Ligand Concentration Regulates the Pathways of Coupled Protein Folding and Binding

Kyle G. Daniels, Nam K. Tonthat, David R. McClure, Yu-Chu Chang, Xin Liu, Maria A. Schumacher, Carol A. Fierke, Scott C. Schmidler, and Terrence G. Oas

Supporting Information

Materials and Methods

Expression and Purification of P^{-Pro}

The gene for P protein variant P39A/P90A/F107W (P^{-Pro}) used in this study was constructed from the previously described gene, F107W variant, by the QuickChange site-directed mutagenesis protocol (Strategene). Both variants were overexpressed in *Escherichia coli* BL21 DE3 pLysS and purified as previously described [1] with the following modification. The pooled fractions of P protein that eluted from the second CM-Sepharose column were dialyzed into water and sodium triphosphate (Sigma-Aldrich) was added to a final concentration of 10 mM to precipitate P protein. P protein was pelleted by centrifugation at 10,000g for 30 minutes. Pellets were resuspended in 6M Guanidine-HCl and exchanged into Guanidine-HCl several times using Pierce Protein Concentrators with 9K MWCO. Protein was stored in 6M Guanidine-HCl at -80°C until being extensively dialyzed against 20 mM sodium cacodylate pH 7.0 for use. In all experiments, the protein concentration was determined using the Edelhoch[2] method using an extinction coefficient of 11460 M^{-1} cm^{-1} at 280 nm.

Crystallography

A 1 ml aliquot of RNase P protein (P^{-Pro}) in 6M Guanidine-HCl storage solution was dialyzed against water using a Pierce 3.5K MWCO Slide-Z-Lyzer dialysis cassette and concentrated using a 3K MWCO Amicon Ultra 0.5 mL centrifugal concentrator to 3.5 mg/mL in 10 mM Tris-HCl (pH 8.0) and mixed in a 1:1 ratio with crystallization solution (15% PEG 550 MME, 100 mM

MES pH 6.5, 27.5 mM $ZnSO_4$ and 5 mM sodium pyrophosphate). Diffracting crystals were formed by hanging-drop vapor diffusion within three days. Cryo-preservation of the crystal was achieved with the addition of 18% PEG 550 MME to the crystallization condition. The crystals formed in the $P6_4$ space group with cell dimensions: a=b= 83.14 Å and c=32.24 Å. All data were collected at The Advanced Photon Source (APS) at Argonne National Laboratory, beamline 22-BM (SER-CAT). Data were processed and scaled with HKL3000[3]. The structure was solved by molecular replacement using the *B*. *subtilis* RNase P protein structure as a starting model [4], and refined using Phenix[5]. Table S1 contains statistics for the structure. PDB ID: 4JG4.

Enzymology

Wildtype *B*. *subtilis* P RNA and a fluorescently labeled *B*. *subtilis* pre-tRNA[Asp] (Fl-pre-tRNA) were prepared as described in (Rueda, Hsieh et al. 2005)[6]. For single-turnover cleavage, 25 nM Fl-pre-tRNA and 500 nM RNase P holoenzyme (1:1 P protein to P RNA) were folded into 50 mM Tris/Mes (pH varies), ~200 mM KCl, 10 mM $CaCl_2$ and 20 mM DTT. Multiple-turnover reactions contains 1 nM P RNA and 4 nM P protein in 50 mM Tris-HCl (pH 8), 100 mM KCl, 10 mM $MgCl_2$, 20 mM DTT and 0.01% NP-40. Cleavage reactions were monitored by fluorescence polarization signal change using TECAN Infinite F500 plate reader [7].

Stopped-Flow Fluorescence

Stopped-flow fluorescence kinetic experiments were performed on an Applied Photophysics SX20 instrument at 25°C. Samples were mixed in an observation cell with a 2 mm pathlength and excited at 285 nm with a slit width of 1mm. Emitted light was detected through a 320 nm high-pass filter. Experiments were performed by mixing protein with PPi or TMAO solutions at a ratio of 1:5. The initial protein concentration was 15 μM and the final protein concentration was 2.5 μM. Final PPi concentrations ranged from 0 to 166 μM and final TMAO concentrations

ranged from 0 M to 0.93 M. Kinetic traces were collected over 30 seconds with logarithmic time sampling. At least three traces were collected and averaged for each concentration of TMAO or PPi.

Analysis of Kinetic Data

The model used to analyze stopped-flow data assumes that each of the three conformational states of P^{-Pro} contains up to two high affinity binding sites (Figure 3 for scheme) and that any conformational sub-ensemble can interconvert with any other sub-ensemble that has the same pattern of occupied ligand binding sites. Rate constants for folding and unfolding are assumed to be dependent on the microscopic liganding state of the molecule and the concentration of TMAO. Following standard practice, we assume that the logarithms of the microscopic rate constants are linearly dependent on co-solute (TMAO) concentration[8]. We expressed many of the rate and equilibrium constants in terms of other rate and equilibrium constants in order to fit for as parameters as possible and satisfy the principle of detailed balance. Parameters expressed in terms of other parameters can be found in the Equations section. The kinetics of ligand association and dissociation are too fast to be observed by stopped flow. In the analysis presented here, we have assumed that the second order association rate constant k_{on} is diffusion limited and set it to 10^8 $M^{-1}s^{-1}$. This assumption is consistent with our observation of a burst phase increase in protein fluorescence upon ligand addition, which requires that k_{on} be at least 10^7 $M^{-1}s^{-1}$. The model is insensitive to changes in k_{on} between 10^7 and 10^{10} $M^{-1}s^{-1}$. The dissociation rate constant k_{off} for any individual binding reaction changes with k_{on} because it is the product of the fixed k_{on} and the fitted parameter for the association constant K_A.

Drift in fluorescence signal between collection of the TMAO-induced and PPi-induced folding data was accounted for by addition of an offset parameter to the model.

Algorithm for Parameter Estimation

Fitted parameters were obtained by Bayesian estimation under uniform prior distributions over plausible parameter ranges. Posterior means and 95% credible intervals were computed by a sequential Monte Carlo technique for static data, as follows:

Let Y denote data and θ a vector representing the d parameters of the model. Independent uniform prior distributions were assigned over the allowable ranges (a_i, b_i) of each parameter θ_i. Experimental measurement noise was modeled as Gaussian with unknown precision τ, itself assigned a standard conjugate gamma prior distribution. The resulting posterior distribution over parameters obtained by conditioning on the observed data takes the form:

$$\pi(\theta \mid Y) \propto \left[\prod_{j \in Y} N(y_j; g(\theta, t_j), \tau) \right] \left[\prod_{i=1}^{d} (b_i - a_i)^{-1} \right] \text{Gamma}(\tau; \alpha, \beta)$$

where $g(\theta, t_i)$ denotes fluorescence at time t_i calculated under the model with parameters θ, and $\alpha = 0.01$, $\beta = 0.0001$ give a non-informative prior for τ. Inference about parameters was made by a "static-data" sequential Monte Carlo technique[9] for sampling from $\pi(\theta \mid Y)$ via construction of a sequence of posteriors with increasing amounts of data, as follows:

1. The first, middle, and last time points for all ligand and osmolyte curves (Y_0) were used to obtain the initial posterior $p(\theta|Y_0)$.

2. A standard Metropolis MCMC algorithm[10] was run for 50k iterations to obtain samples from $\pi(\theta|Y_0)$. An initial "burn-in" period of 10k was discarded, after which the samples were thinned to obtain 1000 to be used as initial particles.

3. The remaining data were randomly divided into sets $\{Y_1, \ldots, Y_p\}$ of size 10 which were used to create the sequence of posteriors $\pi_k(\theta) = \pi(\theta \mid Y_0, \ldots, Y_k)$ for $k = 1, \ldots, p$.

4. A standard sequential Monte Carlo algorithm[11] with importance resampling was used to propagate the particles through the sequence of distributions $\pi_1, \ldots, \pi_p$.

5. The resulting samples from $\pi_p = \pi(\theta \mid Y)$ were used to obtain posterior means and 95% credible intervals.

A small number of nearby maxima were identified, whose differences were statistically but not practically significant; we report only one.

Previously Determined Values

A previous study examined TMAO-induced folding of the *B. subtilis* RNase P protein F107W variant[12]. The current work uses the F107W/P39A/P90A variant (P^{-Pro}) to simplify the folding mechanism. Estimates for several parameters (ΔG_{UI}, ΔG_{IF}, β_{UI}, β_{IF}, k_{UI}, k_{IU}, k_{IF}, k_{FI}, m_{UI}, m_{IF}) differ between the two studies. In the earlier study on F107W, kinetic data were fit to sums of exponentials and the obtained amplitudes and k_{obs} values were fit to a linear folding mechanism. The amplitudes of the F107W kinetic data were poorly fit. Kinetic data for P^{-Pro} were fit directly to a triangular folding mechanism and the fit is better than that of the F107W data. Differences in the rate constants and free energies obtained from the fits are probably primarily the result of the differences in the quality of the fits. Additional discrepancies may arise from the proline-to-alanine substitutions.

Flux Calculation

Equilibrium flux for each reaction in the reaction scheme was calculated by multiplying the concentrations of products by the forward rate constants. Equilibrium flux for the 18 pathways of interconversion between U and FL_2 was calculated using the equation $F_{path}=(\Sigma\ 1/F_i)^{-1}$, where F_{path} is the flux of the pathway and F_i is the flux of the *i*th reaction in the pathway.

Isothermal Titration Calorimetry

Binding of PPi to the N-terminal peptide mimic AHLKKRNRLKKNEDW was measured by ITC. The synthetic peptide AHLKKRNRLKKNEDW was obtained from GenScript and purified by cation exchange chromatography and desalted by size exclusion chromatography. Crude peptide was resuspended in 100 mM NaCl, 5 mM EDTA, 50 mM sodium acetate pH 6.5 and loaded onto a CM Sepharose column. Peptide was eluted with a salt gradient of 100-800 mM NaCl, 5 mM EDTA, 50 mM sodium acetate pH 6.5. Fractions containing the peptide were

pooled and the volume was reduced to 0.5 mL by SpeedVac. The 0.5 mL was loaded onto a

Sephadex G-50 column and eluted with water. Eluted peptide was added to an equal volume of

40 mM sodium cacodylate pH 7.0 and frozen at -20°C until use. 170 μM peptide was titrated

with PPi to a final PPi concentration of 3.5 mM. Binding of PPi to a mixture of P^{-Pro} fragments

produced by formic acid cleavage of P^{-Pro} was measured by ITC. Formic acid cleavage was

performed by incubating 25 μM P^{-Pro} in 2% formic acid for 2 hours in a boiling water bath. The

presence of the predicted fragments (AHLKKRNRLKKNED,

FQKVFKHGTSVANRQFVLYTLD, QAEND,

ELRVGLSVSKKIGNAVMRNRIKRLIRQAFLEEKERLKEKD, and

YIIIARKAASQLTYEETKKSLQHLWRKSSLYKKSSSK) was confirmed by MALDI-ToF

mass spectrometry. P^{-Pro} fragments were recovered by lyophilizing the reaction mixture and

resuspending in 20 mM sodium cacodylate pH 7.0. 20 μM peptide mixture was titrated with PPi

to a final PPi concentration of 4.3 mM. Titrations were performed in 20 mM sodium cacodylate

pH 7.0 at 25°C. Red lines in the Figure S4A and S4B are the best-fit of the data to a single site

binding isotherm. 200 μM P^{-Pro} in 20 mM sodium cacodylate pH 7.0 at 25°C was titrated with

PPi to a final PPi concentration of 740 μM. The data were fit to a three-state, three-site coupled

folding and binding model using parameters obtained from the fit of the kinetic data.
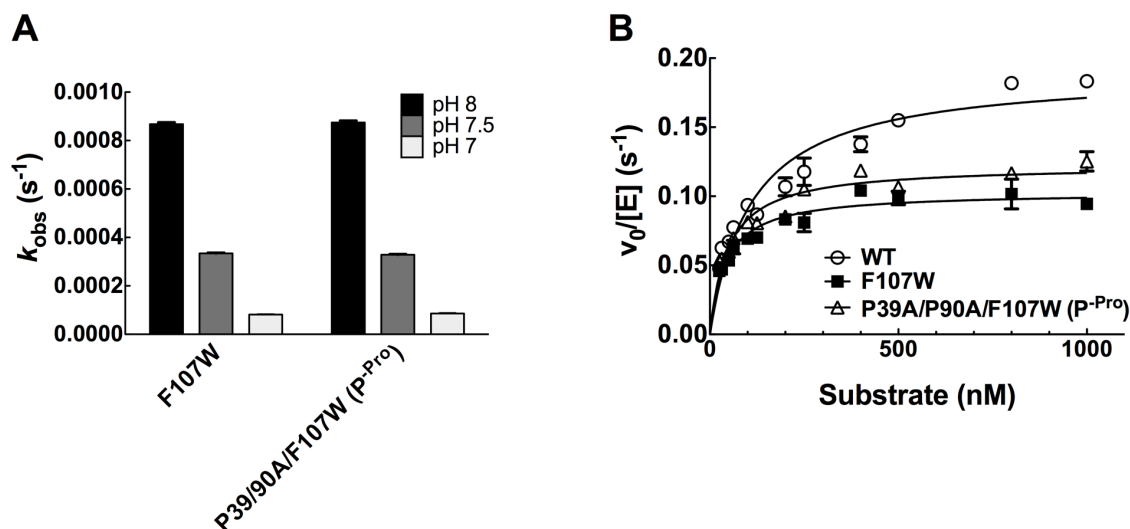
Figure S1. Single and multiple turnover cleavage of fluorescent pre-tRNA by RNase P. Structural integrity of P protein variants was assessed by their ability to form enzymatically active RNase P holoenzyme as described in Materials and Methods. **(A)** Under single-turnover conditions, mutations in P protein (F107W or P^{-Pro}) do not affect the pH-dependent rate for RNase P holoenzyme-catalyzed cleavage of fluorescently labeled pre-tRNA^{Asp}. **(B)** RNase P holoenzyme formed by P RNA and either wildtype (WT), F107W, or P^{-Pro} proteins cleave fluorescently labeled pre-tRNA^{Asp} under multiple turnover conditions. Data were fit to the Michaelis-Menten equation to calculate the following steady-state kinetic parameters: wildtype RNase P holoenzyme, $k_{cat} = 0.19 \pm 0.01$ s^{-1}, $K_M = 110 \pm 16$ nM, and $k_{cat}/K_M = 1700 \pm 200$ mM^{-1}s^{-1}; F107W P protein, $k_{cat} = 0.103 \pm 0.003$ s^{-1}, $K_M = 41 \pm 5$ nM, and $k_{cat}/K_M = 2500 \pm 300$ mM^{-1}s^{-1}; and P^{-Pro} P protein, $k_{cat} = 0.122 \pm 0.003$ s^{-1}, $K_M = 47 \pm 5$ nM, and $k_{cat}/K_M = 2600 \pm 300$ mM^{-1}s^{-1}.
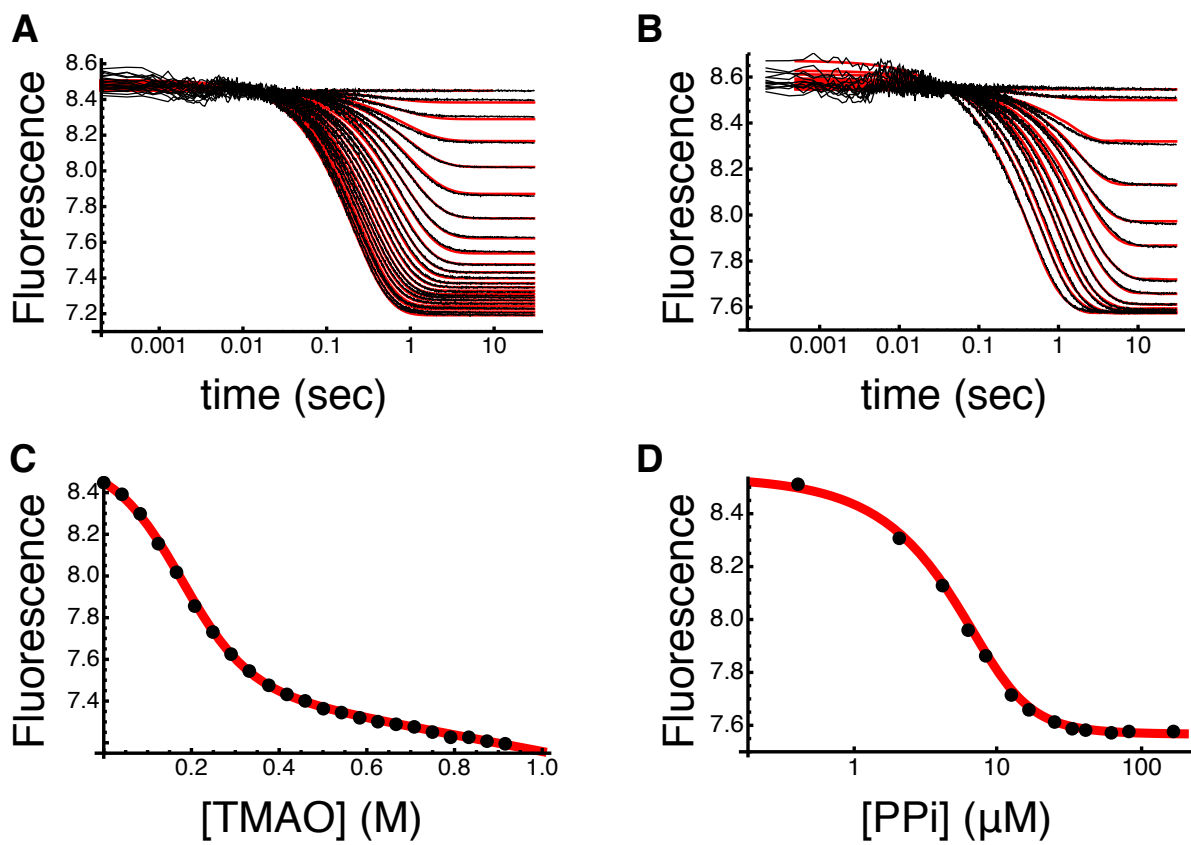
Figure S2. TMAO- and PPi-induced P$^{\text{-Pro}}$ folding were monitored by stopped-flow fluorescence. Two-dimensional projections kinetic traces for TMAO-induced (**A**) and pyrophosphate-induced (**B**) folding of P$^{\text{-Pro}}$ as monitored by stopped-flow fluorescence. Data points are the average of data points from three traces. The red lines are the global best-fit of the data to the coupled folding and binding model. Data points for the equilibrium plots for TMAO-induced (**C**) and pyrophosphate-induced (**D**) folding are the final time points (30 seconds) of each trace in the stopped-flow experiments. The red lines are the ideal equilibrium plots generated using the parameter estimates obtained from fitting the stopped-flow data.
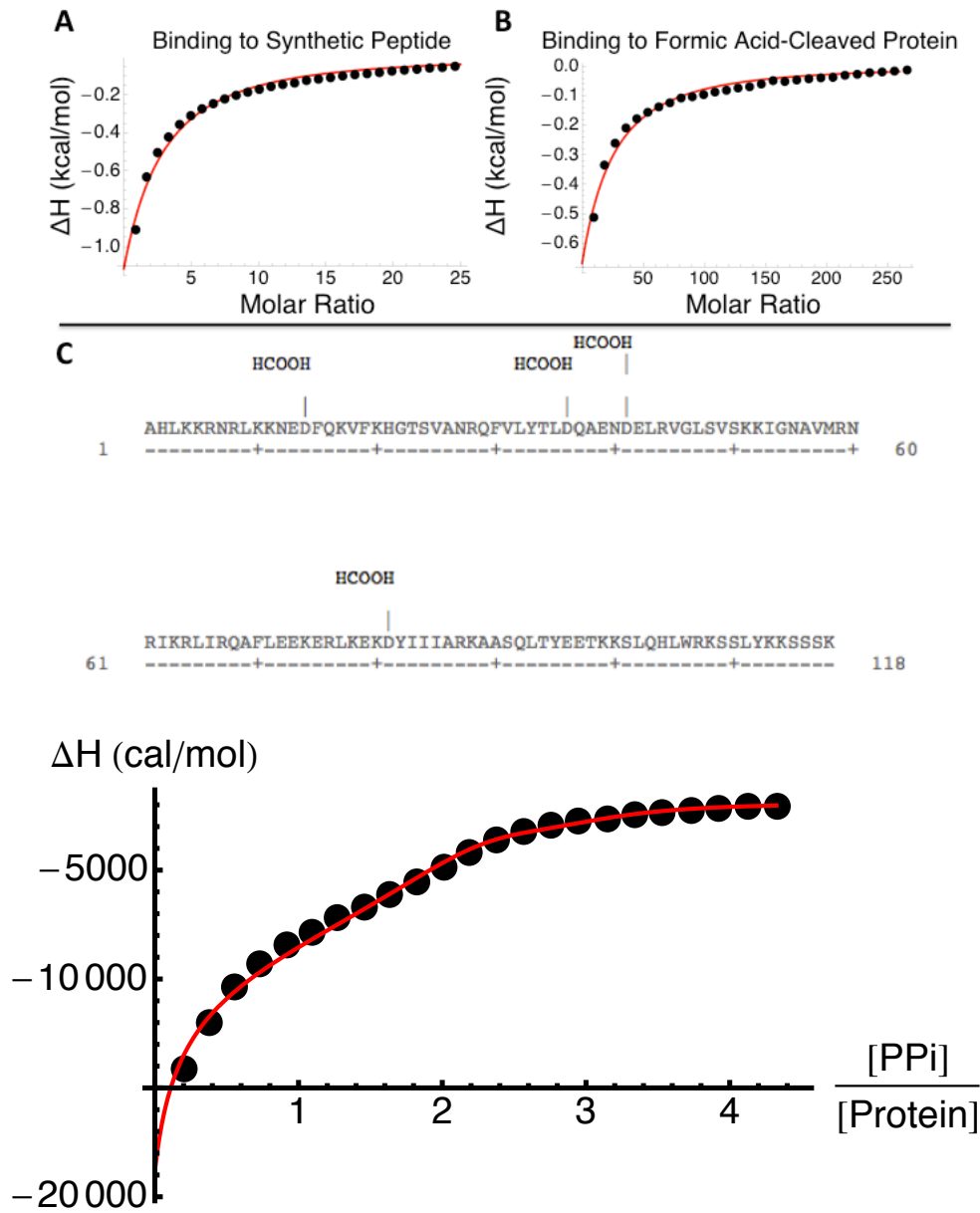
A Binding to Synthetic Peptide

B Binding to Formic Acid-Cleaved Protein

C

```
                    HCOOH
       HCOOH        HCOOH        |
         |            |          |
   AHLKKRNRLKKNEDFQKVFKHGTSVANRQFVLYTLDQAENDELRVGLSVSKKIGNAVMRN
 1 ---------+---------+---------+---------+---------+---------+  60
```

```
               HCOOH
                 |
   RIKRLIRQAFLEEKERLKEKDYIIIARKAASQLTYEETKKSLQHLWRKSSLYKKSSSK
61 ---------+---------+---------+---------+---------+--------  118
```

Figure S3. Binding of pyrophosphate to P[-Pro] and mimics of the P[-Pro] N-terminus was measured by isothermal titration calorimetry. (**A**) Binding of pyrophosphate to the N-terminal peptide mimic AHLKKRNRLKKNEDW was measured by ITC. A sample of 170 μM peptide was titrated with pyrophosphate to a final pyrophosphate concentration of 3.5 mM. (**B**) Binding of pyrophosphate to a mixture of P[-Pro] fragments produced by formic acid cleavage [13] of P[-Pro] was measured by ITC. A sample of 20 μM peptide mixture was titrated with pyrophosphate to a final

pyrophosphate concentration of 4.3 mM. Titrations were performed in 20 mM sodium cacodylate

pH 7.0 at 25°C. Red lines are the best-fit of the data to a single site binding isotherm. (**C**) Formic

acid cleaves to the C-terminal side of aspartate residues. Cleavage sites are indicated by HCOOH

markers. Formic acid cleavage was performed as described in materials and methods. (**D**) A

sample of 200 μM P^{-Pro} in 20 mM sodium cacodylate pH 7.0 at 25°C was titrated with

pyrophosphate to a final pyrophosphate concentration of 740 μM. The data were fit to a three-

state, three-site coupled folding and binding model using parameters obtained from the fit of the
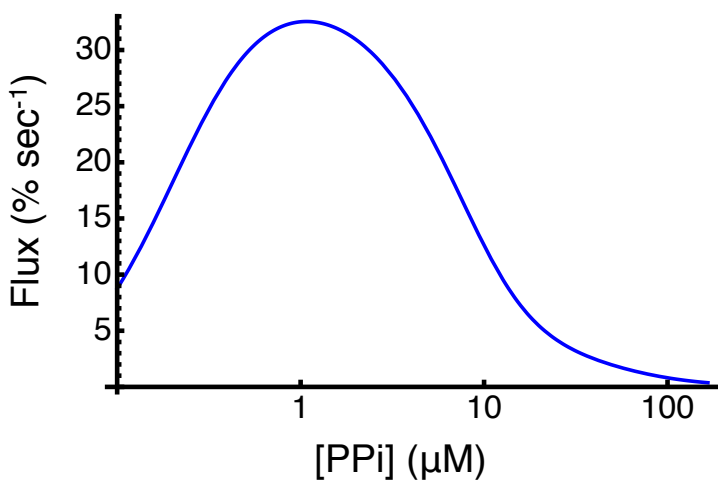
kinetic data.

Figure S4. Total equilibrium flux for the U ⇔ FL$_2$ reaction vs. ligand concentration calculated from best-fit parameter values of the model describe in the text. Peak flux occurs at the apparent midpoint of the binding curve, where the populations of U and FL$_2$ are the same and falls off substantially at lower and higher PPi concentrations. Flux is defined as the percent of the P[-Pro] interconverting between U and FL$_2$ per second.

Table S1. Crystallography statistics for P^{-Pro} in complex with pyrophosphate.

	P^{-Pro} in complex with Pyrophosphate
Space group	P 6$_4$
a, b, c (Å)	83.14, 83.14, 32.24
α, β, γ (°)	90, 90, 120
R$_{sym}$ (%)^a	4.8 (12.6)*
<I>/<σI>	44.3 (11.9)
Completeness (%)	94.4 (78.6)
Multiplicity	8.9 (7.6)
Total reflections (#)	49240
Unique reflections (#)	5545
Resolution (Å)	24.02 (2.30)
R$_{work}$/R$_{free}$ (%)^b	21.26/24.92
Ramachandran statistics	
Favored	116 (98.2%)
Allowed	2 (1.8%)
Outlier	0 (0%)
Rmsd	
Bond lengths (Å)	0.010
Bond angles (°)	1.488

*Values in parentheses are for highest-resolution shell.

$^a R_{sym} = \sum\sum|Ihkl - Ihkl(j)|/\sum Ihkl$, where Ihkl(j) is observed intensity and Ihkl is the final average value of intensity.

$^b R_{work} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$ and $R_{free} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$, where all reflections belong to a test set of 10% data randomly selected in PHENIX.

Table S2. Dissociation constants for P^{-Pro} and pyrophosphate.

Conformational State	K_D (M)		
	α site[a]	β site[a]	γ site[b]
U	**2.6** (2.3, 2.7) $\times 10^{-3}$	ND[c]	ND[c]
I	**1.2** (1.1, 1.3) $\times 10^{-4}$	**4.6** (3.7, 5.3) $\times 10^{-4}$	ND[c]
F	**7.6** (7.3, 7.9) $\times 10^{-7}$	**2.3** (2.2, 2.4) $\times 10^{-6}$	**4.0** (3.6, 4.4) $\times 10^{-5}$

[a]Estimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of stopped-flow data as described in Materials and Methods.

[b]Estimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of isothermal titration calorimetry data as described in Materials and Methods.

[c]Value not determined because no binding was detected.

Table S3. Conformational parameter value estimates from fit of stopped-flow data.

Parameter	(cal mol^{-1})[a]	Parameter	(cal mol^{-1} M^{-1})[a]	Parameter	(unitless)[a]
ΔG_{UI}°	**859** (850, 864)	mui	**-984** (-996, -965)	β_{UI}	**0.41** (0.39, 0.42)
ΔG_{IF}°	**372** (366, 382)	mif	**-6960** (-7000, -6930)	β_{IF}	**0.566** (0.563, 0.570)
				β_{UF}	**0.88** (0.86, 0.93)

[a]Estimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of stopped-flow data as described in Materials and Methods.

Table S4. Rate constant parameter value estimates from fit of stopped-flow data.

Parameter	(s^{-1})[a]
k_{UI}	**1.99** (1.97, 2.00)
k_{IU}[b]	**8.5** (8.4, 8.6)
$k_{UI}\alpha$	**2.4** (2.0, 3.3)
$k_{IU}\alpha$[b]	**0.5** (0.4, 0.7)
k_{IF}	**1.06** (1.03, 1.09)
k_{FI}[b]	**2.00** (1.95, 2.05)
$k_{IF}\alpha$	**7** (6, 8)
$k_{FI}\alpha$[b]	**0.09** (0.06, 0.1)
$k_{IF}\beta$	**22** (19, 25)
$k_{FI}\beta$[b]	**0.21** (0.18, 0.28)
$k_{IF}\alpha\beta$	**120** (100, 140)
$k_{FI}\alpha\beta$[b]	**0.0072** (0.0068, 0.0078)
k_{UF}	**0.036** (0.031, 0.043)
k_{FU}[b]	**0.29** (0.25, 0.35)
$k_{UF}\alpha$	**9** (8, 10)
$k_{FU}\alpha$[b]	**0.021** (0.019, 0.024)

[a]Estimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of

stopped-flow data as described in Materials and Methods.

Table S5. Signal parameter value estimates from fit of stopped-flow data.

Parameter	(AFU^{b}) [a]	Parameter	$(AFU\ M^{-1\,b})$ [a]
$Signal_U$	**8.758** (8.756, 8.761)	$Slope_U$	**0.255** (0.252, 0.259)
$Signal_I$	**7.60** (7.59, 7.62)	$Slope_I$	**-0.498** (-0.499, -0.495)
$Signal_F$	**7.555** (7.553, 7.556)	$Slope_F$	**-0.397** (-0.399, -0.394)
$Signal_{UBound}$	**9.5** (9.3, 9.6)		
$Signal_{IBound}$	**8.44** (8.40, 8.48)		
$Signal_{Free}$	**7.474** (7.473, 7.475)		

Table S6. Dissociation constants for the P[-Pro] unfolded state and pyrophosphate.

Protein Form	K_D (mM)
Unfolded Protein[a]	**2.6** (2.3, 2.7)
N-terminal Synthetic Peptide[b]	**1.0** (0.8, 1.2)
Peptide Mixture[b]	**1.0** (0.8, 1.2)

[a]Estimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of stopped-flow data as described in Materials and Methods.

[b]Estimate (bold) and the 95% confidence interval (parentheses) were obtained from fit of isothermal titration calorimetry data as described in Materials and Methods.

Movie S1. The mechanism of coupled folding and binding is concentration dependent. The mechanism of interconversion between U and FL$_2$ was assessed by calculating the fractional flux through each of 18 pathways. Species populations (spheres) and fractional fluxes (path lines) were calculated using parameter values derived from the global best-fit of TMAO- and PPi-induced folding stopped-flow data. The populations and fluxes were calculated using 2.5 µM total protein concentration and the indicated total pyrophosphate concentrations from 0.03 to 980 µM.

Equations used for fitting the data to the model

General Notes :

Upper case K (Kui, Kif, Kuf, and so on) denotes an equilibrium constant

Lower case k (kui, kif, kuf, and so on) denotes a rate constant

ΔG terms are the free energies of reactions

Any term with α (ILα, kifα) is the equivalent of the term

without α (I, kif) when a ligand is bound to the α site (same for β and $\alpha\beta$)

Terms in brackets [] are concentrations in units of molar (M)

Differential Equations for Addition of Osmolyte

$$\frac{d[U]}{dt} = -kUI\,[U] + kIU\,[I] - kUF[U] + kFU\,[F]$$

$$\frac{d[I]}{dt} = -kIU\,[I] - kIF\,[I] + kUI\,[U] + kFI\,[F]$$

$$\frac{d[F]}{dt} = -kFI\,[F] + kIF\,[I] + kUF\,[U] - kFU\,[F]$$

$$[U]\,(0) = \frac{[Pt]}{1 + Kui\,(1 + Kif)}$$

$$[I]\,(0) = \frac{Kui\,[Pt]}{1 + Kui\,(1 + Kif)}$$

$$[F]\,(0) = \frac{Kui\,Kif\,[Pt]}{1 + Kui\,(1 + Kif)}$$

(*when osmolyte* (*Os*) *is present, kui, kuf, and kif are written as kUI, kUF,*

and kIF. These are the rates as a function of osmolyte concentration. Kui becomes KUI, Kuf becomes KUF and so on.)

(*mui and mif are m − values that describe the osmolyte dependence of the folding free energy. βij are β −*

Tanford values that describe the interaction of osmolyte with the transition state when going from state i to state j.)

$$kUI = e^{\,Log[kui]+(\beta ui-1)\frac{mui}{R\,Temp}[Os]}$$

$$kUF = e^{\,Log[kuf]+(\beta uf-1)\frac{(mui+mif)}{R\,Temp}[Os]}$$

$$kIF = e^{\,Log[kif]+(\beta if-1)\frac{mif}{R\,Temp}[Os]}$$

$$KUI = e^{\,\frac{\Delta Gui + mui\,Os}{-R\,Temp}}$$

$$KIF = e^{\,\frac{\Delta Gif + mif\,Os}{-R\,Temp}}$$

$$KUF = KUI\,KIF$$

$$kIU = \frac{kUI}{KUI}$$

$$kFU = \frac{kUF}{KUF}$$

$$kFI = \frac{kIF}{KIF}$$

Observable

$$FSIGU = SigU + SlopeU\,[Os]$$
$$FSIGI = SigI + SlopeI[Os]$$
$$FSIGF = SigF + SlopeF[Os]$$

$$\text{Observable Signal} = FSIGU\,\frac{[U]}{[Pt]} + FSIGI\,\frac{[I]}{[Pt]} + FSIGF\,\frac{[F]}{[Pt]}$$

Differential equations for ligand induced folding

$$\frac{d[U]}{dt} = -kui\,[U]\ + kiu\,[I] - kau\alpha\,[U]\,[Lf] - kau\beta\,[U]\,[Lf]\ + kdu\alpha\,[UL\alpha]\ + kdu\beta\,[UL\beta] - kuf\,[U] + kfu\,[F]$$

$$\frac{d[UL\alpha]}{dt} =$$
$$-kdu\alpha\,[UL\alpha]\ + kau\alpha\,[U]\,[Lf] - kui\alpha\,[UL\alpha]\ + kiu\alpha\,[IL\alpha]\ - kau\beta\,[UL\alpha]\,[Lf] + kdu\beta\,[UL\alpha\beta] - kuf\alpha\,[UL\alpha] + kfu\alpha\,[FL\alpha]$$

$$\frac{d[UL\beta]}{dt} =$$
$$-kdu\beta\,[UL\beta]\ + kau\beta\,[U]\,[Lf] - kui\beta\,[UL\beta]\ + kiu\beta\,[IL\beta]\ - kau\alpha\,[UL\beta]\,[Lf] + kdu\alpha\,[UL\alpha\beta] - kuf\beta\,[UL\beta] + kfu\beta\,[FL\beta]$$

$$\frac{d[UL\alpha\beta]}{dt} = -kdu\alpha\,[UL\alpha\beta]\ - kdu\beta\,[UL\alpha\beta]\ + kau\alpha\,[UL\beta]\,[Lf] +$$
$$kau\beta\,[UL\alpha]\,[Lf] - kui\alpha\beta\,[UL\alpha\beta]\ + kiu\alpha\beta\,[IL\alpha\beta] - kuf\alpha\beta\,[UL\alpha\beta] + kfu\alpha\beta\,[FL\alpha\beta]$$

$$\frac{d[I]}{dt} == -kiu\,[I]\ - kif\,[I]\ + kui\,[U]\ + kfi\,[F]\ - kai\alpha\,[I]\,[Lf] - kai\beta\,[I]\,[Lf] + kdi\alpha\,[IL\alpha]\ + kdi\beta\,[IL\beta]$$

$$\frac{d[IL\alpha]}{dt} = -kiu\alpha\,[IL\alpha]\ - kif\alpha\,[IL\alpha]\ + kui\alpha\,[UL\alpha]\ + kfi\alpha\,[FL\alpha]\ - kdi\alpha\,[IL\alpha]\ + kai\alpha\,[I]\,[Lf] - kai\beta\,[IL\alpha]\,[Lf] + kdi\beta\,[IL\alpha\beta]$$

$$\frac{d[IL\beta]}{dt} = -kiu\beta\,[IL\beta]\ - kif\beta\,[IL\beta]\ + kui\beta\,[UL\beta]\ + kfi\beta\,[FL\beta]\ - kdi\beta\,[IL\beta]\ + kai\beta\,[I]\,[Lf] - kai\alpha\,[IL\beta]\,[Lf] + kdi\alpha\,[IL\alpha\beta]$$

$$\frac{d[IL\alpha\beta]}{dt} =$$
$$-kiu\alpha\beta\,[IL\alpha\beta]\ - kif\alpha\beta\,[IL\alpha\beta]\ + kui\alpha\beta\,[UL\alpha\beta]\ + kfi\alpha\beta\,[FL\alpha\beta]\ - kdi\alpha\,[IL\alpha\beta]\ - kdi\beta\,[IL\alpha\beta]\ + kai\alpha\,[IL\beta]\,[Lf] + kai\beta\,[IL\alpha]\,[Lf]$$

$$\frac{d[F]}{dt} = -kfi\,[F]\ + kif\,[I]\ - kaf\alpha\,[F]\,[Lf] - kaf\beta\,[F]\,[Lf] + kdf\alpha\,[FL\alpha]\ + kdf\beta\,[FL\beta]\ + kuf\,[U] - kfu\,[F]$$

$$\frac{d[FL\alpha]}{dt} = -kdf\alpha\,[FL\alpha]\ + kaf\alpha\,[F]\ [Lf] - kfi\alpha\,[FL\alpha] + kif\alpha\,[IL\alpha]\ - kaf\beta\,[FL\alpha]\ [Lf] + kdf\beta\,[FL\alpha\beta] + kuf\alpha\,[UL\alpha] - kfu\alpha\,[FL\alpha]$$

$$\frac{d[FL\beta]}{dt} = -kdf\beta\,[FL\beta]\ + kaf\beta\,[F]\ [Lf] - kfi\beta\,[FL\beta]\ + kif\beta\,[IL\beta]\ - kaf\alpha\,[FL\beta]\ [Lf] + kdf\alpha\,[FL\alpha\beta] + kuf\beta\,[UL\beta] - kfu\beta\,[FL\beta]$$

$$\frac{d[FL\alpha\beta]}{dt} = -kdf\alpha\,[FL\alpha\beta]\ - kdf\beta\,[FL\alpha\beta]\ + kaf\alpha\,[FL\beta]\ [Lf] +$$
$$kaf\beta\,[FL\alpha]\ [Lf] - kfi\alpha\beta\,[FL\alpha\beta]\ + kif\alpha\beta\,[IL\alpha\beta] + kuf\alpha\beta\,[UL\alpha\beta] - kfu\alpha\beta\,[FL\alpha\beta]$$

$$\frac{d[Lf]}{dt} = -kau\alpha\,[U]\ [Lf] - kau\beta\,[U]\ [Lf] - kau\alpha\,[UL\beta]\ [Lf] - kau\beta\,[UL\alpha]\ [Lf] - kai\alpha\,[I]\ [Lf] -$$
$$kai\beta\,[I]\ [Lf] - kai\alpha\,[IL\beta]\ [Lf] - kai\beta\,[IL\alpha]\ [Lf]\ - kaf\alpha\,[F]\ [Lf] - kaf\beta\,[F]\ [Lf] - kaf\beta\,[FL\alpha]\ [Lf] -$$
$$kaf\alpha\,[FL\beta]\ [Lf] + kdf\alpha\,[FL\alpha\beta]\ + kdf\beta\,[FL\alpha\beta] + kdf\alpha\,[FL\alpha]\ + kdf\beta\,[FL\beta] + kdi\alpha\,[IL\alpha\beta]\ +$$
$$kdi\beta\,[IL\alpha\beta]\ + kdi\alpha\,[IL\alpha]\ + kdi\beta\,[IL\beta]\ + kdu\alpha\,[UL\alpha\beta]\ + kdu\beta\,[UL\alpha\beta]\ + kdu\alpha\,[UL\alpha]\ +\ kdu\beta\,[UL\beta]$$

Initial Concentrations of Species
All liganded species have initial concentrations (time = 0 seconds) of 0 M

$$[Lf]\,(0) = Lt$$

$$[U]\,(0) = \frac{[Pt]}{1 + Kui\,(1 + Kif)}$$

$$[I]\,(0) = \frac{Kui\,[Pt]}{1 + Kui\,(1 + Kif)}$$

$$[F]\,(0) = \frac{Kui\,Kif\,[Pt]}{1 + Kui\,(1 + Kif)}$$

$$[U] + [UL\alpha] + [UL\beta] + [UL\alpha\beta] + [I] + [IL\alpha] + [IL\beta] + [IL\alpha\beta] + [F] + [FL\alpha] + [FL\beta] + [FL\alpha\beta] = [Pt]$$

$R = 1.9858775\ \text{cal}\,M^{-1}\,K^{-1}$ (∗Gas Law Constant∗)

$Temp = 298$ (∗K∗)

$Kui = e^{\frac{\Delta Gui}{-R\,Temp}}$ (Equilibrium constant for U to I transition)

$Kif = e^{\frac{\Delta Gif}{-R\,Temp}}$ (Equilibrium constant for I to F transition)

$K\alpha u = e^{\frac{\Delta G\alpha u}{-R\,Temp}}$ (Equilibrium association constant of ligand binding to the α site of U)

$K\alpha i = e^{\frac{\Delta G\alpha i}{-R\,Temp}}$ (Equilibrium association constant of ligand binding to the α site of I)

$K\alpha f = e^{\frac{\Delta G\alpha f}{-R\,Temp}}$ (Equilibrium association constant of ligand binding to the α site of F)

$K\beta u = e^{\frac{\Delta G\beta u}{-R\,Temp}}$ (Equilibrium association constant of ligand binding to the β site of U)

$K\beta i = e^{\frac{\Delta G\beta i}{-R\,Temp}}$ (Equilibrium association constant of ligand binding to the β site of I)

$K\beta f = e^{\frac{\Delta G\beta f}{-R\,Temp}}$ (Equilibrium association constant of ligand binding to the β site of F)

(*The free energies above of the free energies that correspond to the transitions described by the equilibrium constants*)

$Kuf = Kui\,Kif$ (Equilibrium constant for U to F transition)

On rates (kon)

Units of $M^{-1} s^{-1}$ (per molar per second)

$kon = kaf\alpha = kaf\beta = kai\alpha = kai\beta = kau\alpha = kau\beta = 1 * 10^8$

List of off rates $kdxz$ where x is the conformational state (u = U i = I or f = F) and z is the liganded site (α or β)
Units are of s^{-1}

$$kdu\alpha = \frac{kau\alpha}{K\alpha u}$$

$$kdu\beta = \frac{kau\beta}{K\beta u}$$

$$kdi\alpha = \frac{kai\alpha}{K\alpha i}$$

$$kdi\beta = \frac{kai\beta}{K\beta i}$$

$$kdf\alpha = \frac{kaf\alpha}{K\alpha f}$$

$$kdf\beta = \frac{kaf\beta}{K\beta f}$$

List of unfolding rate constants in all different ligation states
Units are s^{-1}

$$kfi = \frac{kif}{Kif}$$

$$kfi\alpha = \frac{kif\alpha \, K\alpha i}{Kif \, K\alpha f}$$

$$kfi\beta = \frac{kif\beta \, K\beta i}{Kif \, K\beta f}$$

$$kfi\alpha\beta = \frac{kif\alpha\beta \, K\alpha i \, K\beta i}{Kif \, K\alpha f \, K\beta f}$$

$$kiu = \frac{kui}{Kui}$$

$$kiu\alpha = \frac{kui\alpha \, K\alpha u}{Kui \, K\alpha i}$$

$$kiu\beta = \frac{kui\beta \, K\beta u}{Kui \, K\beta i}$$

$$kiu\alpha\beta = \frac{kui\alpha\beta \, K\alpha u \, K\beta u}{Kui \, K\alpha i \, K\beta i}$$

$$kfu = \frac{kuf}{Kuf}$$

$$\mathrm{kfu}\alpha = \frac{\mathrm{kuf}\alpha\ \mathrm{K}\alpha\mathrm{u}}{\mathrm{Kuf}\ \mathrm{K}\alpha\mathrm{f}}$$

$$\mathrm{kfu}\beta = \frac{\mathrm{kuf}\beta\ \mathrm{K}\beta\mathrm{u}}{\mathrm{Kuf}\ \mathrm{K}\beta\mathrm{f}}$$

$$\mathrm{kfu}\alpha\beta = \frac{\mathrm{kuf}\alpha\beta\ \mathrm{K}\alpha\mathrm{u}\ \mathrm{K}\beta\mathrm{u}}{\mathrm{Kuf}\ \mathrm{K}\alpha\mathrm{f}\ \mathrm{K}\beta\mathrm{f}}$$

Observable

$$\mathrm{UFree} = \frac{[\mathrm{U}]}{[\mathrm{Pt}]}$$

$$\mathrm{UBound} = \frac{[\mathrm{UL}\alpha] + [\mathrm{UL}\beta] + [\mathrm{UL}\alpha\beta]}{[\mathrm{Pt}]}$$

$$\mathrm{IFree} = \frac{[\mathrm{I}]}{[\mathrm{Pt}]}$$

$$\mathrm{IBound} = \frac{[\mathrm{IL}\alpha] + [\mathrm{IL}\beta] + [\mathrm{IL}\alpha\beta]}{[\mathrm{Pt}]}$$

$$\mathrm{FFree} = \frac{[\mathrm{F}]}{[\mathrm{Pt}]}$$

$$\mathrm{FBound} = \frac{[\mathrm{FL}\alpha] + [\mathrm{FL}\beta] + [\mathrm{FL}\alpha\beta]}{[\mathrm{Pt}]}$$

Observable Signal = SignalUFree UFree + SignalIFree IFree +
SignalFFree FFree + SignalUBound UBound + SignalIBound IBound + SignalFBound FBound

References

(1) Niranjanakumari, S.; Kurz, J. C.; Fierke, C. A. *Nucleic Acids Research* **1998**, *26*, 3090.

(2) Edelhoch, H. *Biochemistry* **1967**, *6*, 1948.

(3) Minor, W.; Cymborowski, M.; Otwinowskib, Z.; Chruszcza, M. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 859.

(4) Stams, T.; Niranjanakumari, S.; Fierke, C. A.; Christianson, D. W. *Science* **1998**, *280*, 752.

(5) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213.

(6) Rueda, D.; Hsieh, J.; Storms, J. J.; Fierke, C. A.; Walter, N. G. *Biochemistry* **2005**, *44*, 16130.

(7) Liu, X. Ph.D. Dissertation University of Michigan, Ann Arbor, 2013.

(8) Matthews, C. R. *Meth. Enzymol.* **1987**, *154*, 498.

(9) Chopin, N. *Biometrika* **2002**, *89*, 539.

(10) Gilks, W. R.; Richardson, S.; Spiegelhalter, D. J. *Markov chain Monte Carlo in practice*; Chapman & Hall/CRC, 1996.

(11) Doucet, A.; De Freitas, N.; Gordon, N. In *Statistics for Engineering and Information Science*; Springer: 2001.

(12) Chang, Y. C.; Franch, W. R.; Oas, T. G. *Biochemistry* **2010**, *49*, 9428.

(13) Li, A.; Sowder, R. C.; Henderson, L. E.; Moore, S. P.; Garfinkel, D. J.; Fisher, R. J. *Anal. Chem.* **2001**, *73*, 5395.