of Contents

Supporting Information

Data analysis. The data from the LTQ (.raw files) were converted to .mzML files using the tools within the Proteowizard framework.[1] All data analysis was performed using Python scripts developed in-house and a representative workflow is provided in the Supplementary Information (see **Figure S1**). The mzML files were used to extract ion chromatograms for a given mass window of interest that encodes the mobility information for that particular mass range. The ion chromatograms containing the frequency encoded mobility information were Fourier transformed and converted to drift time spectra.[2] The shifts in mobility for amino acids were assessed by calculating K_i/K_o ratio where K_i is the mobility of the amino acid with some concentration of modifier introduced into the drift gas and K_o is the mobility of the amino acid with no modifier.[3] A ratio of 1 is indicative of situations where the modifier does not interact with the target analyte.

Serving as a system suitability check, cocaine was included in all experimental runs. Across the data sets the reduced mobility of cocaine was continually assessed and found to be 1.16 ± 0.01 in agreement with literature values.[4] If the calculated mobility of cocaine for a given HDX experiment was out of this range, it was excluded from further analysis and the replicate was repeated. Because the functional groups hosting the charge-carrying proton possess comparatively high proton affinities and the resulting proton is further stabilized by an adjacent carbonyl group,[5] there is no observed ion-vapor clustering or HDX.

Including cocaine, all analytes were evaluated in their singly charged form $[M+H]^+$ following electrospray ionization. The *m/z* range encompassing the analyte's monoisotopic *m/z* plus the number of exchangeable hydrogens was used for evaluating HDX kinetics data where each isotopologue peak is resolved (see **Figure S2**). The intensities for each *m/z* in the isotope windows were extracted for each run with varying drift cell voltage and modifier flow rate.

To extract kinetic data from the isotope window, we utilized two different approaches. First, semilog plots were prepared to linearize the data and calculate rate constants in accordance with

pseudo-first order kinetics.[6–8] The linear form of the pseudo-first order kinetics integrated rate law is shown in Equation (1) where $[H_{ex}]$ is the number of remaining hydrogens with some concentration of modifier, $[CH_3OD]$ is the pressure of the modifier in torr, *k* is the rate coefficient, *t* is the total time the amino acid is exposed to the modifier and $[H_{ex}]_0$ is the number of remaining hydrogens with no modifier present.

$$ln[H_{ex}] = -[CH_3OD]kt + ln[H_{ex}]_0 \quad (1)$$

Average *m/z* values for each amino acid replicate with variable CH_3OD flow rates were calculated by a weighted average of the intensities and *m/z* values in the isotopic window. The number of remaining exchangeable hydrogens was calculated by subtracting the average *m/z* from the *m/z* if all theoretically exchangeable hydrogens were deuteriums. Modifier vapor pressures were calculated from the density and flow rate of the liquid modifier when mixed with the total gas flow rate using the ideal gas law. The introduction of all gasses was controlled by two mass flow controllers (main drift gas: MKS Model #M100B53CS1BV--S, modifier inlet: Cole Parmer Masterflex model #EW-32907-69). Deduction of the individual rates and associated reaction times for HDX was accomplished by extrapolating the time associated with the gas-phase ion mobility of the individual ions measured within the drift cell to include the total distance spanning the desolvation and drift regions shown in **Figure 2**. While there is a temporal component between droplet formation from an electrospray emitter and a bare ion entering in the gas phase, recent works by Mortenson and Williams have suggested this time is less than 300 μs using considerably lower flow rates and smaller emitters than the present work.[9] However, their work did not include the presence of a counter-current flow of gas. When factoring in the differences between experiment conditions, it is our estimation that the droplet to gas-phase ion evolution still remains well below 1 ms. It is for this reason that the total length of the drift cell (i.e. 231 mm) was used

for evaluating total reaction times. Inclusion of a small temporal adjustment to the total time does not meaningfully alter the resulting quantitative numbers and the subsequent interpretation.

Most importantly, the resulting semilog plots consistently show two linear regions with a slope equivalent to -*k* of the HDX reaction. A piecewise linear fit was done with the pwlf python package to deduce the respective slopes, intercepts, and breakpoints for each amino acid data set.[10] To aid in the interpretation of experimental results we also elected to fit rate constants using a solution to a system of differential equations. This exercise served to evaluate the theoretical boundaries for HDX under the experimental conditions.

Computationally Informed Exchange Rates. As an alternative form of analysis, we also fit results to a population balance model, i.e. we examine the concentration change of a species containing an isotopologue with *n* hydrogens (C_n) exchanged for deuterium. The change in this concentration with time is described by the equation $\frac{dC_n}{dt} = k_{n-1}C_{n-1}p_v - k_nC_np_v$ where *t* is time, k_n is the rate coefficient for $D_n \rightarrow D_{n+1}$, and p_v is the vapor pressure of the modifier gas. Assuming that k_n is not strongly sensitive to n (mainly to simplify inversion), a solution was determined for these equations for all observable isotopomers for each analyte using 4th order Runge-Kutta method. In solving these equations, we systematically varied a single $k(p_v)$ function to minimize the square error between measured isotopomer distributions and modeled distributions. More specifically, to infer the function $k(p_v)$, we varied $k(p_v)$ at each vapor pressure used in experiments and calculated time evolution of each isotopomer relative concentration C_n by the system of equations for ~27 ms (with 0.1 ms time steps), which is the residence time in experiments from the electrospray source and through the drift tube. Performing this operation for all experimental vapor pressures yielded $k(p_v)$ at all test vapor pressures, with linear interpolation used for values intermediate to experimental results. The tested range of the reaction constant was set as 10^{-2} to

10^1 times of k_0, where k_0 is a reaction constant calculated with the assumption under the reaction constant is not function of the vapor pressure (e.g., k_0 is 558 $Torr^{-1}s^{-1}$ for cysteine ions). While the rate coefficient may be a function of vapor pressure, the rate coefficient $k(p_v)$ should not differ from k_0 by multiple orders of magnitude. This is confirmed in fitting; we found in all instances that the squared error function for the difference between measurements and predictions is a convex downward curve with a local minimum value always within the same order of magnitude of k_0 (as shown in **Figure S8**). For this reason, the applied range in fitting of 10^{-2} to 10^1 times of k_0 is reasonable, and with minimal inputs, this approach provided a level of validation for not necessarily the absolute values observed for exchange rates, but rather, the overarching trends observed in the experimental data.

Fitting with the population balance model (see Figure S7) consistently suggests the exchange rate coefficient decreases with increasing vapor pressure, and there is a clear correlation between the shift in rate coefficient and shift in mobility for the five amino acid ions examined. Methionine's trends in mobility shifts and concentration dependent rate coefficients are nearly identical. Cysteine's rate coefficient interestingly increases slightly before it decreases, matching a phenomenon that was constant in the mobility shift data set. Glycine, serine, and tryptophan also show clear relationships between the concentration dependent nature of HDX reaction rate coefficient and shifts in mobility from transient ion-vapor clusters.

Estimating the Fraction of Time with Methanol Bound

Ion mobility theory shows that the length of the drift tube, L, can be related to the time an ion takes to traverse a drift tube, t_d, by the mobility coefficient, K, and the electric field, E.

$$L = KEt_d \tag{1}$$

If the total length is viewed as a summation of the ion with no vapor bound, one vapor, and so on ($L = L_0 + L_1 + L_2 + \dots$), then each of those environments has its own K and t_d but the same E.

$$KEt_d = K_0Et_0 + K_1Et_1 + K_2Et_2 + \dots = \Sigma_0^\infty K_iEt_i \text{ and so } t_d = \Sigma_0^\infty \frac{K_i}{K} t_i \quad (2)$$

It has been shown in previous publications that the probability of having greater than 2 molecules bound is negligible for these smaller ions and the vapor pressures examined; otherwise, we would observe substantially larger mobility shifts, such as those observed in Li and Hogan, 2019 (doi: 10.1080/02786826.2017.1288285) for salt cluster ions. Equation 2 can then be simplified.

$$t_d = \frac{K_0}{K} t_0 + \frac{K_1}{K} t_1 = t_0 + t_1 \quad (3)$$

Equation 3 can be rearranged and used to calculate the relative time spent in a cluster and as a bare ion.

$$\frac{t_1}{t_0} = \frac{\frac{K_0}{K} - 1}{1 - \frac{K_1}{K}} = \frac{K_0 - K}{K - K_1} \quad (4)$$

All the necessary variables are attained from the experiment: K_0, the mobilities with no modifier present, K, the mobilities with some flow rate of modifier, and K_1 which is estimated to be the maximum percent shift mobility data as shown in Table 2.

Figure S1: Screen shot of the example iPython notebook with the raw processing code. The full processing file and example data are included separately.

Figure S2. Schematics for “exchange competent” ion-neutral complexes for a generic amino acid and methanol-OD using two hydrogen bonds (depicted by dotted lines). The suggested structures

depict exchange using the A) amino- group and B) carboxylic acid where the deuterium on the methanol is labeled in red and the added proton from electrospray on the amino group is labeled in blue.

Analyte	Fit Predicted # of H_{ex}	Theoretical # of H_{ex}	Fast Rate (cm^3 $molecule^{-1}$ s^{-1})	Slow Rate (cm^3 $molecule^{-1}$ s^{-1})	k_{fast}/k_{slow}	Average % Cocaine K_i/K_0
Tryptophan	3.7	5	9.09E-12 (7%)	2.81E-12 (18%)	3.23	100 ± 1
Methionine	3.7	4	1.01E-11 (4%)	2.86E-12 (6%)	3.53	99.7 ± 0.6
Met ME	2.9	3	5.54E-12 (2%)	2.59E-12 (8%)	2.14	100.0 ± 0.4
Serine	4.6	5	9.37E-12 (5%)	4.89E-12 (9%)	1.92	100.0 ± 0.3
Cysteine	4.0	5	2.63E-11 (1%)	8.95E-12 (7%)	2.94	100.1 ± 0.3
Glycine	3.9	4	1.69E-11 (6%)	9.33E-12 (8%)	1.81	100.4 ± 0.6

Table S1: HDX kinetics results for select amino acids examined. The fit predicted number of exchangeable hydrogens is determined from the y-intercept of the fast exchange linear regression. For the fast and slow rates reported, the percentage in the parentheses is the percent error in the slope from the linear regression. The reported error is the standard deviation for the mobility shift. The maximum shift is the percent change in the mobility at the highest concentration of modifier in the drift cell. The average % cocaine K_i/K_0 is included and shows that there is no mobility shift for cocaine in the experiment

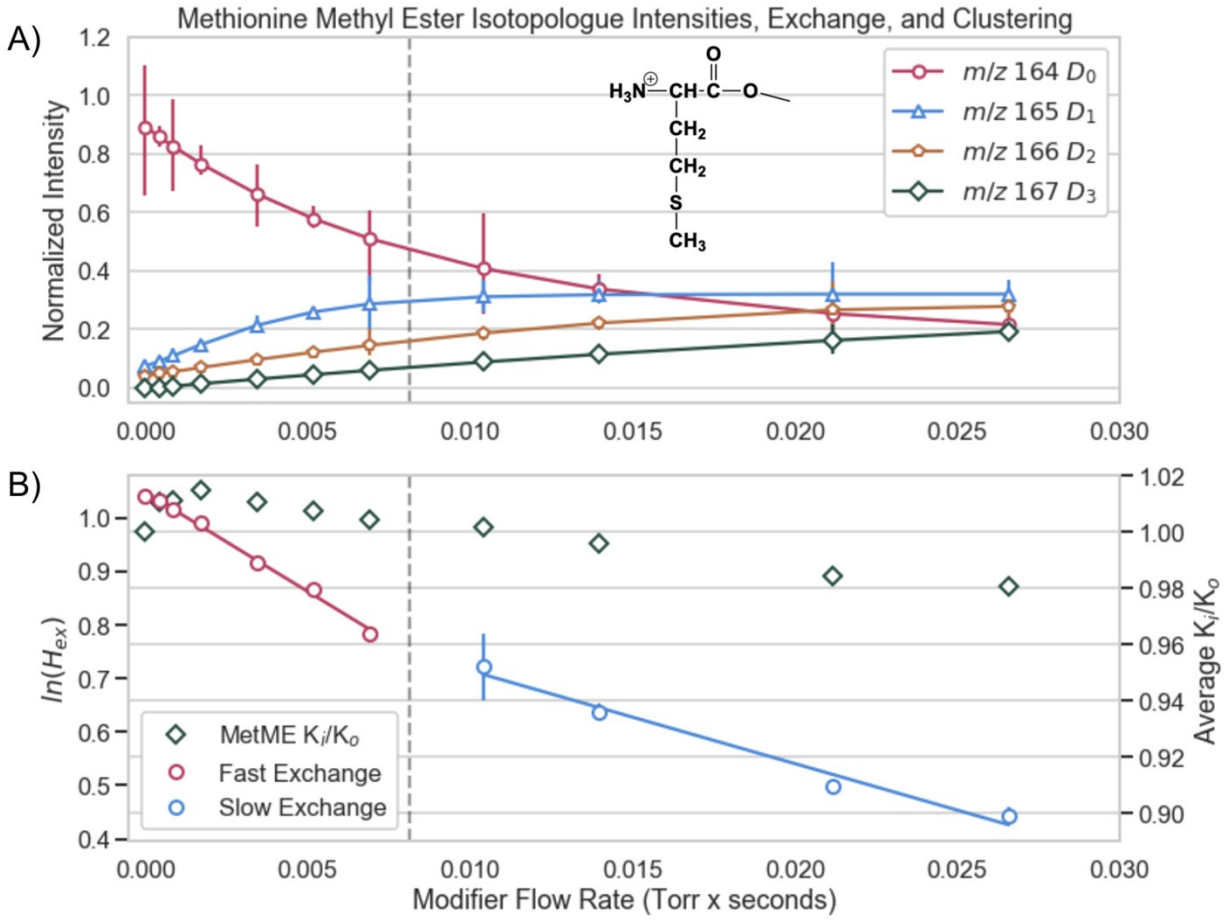

Figure S3: Methionine methyl ester results of the HDX and ion-neutral clustering experiment. The structure is the same as naturally occurring methionine except for the carboxylic acid is converted to a methyl ester so that the analyte has one fewer exchangeable hydrogen. All exchangeable hydrogens must be on the protonated amino group. While the relative rates are quite similar, the individual rates are slower. The carboxylic acid hydrogen does exchange faster, but it is not the source of nonlinearity.

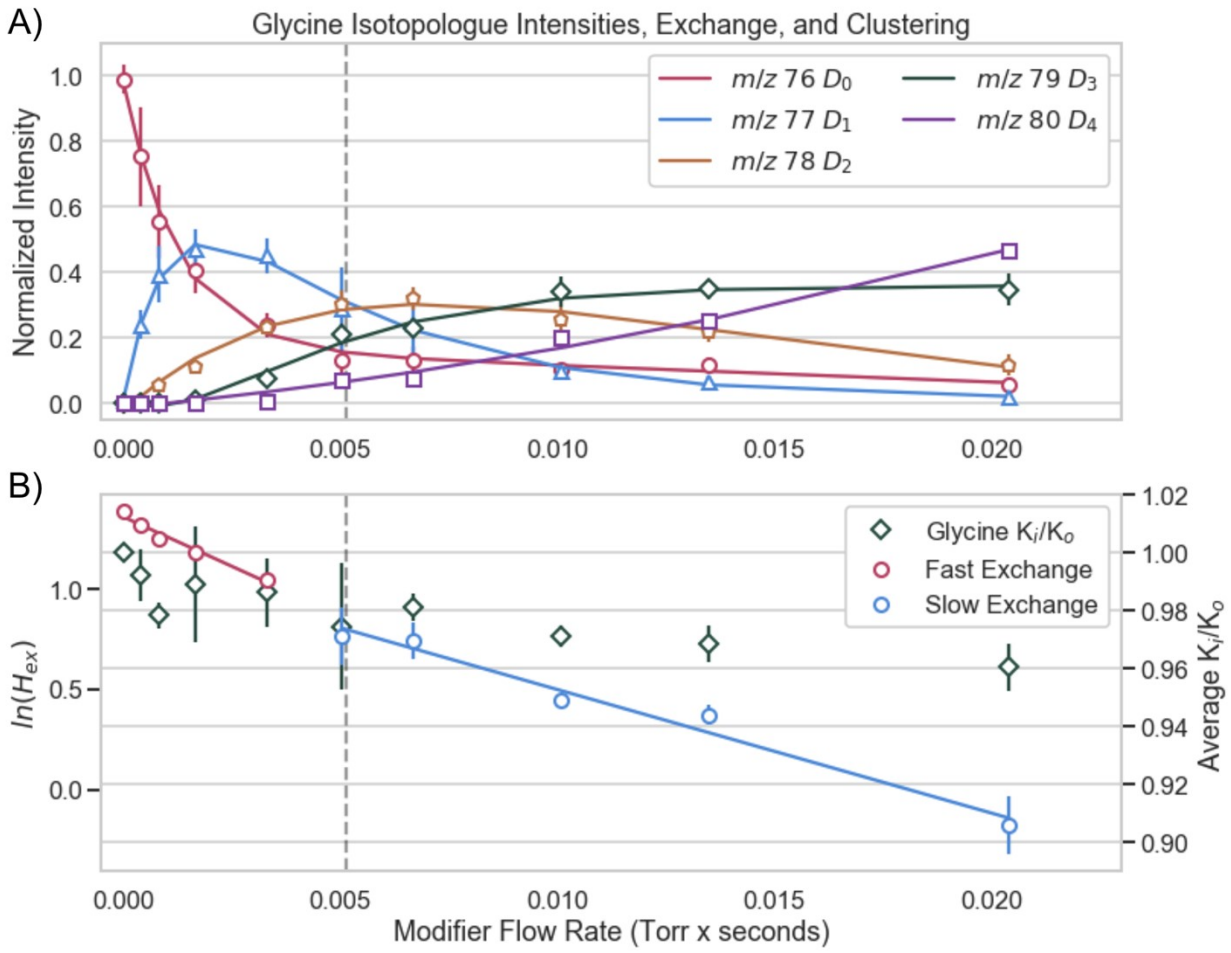

Figure S4: A) Glycine isotopologue intensity changes with a double exponential fit and B) semilog plot calculated from the average number of remaining hydrogens. Glycine does not show a clear breakpoint between two rates. The isotope ratios change rapidly for D_0 and D_1 with gradual increase rather than decrease for D_2 and D_3 while D_4 continues to climb to the most intense peak. While little clustering is observed, mobility shifts are observed after the breakpoint is calculated. The K_i/K_o then slowly declines to a minimum of 0.96, a 4% change in mobility.

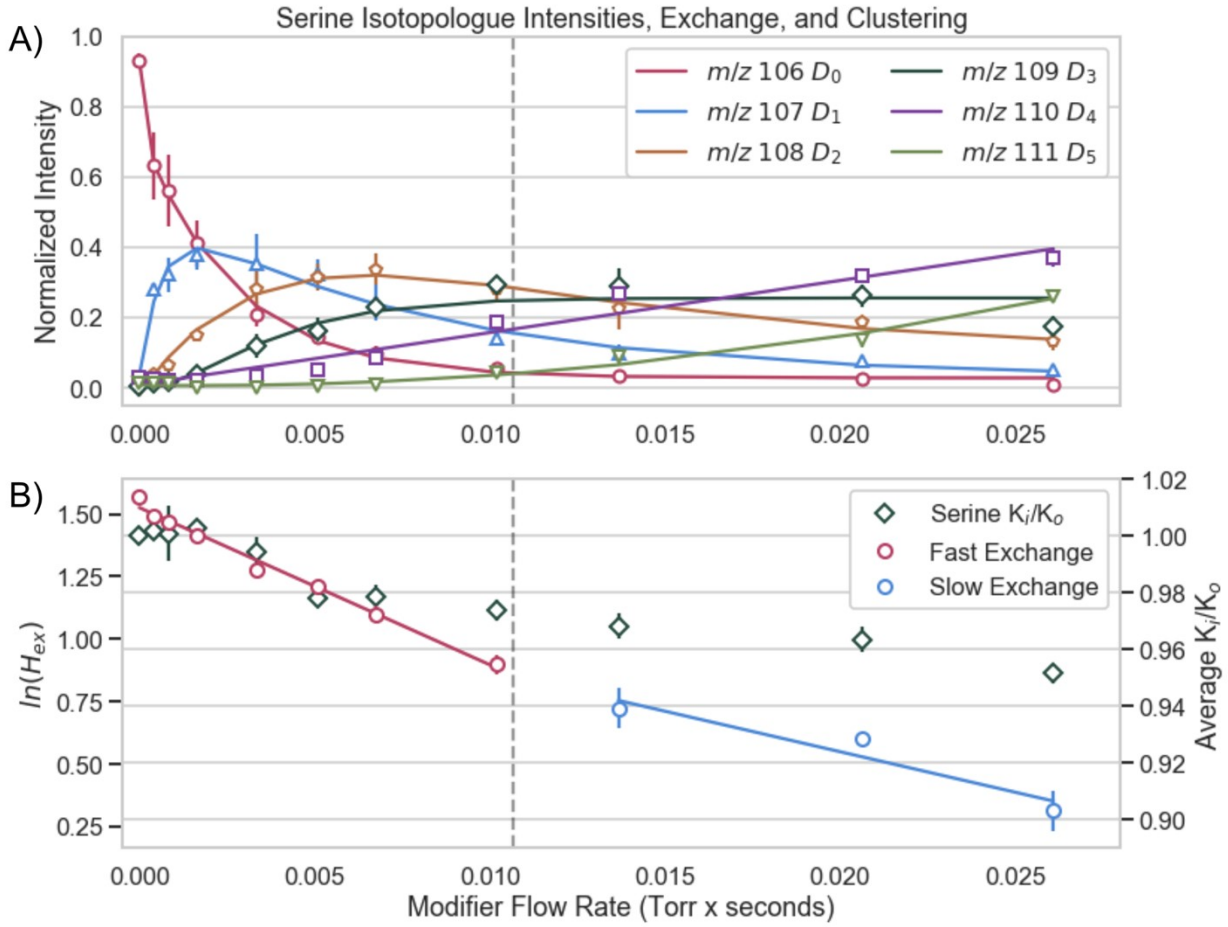

Figure S5: A) Serine isotopologue intensity changes with a double exponential fit and B) semilog plot calculated from the average number of remaining hydrogens. Serine is structurally the same as cysteine with a hydroxyl side chain rather than sulfhydryl and exchanges all five theoretically exchangeable hydrogens. This is explained by the capability of hydrogen bonding of the hydroxyl rather than a sulfhydryl group. The isotope change is similar to that of cysteine where the D_4 and here D_5 isotope increase steadily throughout the experiment while the others reach a peak intensity and then decrease. The clustering data show mobility shifts until 0.95 within the range of the experiment similar to cysteine and glycine.

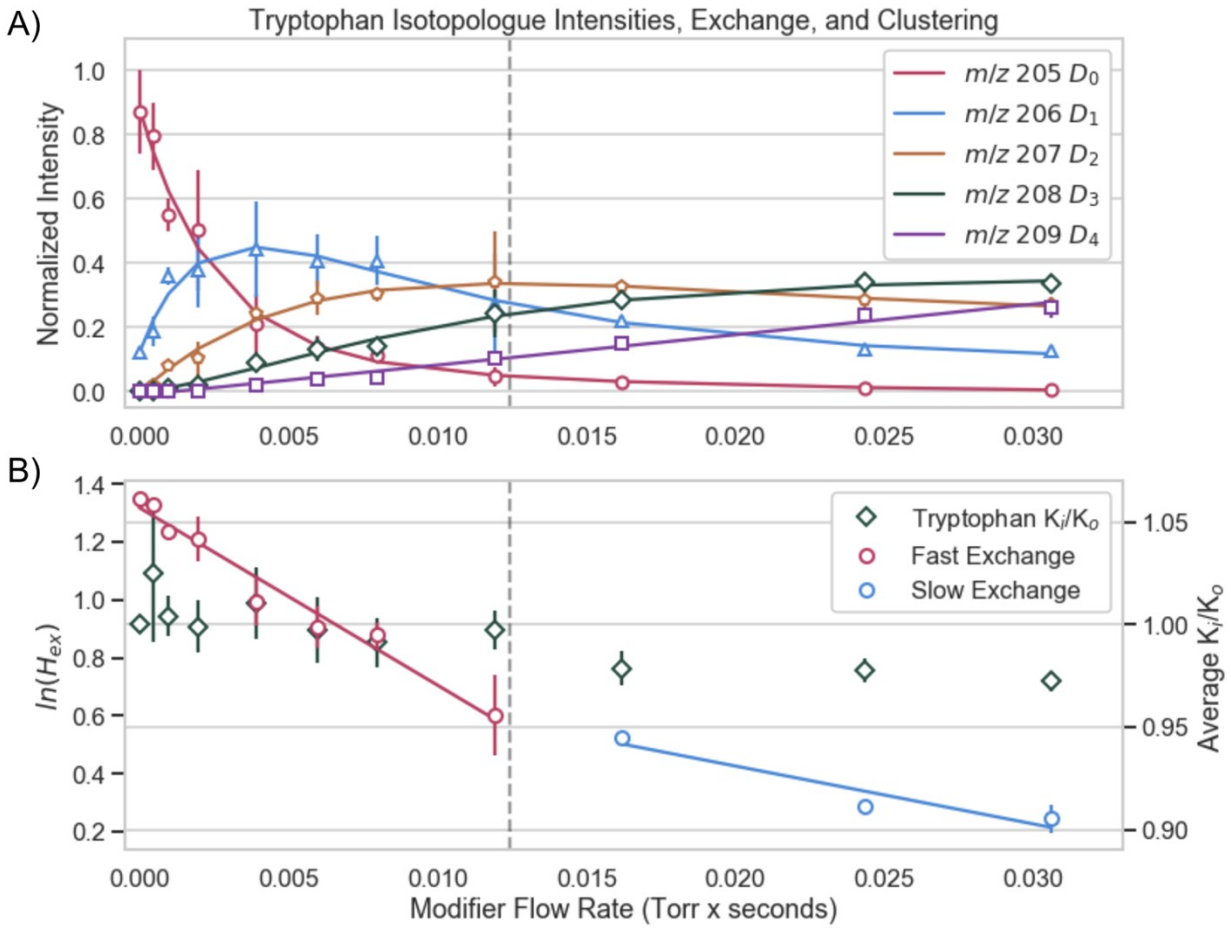

Figure S6: A) Tryptophan isotopologue intensity changes with a double exponential fit and B) semilog plot calculated from the average number of remaining hydrogens. The bulky, aromatic side chain of tryptophan with one possibly exchangeable hydrogen serves an interesting case to assess steric bulk on kinetic rates and clustering. Only four of the five theoretically exchangeable hydrogens are observed on the timescale of the experiment. The side chain indole group is either too far to form a stable hydrogen bond or is sterically blocked from exchange during the time scale of the ion mobility experiment. The isotopologues approach a steady state similar to what is observed with methionine, but the breakpoint is at much higher flow rates and very little clustering occurs. The mobility shift is not statistically different from one until after the breakpoint near 120 µL/hr and only changes about 3%.

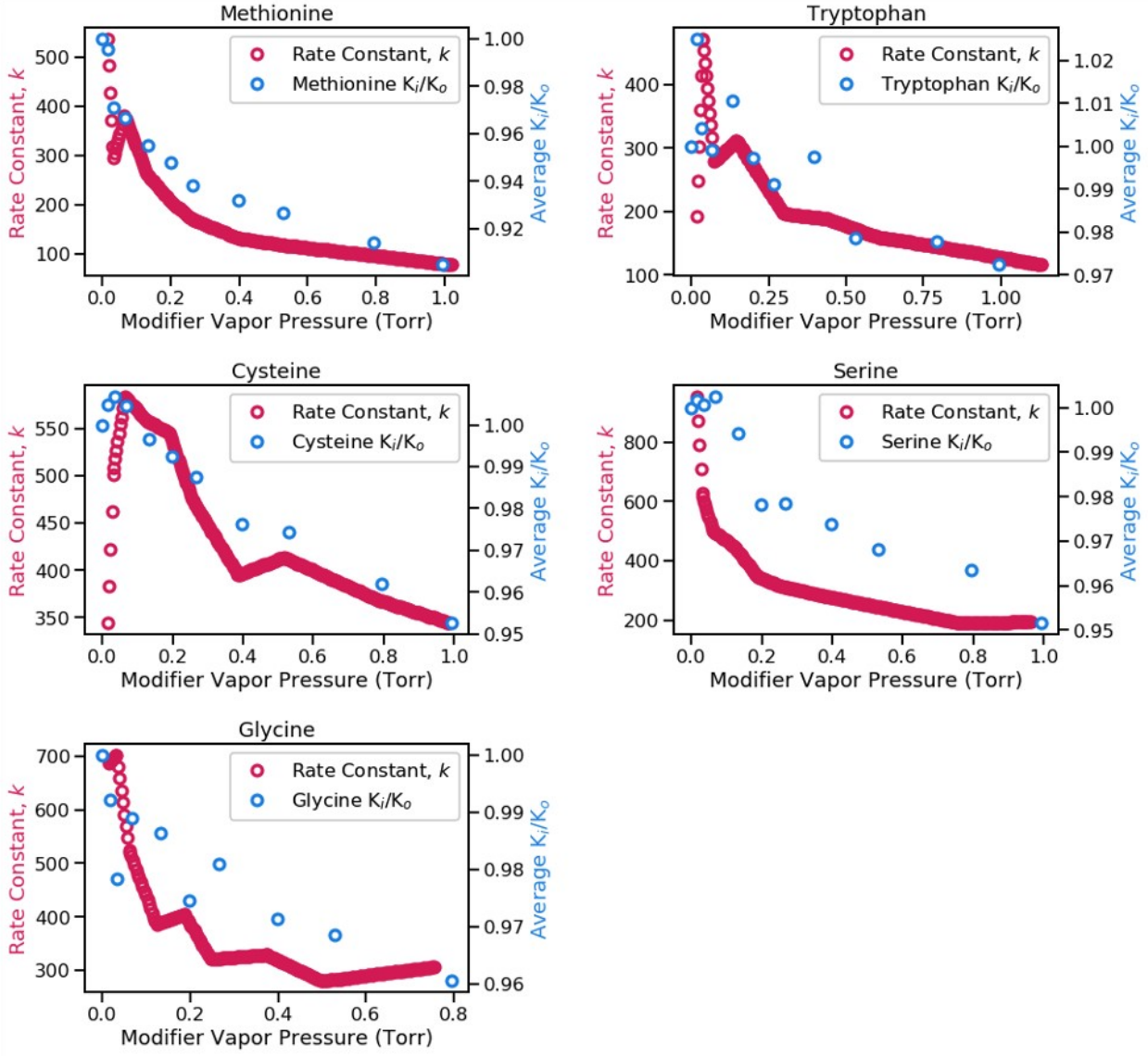

Figure S7: Average mobility shifts for each amino acid compared to the modeled rate constants, *k*. Inferred rate constants are in units of $Torr^{-1}$ s^{-1}.

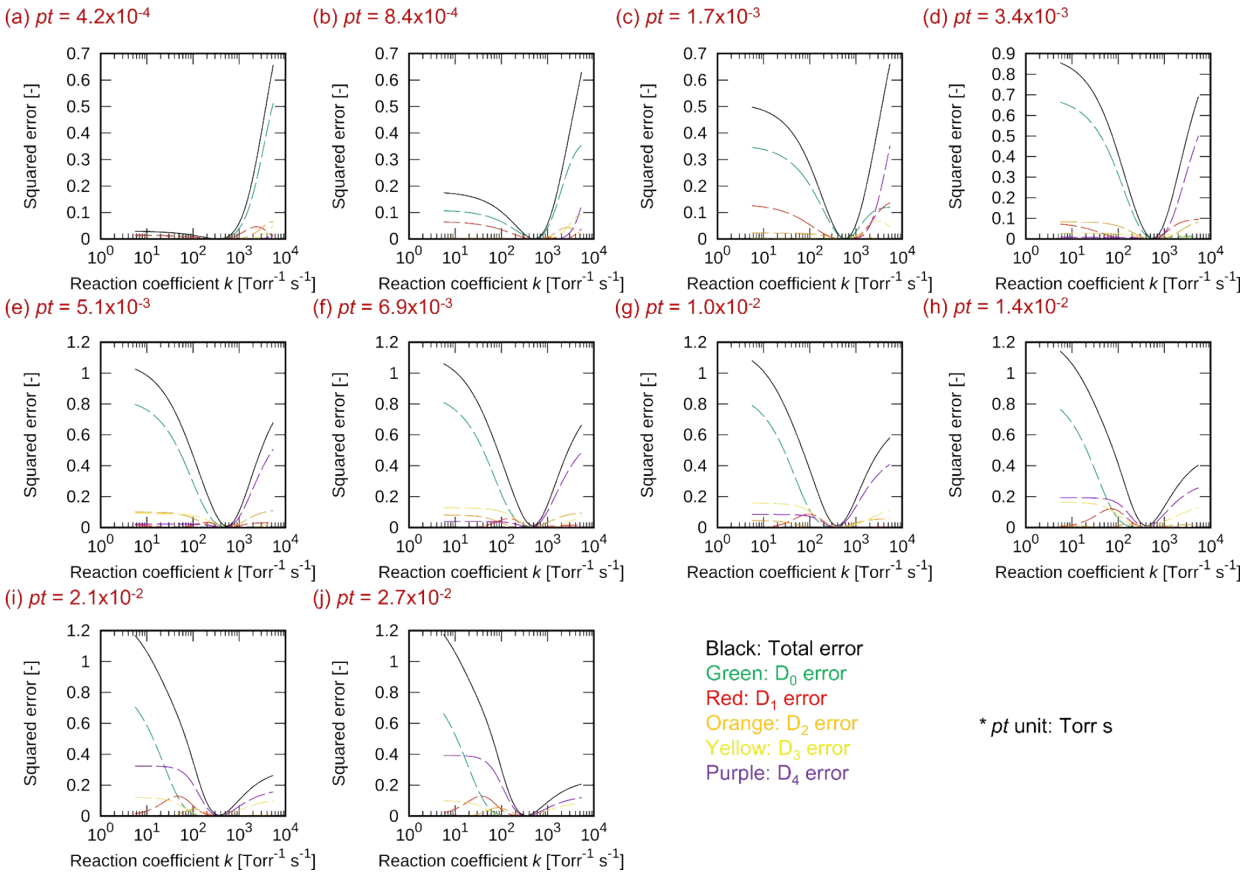

Figure S8: Plots of the squared error for variable trial reaction coefficients for cysteine ion HDX experiments for D_0-D_4 ions.

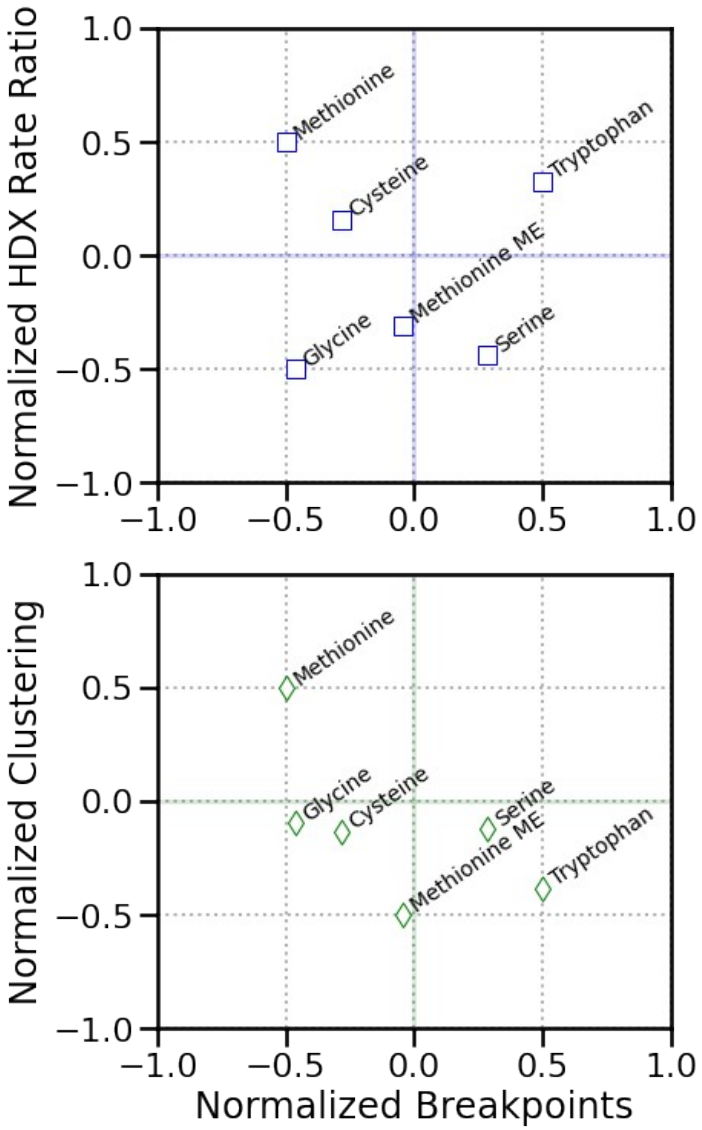

Figure S9: Composite data from respective tables (Tables 2 and S1) illustrating the range of behavior for the target analytes with respect to determined breakpoints between rates and the degree of ion neutral clustering observed. While the compound set probed in this effort is by no means exhaustive, it does suggest that multiple rates of HDX are elusive without ion-neutral clustering. The values shown for the breakpoints and degree of clustering were normalized to the range of values observed for the amino acids evaluated. This normalization was set to span a total value of 1 with a center point of 0.

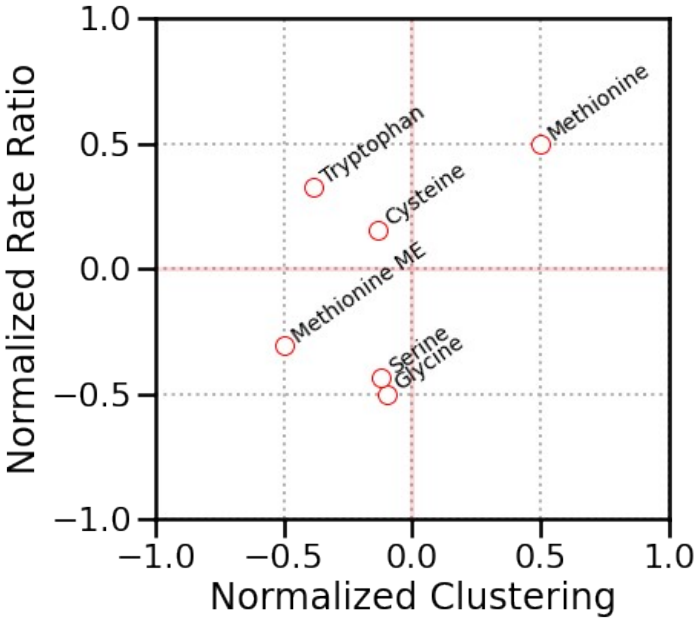

Figure S10: Composite data from respective tables (Tables 2 and S1) illustrating the range of behavior for the target analytes with respect to the degree of clustering and HDX observed. The HDX is expressed as a ratio between the fast and slower exchanging slopes shown in the summary plots (e.g. Figure S3). The values shown for the rate ratio and degree of clustering were normalized to the range of values observed for the amino acids probed. This normalization was set to span a total value of 1 with a center point of 0.

References

[1] M. C. Chambers, B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, K. Hoff, D. Kessner, N. Tasman, N. Shulman, B. Frewen, T. A. Baker, M.-Y. Brusniak, C. Paulse, D. Creasy, L. Flashner, K. Kani, C. Moulding, S. L. Seymour, L. M. Nuwaysir, B. Lefebvre, F. Kuhlmann, J. Roark, P. Rainer, S. Detlev, T. Hemenway, A. Huhmer, J. Langridge, B. Connolly, T. Chadick, K. Holly, J. Eckels, E. W. Deutsch, R. L. Moritz, J. E. Katz, D. B. Agus, M. MacCoss, D. L. Tabb, P. Mallick, *Nat. Biotechnol.* **2012**, *30*, 918–920.
[2] K. A. Morrison, W. F. Siems, B. H. Clowers, *Anal. Chem.* **2016**, *88*, 3121–3129.
[3] P. Kwantwi-Barima, C. J. Hogan Jr, B. H. Clowers, *J. Phys. Chem. A* **2019**, *123*, 2957–2965.
[4] T. P. Forbes, M. Najarro, *Analyst* **2016**, *141*, 4438–4446.
[5] K. A. Morrison, B. R. Valenzuela, E. H. Denis, M. K. Nims, D. A. Atkinson, B. H. Clowers, R. G. Ewing, *Analyst* **2020**, *145*, 6485–6492.
[6] S. J. Valentine, D. E. Clemmer, *J. Am. Chem. Soc.* **1997**, *119*, 3558–3566.
[7] D. S. Wagner, L. G. Melton, Y. Yan, B. W. Erickson, R. J. Anderegg, *Protein Sci.* **1994**, *3*, 1305–1314.
[8] H. A. Sawyer, J. T. Marini, E. G. Stone, B. T. Ruotolo, K. J. Gillig, D. H. Russell, *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 893–905.
[9] D. N. Mortensen, E. R. Williams, *Anal. Chem.* **2014**, *86*, 9315–9321.
[10] Jekel, Venter, *URL: https://github.com/cjekel* **n.d.**