Linked Folding and Anion Binding of the *Bacillus subtilis* Ribonuclease P Protein†

Christopher H. Henkels,‡ Jeffrey C. Kurz,§ Carol A. Fierke,§ and Terrence G. Oas*,‡

Department of Biochemistry, Box 3711, Duke University Medical Center, Durham, North Carolina 27710, and Department of Chemistry, University of Michigan, 930 North University Road, Ann Arbor, Michigan 48109

Received September 1, 2000; Revised Manuscript Received December 22, 2000

ABSTRACT: Ribonuclease P (RNase P) is the endoribonuclease responsible for the 5′-maturation of precursor tRNA transcripts. In bacteria, RNase P is composed of a catalytic RNA subunit and an associated protein subunit that enhances the substrate specificity of the holoenzyme. We have initiated a study of the biophysical properties of the protein subunit from *Bacillus subtilis* RNase P (P protein) toward the goal of understanding the thermodynamics of RNase P holoenzyme assembly. The P protein is predominantly unfolded in 10 mM sodium cacodylate at neutral pH based on circular dichroism and NMR studies and therefore has several characteristics typical of "intrinsically unstructured" proteins. Furthermore, the P protein folds to its native α/β structure upon addition of various small molecule anions. Anion-induced folding is best attributed to the binding of these anions to the folded state of the protein, and a model is presented which describes the observed tightly coupled folding and binding phenomena. The P protein also undergoes a cooperative folding transition upon addition of the osmolyte trimethylamine *N*-oxide (TMAO). The equilibrium constant of folding (K_{fold}) at 37 °C for the P protein was determined to be 0.0071 ± 0.0005 using a two-state folding model to describe the TMAO titration data. Thus, the folding and binding equilibria observed in the anion-induced folding of the P protein can be uncoupled to determine the intrinsic binding affinities (K_a's) of the anionic ligands. Evidence that the osmolyte-induced and the ligand-induced folded conformations of the P protein are structurally similar is also presented.

In a wide variety of systems, the regulation of biological activity is achieved through the thermodynamic linkage of macromolecular ligand binding to changes in activity or conformation. In addition to alterations in enzymatic activity, ligand binding can affect other protein properties such as response to environmental conditions or aggregation state (*1*). Ligand-induced structural changes lead to conformations that are energetically unfavorable in the absence of ligand. This energetic coupling can increase the binding specificity for a particular ligand, and thereby enhance the selectivity of associated regulation mechanisms. Tightly bound ligands can induce even a highly unfavorable conformational change, and some proteins may even remain completely unfolded in the absence of their binding partners. Indeed, recent studies have reported several proteins or protein domains that change from a predominantly unfolded state to a more compact ("native") fold when bound to their associated ligand, including protein (*2*−*7*), RNA (*8*−*10*), or DNA (for a review, see ref *11*) partners. Given the diversity of systems that exhibit this phenomenon (Table 1), more proteins that are "intrinsically unstructured" (*12*) are likely to be found in the future. The reason some, but not all, proteins are intrinsically unstructured when isolated from their associated binding partners is not currently understood. Additionally, the biological relevance of ligand-linked disorder-to-order transitions remains to be elucidated. One hypothesis is that the thermodynamic linkage of protein folding to ligand binding yields enhanced control in cellular macromolecular assembly (*5*). Ligand-induced protein folding may also allow a protein to recognize several binding targets (*3*) or allow for cellular regulation of activity via proteolytic degradation (*12*).

Table 1: List of Some Intrinsically Unstructured Proteins[a]

protein	binding partner	ref
(1) protein binding		
FlgM	σ^{28}	*2*
$p21^{Waf1/Cip1/Sdi1}$	Cdk2	*3*
fibronectin-binding protein	fibronectin	*4*
Mn^{2+}-stabilizing protein	PSII	*5*
NH_2 domain of $TAF_{II}230$	TBP	*6*
Sso1, Snc1, and Sec9	each other	*7*
(2) RNA binding		
protein N of λ phage	box B RNA	*8, 10*
M domain of Ffh	4.5S RNA	*9*
(3) DNA binding		
TrpR; GR DBD; *Eco*RV; λ cI repressor; GCN4	site-specific DNA binding sequences	reviewed in ref *11*

[a] See Wright and Dyson (*12*) for a review of intrinsically unstructured proteins. The extent of protein folding upon ligand binding varies greatly from the ordering of loops or α-helix formation to the folding of a domain or the entire protein.

Implicit in the studies described above is the notion that the folded, unliganded state is energetically less favored than both the liganded, folded state and the unfolded state. Thus, the unliganded, native state represents a metastable state (*13*) in the folding and binding mechanism. To better describe the folding mechanisms of these systems, it is necessary to

† Supported by National Institutes of Health Grants GM45322 (T.G.O.) and GM55387 (C.A.F.). C.H.H. and J.C.K. were supported in part by NIH Training Grant GM08487.

* To whom correspondence should be addressed. Telephone: (919) 684-4363. Fax: (919) 681-8862. E-mail: oas@biochem.duke.edu.

‡ Duke University Medical Center.

§ University of Michigan.

10.1021/bi002078y CCC: $20.00
Published on Web 02/06/2001

develop a means of separating the tightly coupled folding and binding equilibria. Previous studies have primarily investigated the structures of these proteins in their free and bound forms rather than the thermodynamics of the coupled process. Here, we develop an experimental method for studying the metastable unliganded native state of a protein that is predominantly unfolded in the absence of ligand using the protein component of the *Bacillus subtilis* ribonuclease P (P protein).[1]

Ribonuclease P (RNase P) is the enzyme responsible for the processing of the 5′-leader sequences from precursor tRNA (pre-tRNA) substrates to produce tRNA molecules containing the mature 5′-end (for reviews, see refs *14*−*17*). This catalytic activity appears to be ubiquitous because RNase P has been identified in Archea, Bacteria, and Eukarya. Originally identified in *Escherichia coli* (*18*), bacterial versions of RNase P are ribonucleoprotein complexes composed of an RNA subunit of ∼400 nucleotides and a protein subunit of ∼120 amino acids. In bacteria, all of the components necessary for catalysis are found in the RNA subunit (*19*). Unlike that from bacteria, eukaryotic RNase P consists of an RNA subunit and nine protein subunits (*15*, *20*), and all of the subunits are required to achieve catalytic activity in vivo (*14*, *20*). Given the compositional differences in RNase P between bacteria and eukaryia, bacterial RNase P is an attractive target for the development of novel antibiotics (*21*).

Although the P RNA component alone can cleave pre-tRNA substrates under elevated salt conditions, the protein subunit is essential for catalysis in vivo (*22*). The P protein enhances holoenzyme efficiency by increasing the affinity of pre-tRNA via direct interactions with the 5′-leader region of pre-tRNA (*23*−*26*). These interactions also modulate the substrate specificity of RNase P (*27*). The protein component may also enhance the magnesium affinity (*23*, *28*) and alter the structure of the RNA component (*29*).

The crystal structure of the *B. subtilis* P protein (*30*) indicates that the P protein is a globular mixed α/*β* protein (Figure 1) that contains an overall secondary structure of α*ββββ*αβα and is structurally similar to other RNA-binding protein domains (*31*, *32*). There is an unusual left-handed crossover that is conserved among these RNA-binding proteins; this crossover region also represents the region of greatest sequence identity among bacterial P proteins and is known as the RNR motif (*33*). A central cleft is formed by the NH_2-terminal α-helix and the central *β*-sheet that exposes aromatic residues to the solvent (F16, F20, and Y34), and cross-linking studies suggest that this region of the protein interacts with pre-tRNA (*26*). Regions of the ribozyme that are altered upon formation of the holoenzyme have been identified using RNase T1 protection experiments (*34*), deletion analysis, and chemical protection studies (*29*, *35*, *36*). Unlike other RNA-binding proteins that bind a stem−loop structure (*8*, *37*), the RNase P protein recognizes several regions distant in the ribozyme primary sequence and secondary structure (*35*, *36*, *38*). Amino acids in the protein subunit that are important for the holoenzyme interaction have also been analyzed by mutagenesis (*39*, *40*), cross-linking (*30*), and directed affinity cleavage experiments (*38*). The results suggest that conserved basic and aromatic residues are important for the holoenzyme RNA−protein assembly. These residues include amino acids located in the RNR motif (R60, N61, and R65; *B. subtilis* numbering) and the central cleft (F16 and F20), as well as residues on the protein surface (R45) and core (F107).

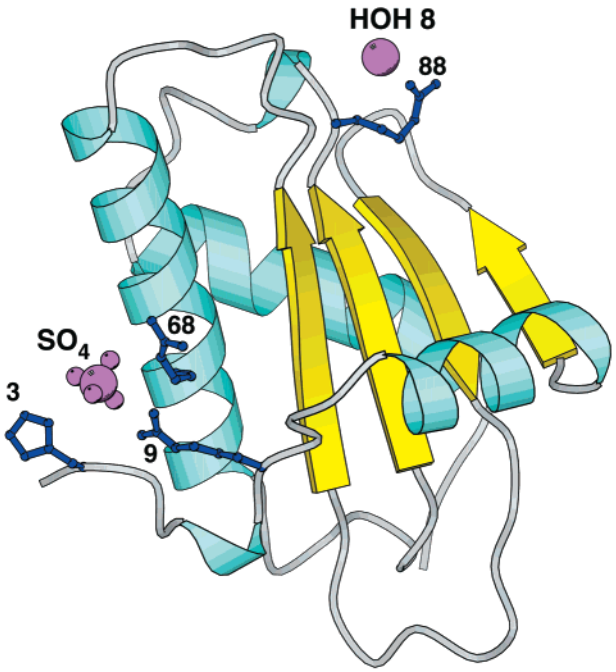


FIGURE 1: Ribbon diagram of the *B. subtilis* RNase P protein based on the crystal structure of Christianson and co-workers (*30*). The originally observed sulfate is indicated with the position of a water molecule that may represent the second sulfate binding site (see the text). The Protein Data Bank code is 1A6F. Ribbon diagrams were generated using the program MOLSCRIPT (*82*).

Here we demonstrate that the protein component of *B. subtilis* RNase P is predominantly unfolded when isolated from RNA and small molecule anions in a low-ionic strength buffer. The P protein can be induced to refold by increasing the anion concentration in solution or by the addition of the osmolyte trimethylamine *N*-oxide (TMAO). On the basis of these findings, we propose a model that describes the anion-induced folding of the P protein. This study was initiated to discern the thermodynamics of P protein folding as well as to develop a means of calculating intrinsic association constants for P protein ligands toward the goal of elucidating the thermodynamics of RNase P holoenzyme assembly.

MATERIALS AND METHODS

Expression and Purification of the P Protein. Plasmid pPWT-28b contains the gene for the *B. subtilis* RNase P protein subcloned from pPWT1 (*41*) and inserted into the

[1] Abbreviations: apoC1, apolipoprotein C-1; CD, circular dichroism; (d)CMP, (dexoy)cytidine monophosphate; (d)CTP, (dexoy)cytidine triphosphate; EDTA, (ethylenedinitrilo)tetraacetic acid; EPR, electron paramagnetic resonance; F_n, fraction native; HSQC, heteronuclear single-quantum coherence; LB, Luria-Bertani; MES, 2-(*N*-morpholino)-ethanesulfonic acid; NMR, nuclear magnetic resonance; pI, isoelectric point; PP_i, pyrophosphate; P protein, protein subunit of *B. subtilis* ribonuclease P; pre-tRNA, precursor tRNA; P RNA, RNA subunit of *B. subtilis* ribonuclease P; RNase P, ribonuclease P; RNase T1, ribonuclease T1; Sac7d, member (denoted d) of the 7 kDa DNA-binding proteins from *Sulfolobus acidocaldarius*; TMAO, trimethylamine *N*-oxide; TMSP, sodium 3-(trimethylsilyl)propionate; Tris, tris(hydroxymethyl)aminomethane.

pET-28b T7 expression vector. Recombinant RNase P protein was expressed and purified as previously described (*41*; method 1) with the following modifications. The P protein fractions from the second CM-Sepharose (Pharmacia) column were pooled, dialyzed against water extensively, and subsequently lyophilized. The dried protein was resuspended in 5% acetic acid and applied over a Sephadex G50 (Pharmacia) column (2 cm × 50 cm) equilibrated with 5% acetic acid. Fractions off the sizing column were monitored at 234 nm and yielded an elution profile containing two resolvable peaks. The fractions corresponding to the first (protein) peak were pooled, lyophilized, and stored as a dry powder in a desiccated environment. In all experiments, protein quantitation was based on the method of Edelhoch (*42*) using an extinction coefficient of 6000 M^{-1} cm^{-1} at 275.5 nm.

The ^{15}N-labeled P protein was purified using the protocol as described above; however, the expression parameters are varied. An overnight culture [grown in LB medium (*41*)] was diluted 1:20 into an unlabeled M9 minimal culture and grown at 37 °C to an OD_{600} of 0.5−0.6. The unlabeled minimal culture was then diluted (1:50) into a 1 L M9 minimal medium culture enriched with ^{15}N-labeled ammonium chloride. Once the ^{15}N-labeled culture reached an OD_{600} of 0.4−0.5, overexpression of the P protein was induced by the addition of 1 mM isopropyl β-D-thiogalactopyranoside and the culture was incubated at 37 °C for an additional 8−12 h prior to harvesting the cells.

Circular Dichroism (CD) Spectroscopy. CD spectra were collected on an Aviv model 202 CD spectrophotometer with either 5 or 10 μM P protein in 10 mM sodium cacodylate buffer (pH 7) or 10 mM Tris cacodylate (pH 8) in thermostated quartz cells. Wavelength traces were scanned from 300 to 190 nm in a 1 mm cuvette at 1 nm increments with a bandwidth of 1 or 1.5 mm and averaged over a time of 5 s; the temperature was held constant at 37 °C. Near-UV CD wavelength spectra were collected at two different experimental conditions. Preliminary scans were carried out using 200 μM P protein in 10 mM sodium cacodylate (pH 7) in a thermostated 1 mM cuvette with the temperature held constant at 37 °C. Subsequent near-UV CD spectra were collected using 25 μM P protein in either 10 mM sodium cacodylate (pH 7) or 10 mM Tris cacodylate (pH 8) in a nonthermostated 10 cm cell (see below); these latter conditions were chosen to minimize the protein concentration required to observe the signal. In both experimental conditions, the near-UV CD scans were carried out from 350 to 250 nm in 1 nm increments with a bandwidth of 1 or 1.5 mm and a signal averaging time of 5 s.

Manual titrations of the P protein with various anions were carried out at 37 °C using a bandwidth of 1.5 mm. A 2 mL solution containing 5 or 10 μM P protein in 10 mM sodium cacodylate (pH 7) was placed in a 1 cm cuvette and allowed to reach thermal equilibrium at 37 °C for ∼15 min. The initial CD signal at 222 nm was followed for 30 s at 0.5 s intervals; the average CD signal was then calculated. High-molarity salt solutions (1 M or 100 mM) in 10 mM sodium cacodylate (pH 7) were prepared, and the pH was adjusted with either NaOH or cacodylic acid. Serial dilutions from the high-concentration stocks were also made in 10 mM sodium cacodylate buffer (pH 7). An aliquot of an appropriate salt solution was added to the cuvette, and the CD signal was measured again over an additional 30 s at 0.5 s intervals. At any given anion concentration, the CD signal was followed in triplicate. The change in protein concentration was used to correct the raw CD signal. An additional sulfate titration was also carried out at a constant temperature of 28 °C; this experiment was used in comparison with the near-UV CD titration experiment (see below) to analyze the sulfate-induced extent of protein folding.

A sulfate titration following the near-UV CD signal at 270 nm was also carried out in the kinetic mode for 30 s where the signal was followed at 0.5 s intervals. A nonthermostated 10 cm cuvette was used to carry out the near-UV CD titration experiment, and the protein concentration was 25 μM; the temperature was followed manually and determined to be constant at 27.9 ± 0.4 °C.

Thermal melts of the P protein were monitored at 222 nm from 2 to 90 °C at 1 °C increments. The samples were initially allowed to equilibrate at 2 °C for at least 15 min prior to melting. An equilibration time of 30 s and a signal averaging time of 30 s were used throughout the melt, and the temperature bandwidth was 0.2 °C. Previous experience with similar samples on the same instrument indicates that thermal equilibrium is reached within the equilibration time. Thermal melts were fit to eq 1:

$$[\theta]_{222}(T) = \{(b_N + m_N T) + (b_D + m_D T)[\exp(-\Delta G_d/RT)]\}/[1 + \exp(-\Delta G_d/RT)] \quad (1)$$

where

$$\Delta G_d = \Delta H_m(1 - T/T_m) + \Delta C_p[T - T_m - T\ln(T/T_m)] \quad (2)$$

where $[\theta]_{222}(T)$ is the observed CD signal as a function of temperature, b_N and b_D and m_N and m_D are the *y*-intercept signals (0°) and temperature-dependent slopes of the native and denatured baselines, respectively. ΔG_d is the free energy of thermal denaturation using the modified Gibbs−Helmholtz equation where ΔH_m is the enthalpy change at the midpoint of thermal transition, T_m, and ΔC_p is the change in heat capacity upon thermal denaturation (*43*). Thermal denaturations of the P protein were at least 90% reversible under the experimental conditions used in this work, based on the low-temperature CD signal of a native sample before and after thermal denaturation (data not shown).

Titration of the P protein with trimethylamine *N*-oxide (TMAO) was performed at 28, 37, 50, and 60 °C using 10 μM P protein in 10 mM Tris cacodylate (pH 8) buffer. The pH is adjusted to 8 rather than 7 in all TMAO experiments to limit the amount of cacodylic acid needed to adjust the pH; the final cacodylate concentration in all experiments was at most 25 mM, less than the apparent K_{app} for cacodylate (∼200 mM, data not shown). Solutions were allowed to reach thermal equilibrium by waiting at least 15 min at the appropriate temperature prior to the experiment. The CD signal at 222 nm was followed as a function of TMAO concentration using a mixing equilibration time of 1 min followed by an averaging time of 30 s. The TMAO concentration was measured using the refractive index as described by Wang and Bolen (*44*). The resulting titration curve was fit to a two-state *folding* model assuming a linear

$\Delta G_{folding}$ versus TMAO dependence (eq 3):

$$\theta_{222}([TMAO]) = ((b_D + m_D[TMAO]) + (b_N + m_N[TMAO])\{\exp[m([TMAO] - [TMAO]_{1/2})/RT]\})/\{1 + \exp[m([TMAO] - [TMAO]_{1/2})/RT]\} \quad (3)$$

where $\theta_{222}([TMAO])$ is the observed CD signal as a function of TMAO concentration, b_D and b_N and m_D and m_N are the intercepts (0 M TMAO) and the TMAO-dependent slopes of the denatured and native baselines, respectively, $[TMAO]_{1/2}$ is the midpoint of renaturation, and m is the slope of the linear dependence of the free energy of folding ($\Delta G_{folding}$) on the TMAO concentration.

NMR Spectroscopy. For HSQC experiments, the lyophilized ^{15}N-labeled P protein was dissolved in 10 mM sodium cacodylate (pH 7.0), 20 μg/mL TMSP, and 10% D_2O. Gradient-enhanced, sensitivity-enhanced $^1H-^{15}N$ HSQC experiments (*45*) were carried out on a Varian Unity Plus 800 MHz spectrometer at 28 °C. A total of 2602 points were collected to encompass a sweep width of 14 005.6 Hz in the proton dimension, while 128 increments were collected in the nitrogen dimension having a sweep width of 4000 Hz. The proton dimension was zero-filled to 4096 points, while the indirectly detected ^{15}N dimension was zero-filled to 512 points; both of the dimensions were apodized with shifted sine-bell functions. The data were processed with forward and backward linear prediction in the ^{15}N dimension. All of the HSQC spectra were processed using NMRPipe (*46*), and the carrier frequencies for the proton and the nitrogen dimensions were 4.734 and 119.72 ppm, respectively.

RNase P Activity Assay. Single-turnover ([E]/[S] > 5) measurements of RNase P activity were performed essentially as described previously (*24*) at 37 °C in buffer containing 10 mM $MgCl_2$, 100 mM KCl, 50 mM Tris, and 50 mM MES (pH 6.1), with 2 M TMAO added as indicated. The RNase P holoenzyme (0.5−1.0 μM) was reconstituted by incubating RNase P RNA with equimolar RNase P protein for 10 min just prior to the reaction. Reactions were initiated by the addition of *B. subtilis* pre-$tRNA^{Asp}$ [containing a 14-nucleotide 5′-precursor sequence (*25*)] to RNase P. Single-turnover rate constants were obtained from the appearance of a cleaved product over time, by fitting the data to the equation for a single exponential (eq 4):

$$[P] = [P]_\infty[1 - \exp(-k_{obs}t)] \quad (4)$$

Analytical Techniques. R. Stevens performed electrospray ionization mass spectrometric analysis of the P protein at the Mass Spectrometry Facility, Duke University Medical Center.

RESULTS

The B. subtilis RNase P Protein Is Unfolded in Low-Ionic Strength Cacodylate Buffer

The P protein purified using the protocol of Fierke and co-workers (*41*) has the following characteristics in 20 mM sodium phosphate buffer (pH 6.5): (1) a far-UV CD spectrum consistent with the published α/β crystal structure and (2) a thermal melting point of 66 °C (*41*). Mass spectroscopic analysis of this protein preparation reveals two extra peaks with molecular masses 316 and 612 Da larger than the expected mass (data not shown). The relative peak heights of the higher-molecular mass species were dependent upon the ionization energy used in the mass spectroscopic analysis, suggesting that the 316 Da species is noncovalently bound to the protein. Removal of the small molecule contaminant was achieved by gel filtration under denaturing conditions with 5% acetic acid as a solvent. The contaminant was identified as EDTA by mass spectroscopy and NMR analysis; EDTA is a component of all of the purification buffers used in the procedure of Fierke and co-workers (*41*).

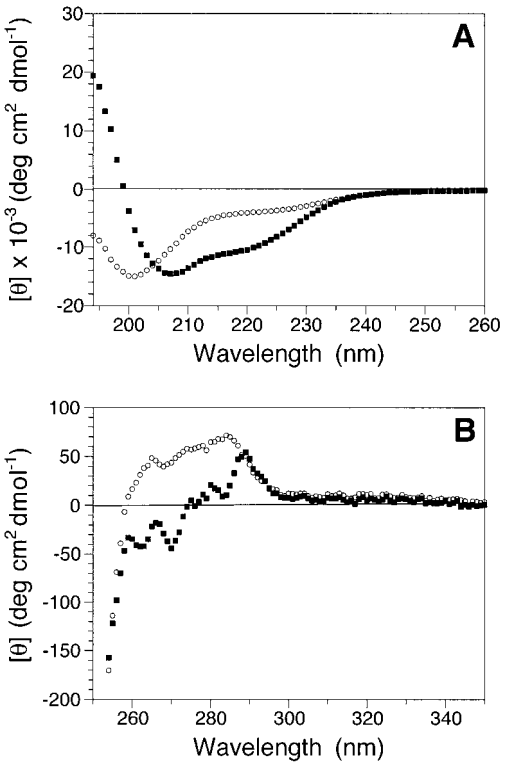


FIGURE 2: Biophysical characterization of the P protein in the presence and absence of sulfate. (A) Far-UV CD wavelength scans of the P protein (10 μM) in cacodylate buffer in the absence of sodium sulfate (○) and in 20 mM sulfate (■). (B) Near-UV CD spectra of the P protein (200 μM; 1 mM cuvette) in 10 mM cacodylate, pH 7 buffer. Spectra of the P protein in the absence of sulfate (○) and in the presence of 10 mM sulfate (■) were scanned from 350 to 250 nm at 1 nm increments.

A far-UV CD wavelength spectrum of the EDTA-free P protein (termed unliganded P protein) in 10 mM sodium cacodylate (pH 7) is shown in Figure 2A. Sodium cacodylate was used as the buffer system because cacodylate, in contrast to phosphate, does not bind to DNA-binding proteins (*47*). The spectrum of the unliganded P protein is substantially different from the spectrum of the P protein previously described (*41*). The unliganded P protein has a wavelength minimum at 202 nm in cacodylate buffer, while the previously published minimum is ~208 nm (Figure 2A). Furthermore, the overall CD signal is much weaker in the 208−222 nm range compared to that of the previously reported spectrum. The unliganded protein contains less secondary structure as monitored by far-UV CD, indicating that the P protein undergoes an unfolding transition upon removal of bound EDTA. Addition of EDTA, or a variety of other small

anions (sulfate is shown in Figure 2A), restores the "native" mixed α/β protein CD spectrum previously observed. The P protein therefore undergoes a major structural rearrangement upon removal of EDTA (and other anions) that can be reversed by addition of these same anions. The unliganded P protein does not contain the secondary structural characteristics expected for a folded α/β protein as assayed by far-UV CD. Using the k2d algorithm (*48*), the unliganded P protein is predicted to be 6% helix, 37% sheet, and 56% other while in the reported crystal structure (*30*) the P protein is composed of 38% helix, 22% sheet, and 40% other. The k2d algorithm predicts that the P protein refolded in 20 mM sulfate is 30% helix, 17% sheet, and 53% other.

To determine whether the P protein maintains its function as a cofactor in RNase P, the catalytic activities of the reconstituted holoenzyme (P RNA and P protein) and the P RNA ribozyme under single-turnover conditions were assayed in low salt (10 mM $MgCl_2$ and 100 mM KCl). The holoenzyme reconstituted with protein purified using the method described above had the same single-turnover cleavage rate constant as that prepared by the published method (*41*) (data not shown), indicating that there were no deleterious effects due to exposure of the protein to 5% acetic acid or to removal of EDTA.

Near-UV CD monitors changes in the stereochemical environment of aromatic residues (*49*). The near-UV wavelength CD spectra of the unliganded P protein and P protein refolded with 10 mM sulfate are shown in Figure 2B. There is a marked change in the wavelength trace in this region upon addition of the sulfate anion as evidenced by the differences in the spectra. Furthermore, the difference spectrum (data not shown) resembles the trace expected for a change in the hydrophobic environment of tyrosine residues.

Two-dimensional $^1H-^{15}N$ HSQC spectra for the P protein in the absence and presence of 20 mM sulfate in cacodylate buffer are shown in panels A and B of Figure 3, respectively. In an HSQC experiment, each cross-peak corresponds to an amide proton and its covalently attached nitrogen (*50*). Each residue in the protein (except for prolines) should therefore give rise to an observable cross-peak due to the peptide backbone (other cross-peaks arise due to amide- and amino-containing side chains). As evident from the spectra shown in panels A and B of Figure 3, the unliganded P protein has a larger degree of chemical shift overlap while the spectrum of the P protein with 20 mM sulfate clearly contains more chemical shift dispersion. These results are consistent with the conclusion that sulfate induces a folding transition. Interestingly, the unliganded P protein HSQC spectrum has fewer than the expected number of cross-peaks (see the Discussion).

To further examine the effects of anions, the conformational stability of the P protein was monitored by thermal denaturation from 2 to 90 °C (Figure 4). The observed far-UV CD signal of the unliganded P protein remains low at all temperatures, while the signal of the P protein refolded with 20 mM sulfate undergoes a cooperative unfolding thermal transition with a midpoint of 58.4 °C. Thus, the removal of RNA, sulfate, and/or other anions greatly destabilizes the folded, native state of the P protein.

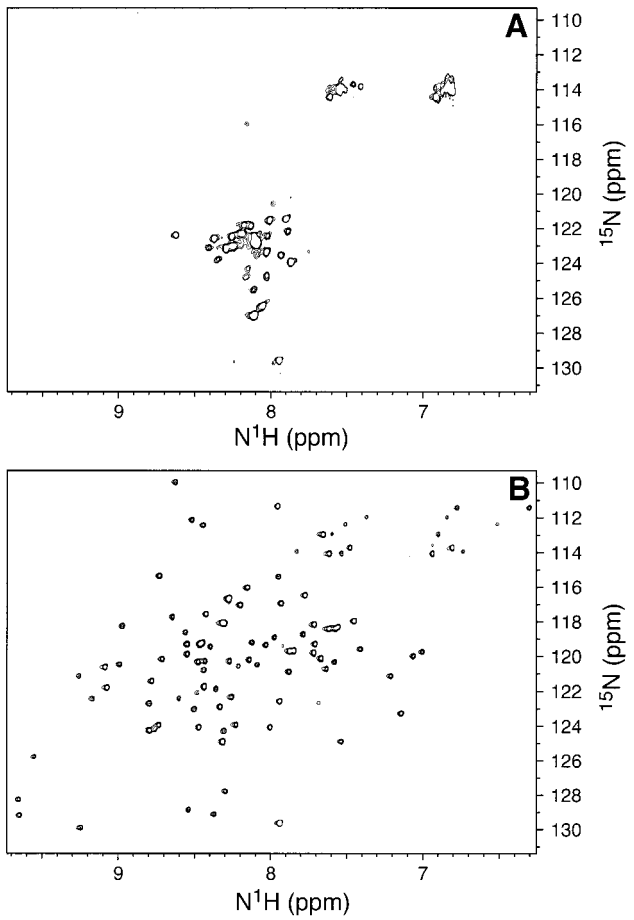


FIGURE 3: $^1H-^{15}N$ HSQC spectra of the P protein in the absence and presence of 20 mM sulfate are shown in panels A and B, respectively. The protein sample in panel A contained 25 μM ^{15}N-labeled P protein in cacodylate buffer and 10% D_2O, while the sample in panel B contained 250 μM ^{15}N-labeled P protein and 20 mM sodium sulfate in the same buffer conditions as for panel A.

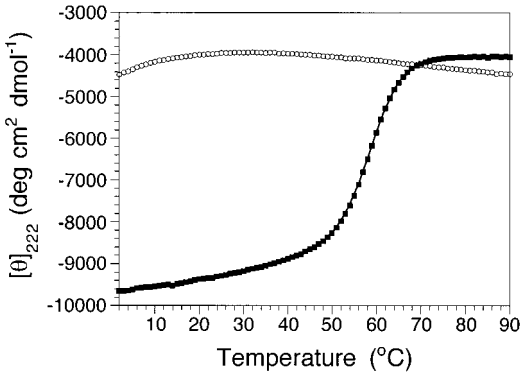


FIGURE 4: Far-UV CD thermal melt of the P protein (10 μM) in cacodylate buffer in the presence (■) and absence (○) of 20 mM sulfate. The solid line is a nonlinear least-squares fit of the data using eqs 1 and 2; the fit melting point (T_m) for the 20 mM sulfate curve is 58.4 ± 0.1 °C.

Anion-Dependent Folding of the P Protein

Anion Survey. The P protein is a basic protein having a predicted pI of 10.2 based on amino acid composition (*41*); therefore, the destabilization of the P protein in the low-ionic strength buffers used here is not surprising. There are three principal means through which the addition of anions can induce the folding of the P protein (reviewed in ref *51*).

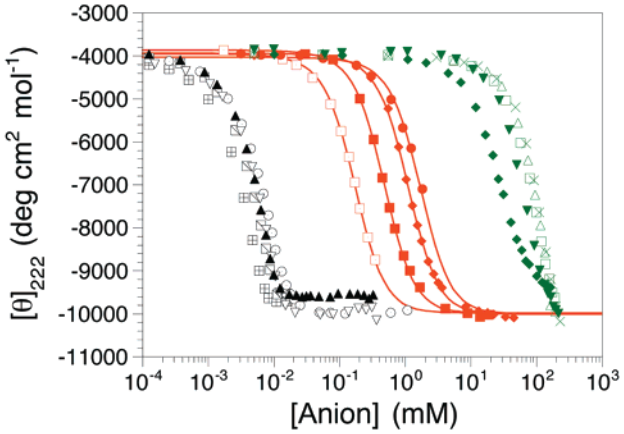


FIGURE 5: Anion-induced folding transitions of the P protein followed by the change in the far-UV CD signal at 222 nm at 37 °C. Protein samples (5 μM) were initially equilibrated in a 10 mM sodium cacodylate buffer (pH 7). Anions used in the titration include tetrapolyphosphate (⊞), tripolyphosphate (⧅), pyrophosphate (⊙), CTP (▽), dCTP (▲), CMP (red ⊡), dCMP (red ●), phosphate (red ◆), sulfate (red ■), acetate (green △), fluoride (green ×), chloride (green ▼), perchlorate (green ◆), and formate (green □). The total anion concentration is shown on the *x*-axis. The lines represent nonlinear least-squares fits of the intermediate affinity anion titration data to eqs 8−12, and the parameters are listed in Table 3.

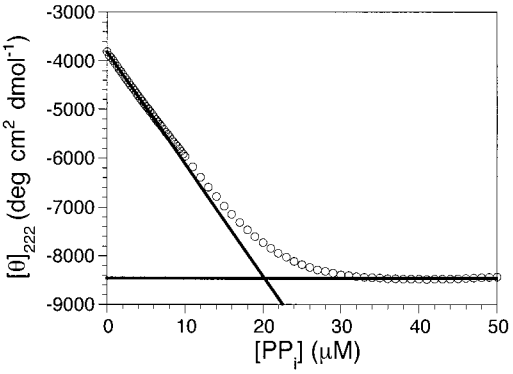


FIGURE 6: Stoichiometry of binding of pyrophosphate to the P protein determined from the binding curve based on the CD signal at 222 nm as a function of total pyrophosphate (PP_i) concentration. The P protein concentration is 10 μM. Lines representing the initial change in signal (when [P protein] ≥ [PP_i]) and the saturated baseline signal ([PP_i] ≥ [P protein]) intersect at the pyrophosphate concentration needed to occupy all of the anion-binding sites of the P protein. This shows that the stoichiometry of binding of PP_i to P protein is 2:1.

First, anions can induce folding through a general ionic strength effect whereby the addition of anions weakens the expected positive charge−charge repulsions present in a compact folded structure. Second, anions can affect the stability of proteins through a Hofmeister effect; addition of anions (and cations) can either stabilize or destabilize proteins through their ability to alter the structure of water (*52*). Last, anions can induce protein folding through the specific binding of the anion to the protein native state conformation, thereby shifting the folding equilibrium toward the native state. Specific anion binding has been shown to induce the refolding of acid-denatured proteins (*51*, *53*−*55*).

To ascertain which of these mechanisms is responsible for the anion-induced folding of the P protein, the far-UV CD signal of the P protein at 222 nm was measured as a function of the concentration of a number of anions. Figure 5 shows that a wide variety of anions can induce the folding of the P protein, although the anion concentration at the midpoint of the folding transition varies greatly. These anions can be divided into different classes on the basis of the concentration required to fold the protein. A number of anions bind stoichiometrically and induce folding at 5−15 μM, near the P protein concentration (5 μM); intermediate-affinity anions stabilize the P protein native structure in the range of 1−50 mM, and the weakly binding anions require ≥100 mM to stabilize the folded P protein. The folding curves of several chloride salts (NaCl, KCl, $MgCl_2$, and $CaCl_2$) overlap when corrected for the total chloride concentration, suggesting that the identity of these cations is not responsible for the observed salt-induced folding (data not shown).

The effectiveness of an anion at stabilizing the folded P protein is enhanced as the negative charge of the anion increases (Figure 5). However, the wide range of anion concentrations at the midpoint of the folding transition (≤4 μM to 100 mM) is more than would be expected if a simple ionic strength effect is the sole cause of folding [the ionic strength varies only 10-fold between L^- and L^{4-} (*51*)]. Furthermore, the order of the affinities for the intermediate and weak anions for P protein follows the anion selectivity series. This series corresponds to the binding of these anions onto anion exchange resins (*56*):

$$SO_4^{2-} > ClO_4^- > H_2PO_4^- > I^- > Br^- > Cl^- > OAc^- = F^- \quad (5)$$

This order of affinities has also been observed for acid-induced protein folding of apomyoglobin and cytochrome *c* (*51*). These results suggest that weak and intermediate anions induce the folded structure of the P protein primarily by binding specifically to the folded conformation of the P protein. Further support for this mechanism comes from the fact that high-affinity anions fold the P protein in the stoichiometric limit of the experiment.

Stoichiometry. The stoichiometry of binding pyrophosphate, a member of the tight binding class of anions, to the P protein was determined by following the CD signal at 222 nm as a function of the total anion concentration (Figure 6). In the stoichiometric limit, when the P protein is in excess and its concentration above the K_d ([P protein] = 10 μM), all of the added pyrophosphate binds to the protein. This results in an initial linear dependence of the CD signal on ligand concentration, starting from an initial value corresponding to that of the deliganded protein. Further addition of pyrophosphate causes a further change in the CD signal until the signal reaches the saturated value corresponding to that of the fully folded protein. The intersection of the initial line extrapolated to the fully saturated signal represents the anion concentration at which the high-affinity P protein anion binding sites are saturated. This intersection occurs at 20 μM pyrophosphate, yielding a stoichiometry of two ligand molecules per P protein molecule.

Two-State Ligand-Induced Folding. Far-UV CD spectra of P protein samples containing various concentrations of sulfate ranging from 0 to 20 mM are shown in Figure 7A. An isodichroic point is observed at 204 nm, which suggests that anion-induced folding is a two-state process such that only the spectra of the unliganded, non-native protein and

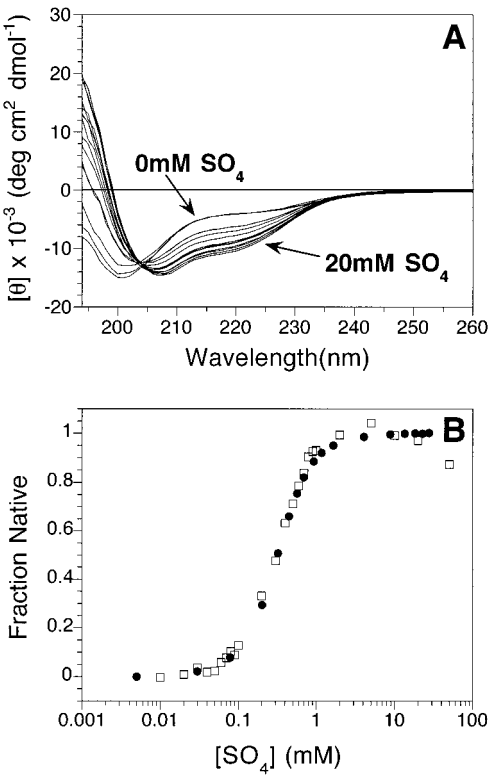


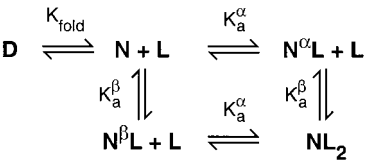
FIGURE 7: Evidence that P protein folding and binding is two-state. (A) Far-UV CD spectra of 10 μM P protein in cacodylate buffer and the following concentrations of sodium sulfate: 0.00, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 2.0, 5.0, 10.0, and 20.0 mM. An isodichroic point is observed at 204 nm. (B) Sodium sulfate-induced folding of the P protein followed by the near-UV ($[\theta]_{270}$, □) or far-UV ($[\theta]_{222}$, ●) CD signal showing coincident renaturation curves.

the liganded, native protein contribute to the observed CD spectra. However, an isodichroic point can also be observed if a third component has a spectrum that is identical to one of the other two states. Therefore, more evidence supporting two-state folding is required. An additional test of two-state folding is the coincidence of de(re)naturation curves (F_n vs perturbant) when using various physical probes to monitor the folding progress (*57*). Figure 7B shows the sulfate-induced folding of the P protein monitored by far-UV (222 nm) and near-UV (270 nm) CD, which are primarily probes of secondary and tertiary structure, respectively; these cooperative folding curves overlap. Taken together, the overlapping folding and binding curves in Figure 7B and the isodichroic point in Figure 7A provide strong evidence that anion-induced folding of the P protein is a two-state process. Thus, under the equilibrium conditions of these experiments, there are only two significantly populated states: the unliganded, predominantly unfolded state and the ligand-bound native state.

Model of Coupled Folding and Ligand Binding

These data demonstrate that the binding of anions to the P protein is tightly coupled to folding, and this linked process can be modeled as shown in Scheme 1. This scheme is intended to apply to high-affinity ligands (K_d < 1 mM), which presumably bind preferentially to the native state. However, it should be noticed that formally, D in Scheme 1 could include both liganded and unliganded forms of the

Scheme 1



denatured protein (*1*). In the treatment that follows, we assume that tight binding occurs only to the native binding sites. Further work will be required to determine the kinetic and thermodynamic role of ligand binding to D, which, if significant, will require a much more complicated model. Given an apparent two-state process and the assumption that the anion-binding sites are independent, the predicted effect of ligand concentration on the overall apparent folding and binding equilibrium constant, $K_{app}([L])$, is given in eqs 6 and 7.

$$K_{app}([L]) = \sum[N]_i/\sum[U]_i \quad (6)$$

$$= ([N] + [N^{\alpha}L] + [N^{\beta}L] + [NL_2])/[U]$$

$$= ([N]/[U])(1 + K^{\alpha}_{a}[L] + K^{\beta}_{a}[L] + K^{\alpha}_{a}K^{\beta}_{a}[L]^2)$$

$$= K_{fold}(1 + K^{\alpha}_{a}[L])(1 + K^{\beta}_{a}[L]) \quad (7)$$

where K_{fold} is the folding equilibrium constant, K^{α}_{a} and K^{β}_{a} are the microscopic association constants of the ligand for the P protein at binding sites α and β, respectively, and [L] is the free ligand concentration. Numerical analysis of the general equation for two independent binding sites given in eq 7 indicates that unless K^{α}_{a} and K^{β}_{a} differ by more than a factor of 50, they cannot be distinguished. Our initial analysis of sulfate binding data (see below) indicates that K^{α}_{a} and K^{β}_{a} for the two anion binding sites in the P protein differ by ≤50, so we further simplify the model by assuming that $K^{\alpha}_{a} \sim K^{\beta}_{a} \sim K_a$. This yields the following equation for the apparent overall equilibrium constant, $K_{app}([L])$:

$$K_{app}([L]) = K_{fold}(1 + K_a[L])^2 \quad (8)$$

where K_a is the intrinsic microscopic association constant for both sites. $K_{app}([L])$ and $F_n([L])$, the fraction native protein, are measured directly from the CD titrations (eqs 9 and 10)

$$K_{app}([L]) = [\theta([L]) - B_d]/[B_n - \theta([L])] \quad (9)$$

$$F_n([L]) = K_{app}([L])/[1 + K_{app}([L])] \quad (10)$$

where B_d and B_n represent the CD signal of the unliganded, unfolded protein and the folded, liganded protein, respectively, and $\theta([L])$ is the observed signal at any given ligand concentration. These equations assume that there is no dependence of the CD signal on ligand concentration. In this model, the free ligand concentration is

$$[L] = [L]_{tot} - [NL] \quad (11a)$$

$$[NL] = \{(2[P]_{tot} + [L]_{tot} + K) - [(2[P]_{tot} + [L]_{tot} + K)^2 - 8[P]_{tot}[L]_{tot}]^{1/2}\}/4 \quad (11b)$$

where

$$K = [(K_{fold} + 1)/K_{fold}K_a] \quad (12)$$

$[P]_{tot}$ and $[L]_{tot}$ are the total protein and ligand concentrations, respectively. The model assumes that the CD signal of each native form of the protein (i.e., N, $N^{\alpha}L$, $N^{\beta}L$, and NL_2) is identical. Further work is underway to rigorously test these assumptions. The utility of the simple proposed scheme is that it successfully models all of the anion-induced folding data we have collected thus far. In particular, even this simple model predicts the pseudocooperativity that results from the shifting of the unfavorable D → N equilibrium shown in Scheme 1 by the first ligand binding event. This effect can be reproduced using identical microscopic binding constants for all of the ligand binding events depicted in Scheme 1.

Since binding and folding are tightly coupled (eq 8), K_{fold} and K_a cannot be independently determined from the ligand CD titration data. For the P protein, K_{fold} is so small that no significant folding is detected until the ligand concentration is $\sim 3K_a$, and under these conditions, $K_{app} \approx K_{fold}K_a[L]^2$, making K_{fold} and K_a impossible to separate. However, if we could determine either K_{fold} or K_a independently using other experiments, then we could determine the other by fitting K_{app} versus [L] to eqs 8 and 9. As described below, we have used the osmolyte TMAO to induce folding of the P protein in the absence of ligand. Extrapolation of the linear free energy versus [TMAO] curve allows us to estimate K_{fold}. With an estimate of K_{fold} in hand, the above model can be used to determine the intrinsic ligand association constant.

Osmolyte-Induced Folding of the P Protein, Uncoupling K_{fold} and K_a

Osmolytes are naturally occurring low-molecular weight organic molecules that stabilize proteins. Three classes of osmolytes are known: polyols, certain amino acids, and methylamines (*58*). Of particular interest is the osmolyte trimethylamine *N*-oxide (TMAO), found in cartilaginous fish and coelacanth, which has been shown to counteract the detrimental effect of urea upon protein activity (*59*). Work of Bolen and co-workers (*44*) has suggested that the primary source of protein stabilization by TMAO is an unfavorable interaction with the peptide backbone that increases the energy of the denatured state. Furthermore, the stabilizing effect of TMAO appears to be independent and additive to the destabilizing effects of urea on protein stability (*44*, *60*). TMAO also causes the contraction of denatured ensembles (*61*) which both lowers the conformational entropy of folding and brings aliphatic side chains closer together, allowing hydrophobic collapse.

Figure 8A displays the effect of TMAO on P protein folding as followed by far-UV CD at 222 nm at pH 8 and various temperatures. Addition of TMAO to the unliganded P protein apparently induces cooperative refolding to a state with a CD signal similar to that of the ligand-bound P protein (see Figure 2A). The titration curves collected at various temperatures were fit to a two-state folding model (eq 3). Table 2 lists the thermodynamic parameters obtained from these fits. As expected, the free energy of folding becomes less favorable at increasing temperatures. The free energy of folding for the P protein at 37 °C using a nonlinear least-squares fit is 3.05 ± 0.04 kcal/mol, which yields an equilibrium constant of folding (K_{fold}) of 0.0071 ± 0.0005. Figure 8B shows the linearity of the free energy of folding versus [TMAO] curves using the baseline parameters derived from the global fit analysis (eq 3). The observation that the linear extrapolation method, which is based on the two-state folding model, fits the data quite well over a wide range of TMAO concentrations suggests that there are no populated intermediates in the TMAO-induced refolding equilibria of the P protein.

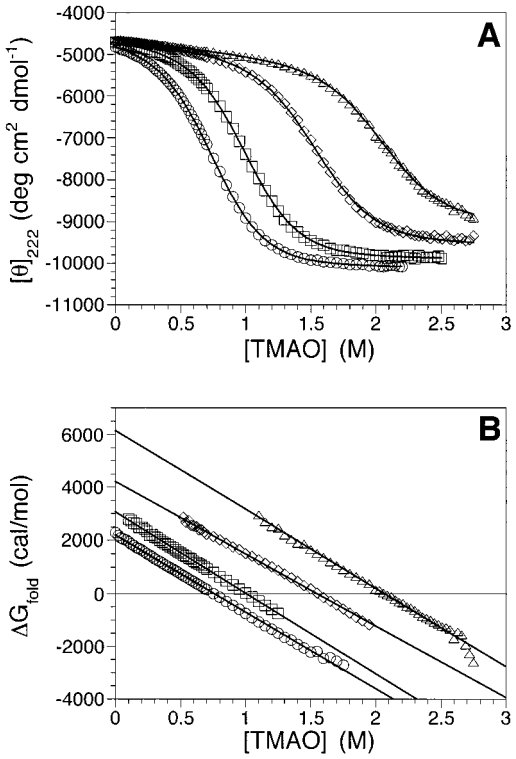


FIGURE 8: TMAO can induce the P protein to fold in the absence of ligand. (A) CD signal at 222 nm of the P protein (10 μM) in 10 mM Tris cacodylate (pH 8) followed at four different temperatures: 28 (○), 37 (□), 50 (◇), and 60 °C (△). Solid lines represent the nonlinear least-squares fit to the data using the linear extrapolation model assuming two-state folding (eq 3). (B) The same data converted to a $\Delta G_{folding}$ vs TMAO concentration profile.

Also shown in Table 2 are the best-fit *m* values ($\delta\Delta G_{fold}/\delta$[TMAO]), which are identical within experimental error over the temperature range that was studied. It is interesting to note that the average *m* value of TMAO-induced refolding is 2.90 ± 0.14 kcal mol^{-1} M^{-1}, which is about equal to and opposite in sign from the value expected for guanidinium chloride denaturation for a protein the size of the P protein (*62*). Thus, TMAO is a powerful renaturant of P protein and may be generally useful in the study of other intrinsically unstructured proteins.

With an estimate of K_{fold}, the intrinsic affinities of the P protein for some anions (Figure 5) can be calculated. Table 3 lists the intrinsic dissociation constants of the intermediate affinity anions for the P protein calculated from a fit to the model (eqs 8−12) using the K_{fold} derived from the TMAO titration data. The intermediate affinity anion class was the only set of anions that was fit in Figure 5 because (1) the free ligand concentration cannot accurately be determined for the stoichiometric binding anions and (2) other effects (i.e., general ionic screening or Hofmeister effects) may sufficiently contribute to the anion-induced folding of the P

Table 2: CD-Detected TMAO Titration Curves[a]

temperature (°C)	$[TMAO]_{1/2}$ (M)[b]	m (kcal mol^{-1} M^{-1})[b]	$\Delta G_{D\rightarrow N}$ (kcal/mol)	$K_{foldD\rightarrow N}$
28	0.75 ± 0.003	2.96 ± 0.04	2.22 ± 0.03	0.024 ± 0.001
37	1.006 ± 0.004	3.05 ± 0.04	3.05 ± 0.04	0.0071 ± 0.0005
50	1.554 ± 0.009	2.80 ± 0.07	4.35 ± 0.11	0.0011 ± 0.0002
60	2.07 ± 0.03	2.9 ± 0.4	6.0 ± 0.4	(1.2 ± 0.7) × 10^{-4}

[a] Determined in 10 mM Tris cacodylate (pH 8). [b] Values determined by fitting the data in Figure 8 to eq 3. The TMAO-dependent slopes of the denatured (m_D) and native (m_N) baselines were fixed at −393 and −62.8 deg cm^2 dmol^{-1} M^{-1}, respectively, based on fits to the denatured and native baselines of curves obtained at the higher temperatures (averaged) and the lower temperatures (averaged), respectively.

Table 3: Intrinsic Binding Constants for the Intermediate Affinity Anions[a]

anion	K^b (mM^{-1})	K_d (μM)[c]	anion	K^b (mM^{-1})	K_d (μM)[c]
SO_4^{2-}	6.06 ± 0.08	43 ± 3	CMP^{2-}	2.17 ± 0.08	15 ± 1
HPO_4^{2-}	14.0 ± 0.2	99 ± 7	$dCMP^{2-}$	21.6 ± 0.6	153 ± 12

[a] Initial conditions were 5 μM P protein, 10 mM sodium cacodylate buffer, pH 7, and 37 °C. [b] Based on the global fit (eqs 8−12) of the intermediate ligand-induced CD titration data depicted in Figure 5; definition of K given in eq 12. [c] Determined using eq 12 where a value of 0.0071 is used for the equilibrium constant of folding (K_{fold} at 37 °C).

protein for the weak affinity ligands. The resulting dissociation constants for these simple anions are in the mid-micromolar range (10−150 μM; see Table 3). Similar anion dissociation constants have been observed for the simple anions used to refold acid-denatured proteins (*51*).

Structures of TMAO-Folded and Ligand-Folded States Are Similar

This model assumes that the osmolyte-folded state of the P protein represents the unliganded, natively folded conformation of the P protein (eqs 6 and 7). Panels A and B of Figure 9 display the far- and near-UV CD wavelength traces, respectively, of both the ligand-folded state (folded with sulfate) and TMAO-folded state. There is extensive spectral overlap in both of the CD spectra, suggesting that there are nearly identical degrees of secondary structure as well as similar aromatic chiral environments for the two folded states. The HSQC spectrum of the P protein folded by TMAO is shown in Figure 9C. Comparison of the sulfate-folded (Figure 3B) and the osmolyte-folded HSQC spectra supports the proposal that the two conformations maintain similar backbone folds because there are several spectral regions that have cross-peak pattern similarity. However, there are some apparent differences between the two spectra that may arise from differing changes in the P protein chemical shift induced by TMAO and sulfate. In aggregate, these data suggest that the TMAO-folded and anion-folded conformations of the P protein are similar.

TMAO Does Not Bind to the High-Affinity Anion Binding Sites

Some or all of the stabilization of the P protein by TMAO could be the result of TMAO binding to the anion binding sites. If this were the case, then at molar concentrations TMAO would act as a competitive inhibitor of the binding of other ligands. To test for such competition, we assessed sulfate-induced folding of the P protein in the presence and absence of 1 M TMAO in 10 mM Tris cacodylate (pH 8) buffer, shown in Figure 10. In the absence of TMAO and

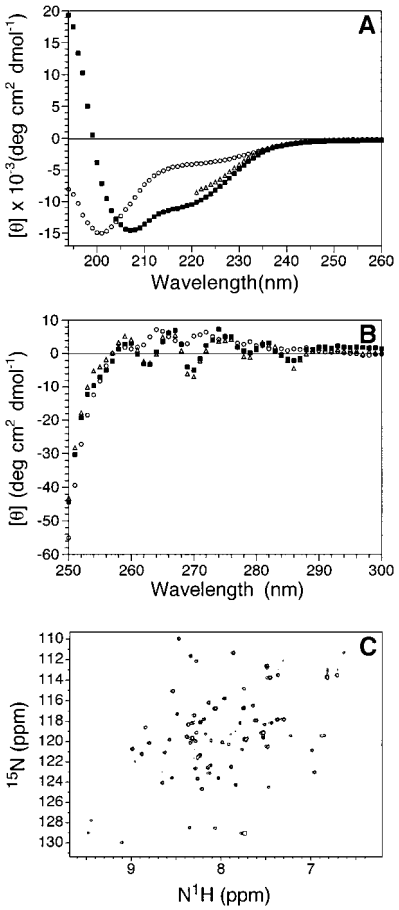


FIGURE 9: Comparison of the unliganded P protein, the anion-induced folded P protein, and the osmolyte-folded P protein. (A) Far-UV CD wavelength scans of the P protein (10 μM) in 10 mM sodium cacodylate buffer (pH 7) without (○) and with 20 mM sulfate (■) (as in Figure 2A). P protein folded with 2 M TMAO (△) in 10 mM Tris cacodylate (pH 8). (B) Near-UV CD traces of the P protein (25 μM; 10 cm cuvette) in 10 mM cacodylate, pH 7 buffer in the unliganded state (○) and the liganded state with 10 mM sulfate (■) (as in Figure 2B). Spectrum of the P protein (25 μM) folded in 2.1 M TMAO, 10 mM Tris cacodylate (pH 8) buffer (△). (C) ^{1}H−^{15}N HSQC of the P protein folded in 2 M TMAO. The sample shown in panel C contained 25 μM ^{15}N-labeled P protein dissolved in 10 mM sodium cacodylate (pH 7.4) and 10% D_2O.

sulfate, the CD signal is that of the denatured protein and addition of sulfate produces a complete renaturation curve. In the presence of 1 M TMAO and the absence of sulfate, the protein has a CD signal that is 67% of the way to that of the fully native value, in approximate agreement with the data listed in Table 2. Using this to estimate K_{fold} under these conditions (K_{fold} = 2.03), and combining it with the K_a [K_a = (57.5 μM)$^{-1}$] deduced from fitting the sulfate titration data in the absence of TMAO, we have used eqs 8−10 to simulate

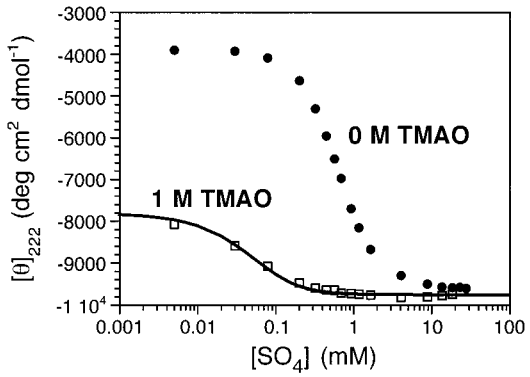


FIGURE 10: TMAO does not compete with sulfate for binding to the high-affinity anion binding sites. The CD signal of 10 μM P protein in 10 mM Tris cacodylate titrated with sodium sulfate at pH 8, 37 °C in 0 M TMAO (●) and 1 M TMAO (□). The line represents a simulation of the binding curve in the presence of TMAO using a K_{fold} of 2.03 and a K_a of (57.5 μM)$^{-1}$. The latter value is the same intrinsic affinity measured in the absence of TMAO, assuming $K_{fold} = 0.0071$.

a sulfate binding curve, shown as the line in Figure 10. There is excellent agreement between the 1 M TMAO experimental data and the binding curve simulated using the sulfate binding constant obtained from the 0 M TMAO data. The observation that intrinsic sulfate binding affinity (K_a) is unperturbed by the presence of 1 M TMAO is strong evidence that the mechanisms of TMAO-induced folding and anion-induced folding are independent and that TMAO does compete with anions for the two high-affinity anion binding sites. In fact, had TMAO competed for sulfate binding, this would have shifted the binding curve in Figure 10 to the right rather than to the left. We have also tested the activity of the RNase P holoenzyme in the presence of 2 M TMAO and find no reduction in activity (data not shown.)

DISCUSSION

The P Protein Is Intrinsically Unstructured in Low-Ionic Strength Buffer. We have carried out several biophysical analyses on the *B. subtilis* RNase P protein to characterize the basic thermodynamics of its folding and ligand binding reactions. The ultimate goal of this study is to better describe the thermodynamics of a ribonucleoprotein assembly reaction. The complete purification of the P protein results in a conformational state that is not consistent with the compact, "native", mixed α/β fold that has been identified by both X-ray crystallography (*30*) and NMR solution structure (*21*) studies. Rather, when analyzed in a low-ionic strength buffer [10 mM sodium cacodylate (pH 7)], the P protein contains several characteristics of a "natively unfolded" protein (*63*). Specifically, the unliganded P protein maintains a conformation containing very little secondary structure as detected by far-UV CD; furthermore, this structure does not melt at elevated temperatures (*3*, *63*). The P protein has a number of amino acids that are positively charged at neutral pH which is also observed in several other natively unfolded proteins (*5*, *64*). However, the unliganded P protein also possesses some characteristics of a molten globule species, which are also often characterized by a large net positive charge (*65*). The HSQC spectrum of the unliganded P protein does not contain the expected number of cross-peaks; more cross-peaks are observed when the P protein is placed in a 6 M urea solution (data not shown) or refolded with 20 mM sulfate (Figure 3). The apparent loss of cross-peaks in the unliganded P protein spectrum may be due to conformational exchange on the time scale of the NMR experiment. The loss of observable cross-peaks relative to the folded state is a common characteristic of molten globule proteins (*65*). Furthermore, the weak negative far-UV CD signal may be consistent with an ensemble of conformations containing residual β-sheet, which is known to give rise to weak far-UV CD signals (*66*). The presence of a non-zero near-UV CD signal also suggests that the unliganded protein is hydrophobically collapsed, another characteristic of molten globule conformations.

Anion-Induced Folding of the P Protein. Regardless of its designation as natively unfolded or molten globule, the unliganded P protein can be induced to fold to the native α/β conformation through the addition of anions. The P protein can be induced to fold over a large range of anion concentrations, depending on the identity of the anion. Furthermore, the effective order of anions that induce protein folding follows the anion binding series and not the Hofmeister series, suggesting that specific ionic interactions, as opposed to solvation effects, are critical for P protein folding in low-ionic strength buffer. These results best support a model whereby the specific binding of anions to the protein stabilizes the folded conformation of the P protein, thereby shifting the folding equilibrium toward the native α/β fold (Scheme 1). A similar anion-induced folding transition to the native conformation has been shown for the refolding of acid-denatured Sac7d (*67*).

The stoichiometry of pyrophosphate binding to the P protein is two, although it is not known whether one or both anion-binding sites are necessary for folding. There are extensive regions of basic charge on the surface of the P protein that are potential anion binding sites. However, both the NMR and crystal structures of the P protein identify two specific anion binding regions. In the solution structure, Spitzfaden et al. (*21*) localize nucleotide binding sites to the crevice formed with the NH_2-terminal helix and the central β-sheet as well as the loop between β-strand 3 and α-helix 2 (the RNR motif). The crystal structure of the P protein reveals a bound sulfate molecule between α-helix 2 and the N-terminus (H3, R9, and R68) (*30*). Subsequent analysis has identified a second possible sulfate and/or phosphate binding site at the position formerly interpreted as solvent molecule HOH8, which is adjacent to R88 and K89 (J. D. Cox and D. Christianson, personal communication). These sites may reflect the positions of the bound anions that stabilize the folded state of the P protein (see Figure 1).

To understand the ligand titration data, it is important to identify all significantly populated ensembles and the various equilibria that interconvert them. On the basis of this work, the coupled process of P protein folding and binding is apparently two-state as a function of ligand with only two ensembles present under equilibrium conditions: the unliganded, predominantly unfolded state and the liganded, folded conformation. The folded conformation is thermodynamically favored only in the presence of ligand. Otherwise, the P protein is predominantly unfolded; all other states have higher energy. This folding−binding reaction can be considered an extreme case of an "induced fit" mechanism wherein an otherwise unfavorable conformation (the native

state) is induced in response to the binding of a partner. Given the degree of coupling between the two equilibria, it is impossible to separate the folding equilibrium constant (K_{fold}) from the ligand association constants (K_a's) directly from ligand titration data.

To separate the thermodynamics of binding and folding, we used the small organic osmolyte TMAO as a means of determining the equilibrium constant of folding (K_{fold}). Titration of the P protein with TMAO resulted in a single cooperative folding transition and indicates that the folded state is 3 kcal/mol less stable than the unfolded state at 37 °C in 10 mM sodium cacodylate at neutral pH. Cooperative folding curves as a function of TMAO have been observed previously for several proteins, including mutant peptides of apoC-1 (*68*), thermodynamically unfolded variants of RNase T1, and staphylococcal nuclease (*69*), as well as the intrinsically unfolded transactivation domain of human glucocorticoid receptor (*70*). Using the calculated value for K_{fold}, we were able to uniquely determine that the intermediate affinity anionic ligands have dissociation constants in the 10−150 μM range. The simplified model used to fit these data assumes that the ligand binding sites are independent and identical. However, it is likely that the binding sites are not identical given the asymmetry of the P protein and the fact that the two sulfate molecules identified in the crystal structure interact with different numbers of amino acids. Further experiments are required to determine whether the calculated dissociation constants for the various ligands are accurate.

Far- and near-UV CD, as well as two-dimensional NMR evidence, suggests that the TMAO-folded state of the P protein is similar structurally to the anion-induced folded form of the protein. This result is striking when one considers the fact that the unliganded, folded state can be considered a metastable state in the folding−binding pathway. The structural similarity between the two folded states is also surprising since the molecular mechanisms for P protein folding are different for the two solvent additives. TMAO stabilizes the native state of proteins due to an unfavorable interaction with the peptide backbone, while anions favor the folded conformation by directly binding the native state. Further experiments are underway to discern the structural differences between these two folded states of the P protein.

Implications for RNase P Assembly. We have demonstrated that the conformational stability of the *B. subtilis* P protein is enhanced by increasing salt concentrations as anions preferentially bind to the native state of the P protein (Figure 4). This phenomenon has also been observed for the *Staphylococcus aureas* P protein wherein the fraction of the denatured protein at 1 M urea was found to decrease as the salt concentration increased (*21*). These data suggest that this behavior is shared by a number of bacterial P proteins. This model may therefore provide additional understanding of previous studies of RNase P assembly.

Talbot and Altman (*71*) attributed the increase in the apparent P protein−RNA affinity with increasing salt to a "hydrophobically driven association reaction". Our results with the *B. subtilis* P protein indicate that the salt dependence could reflect an increased population of folded P protein to which the RNA ligand can bind. This suggestion is further bolstered by the striking observation that the observed dependence of the holoenzyme dissociation constant (K_d) on sodium acetate ($K_{1/2}$ = 200 mM) described in Talbot and Altman (*71*) is predicted well by the acetate-induced folding curve depicted in Figure 5. Thus, further thermodynamic studies are needed to determine whether the P protein−P RNA interaction is in fact a hydrophobically driven process.

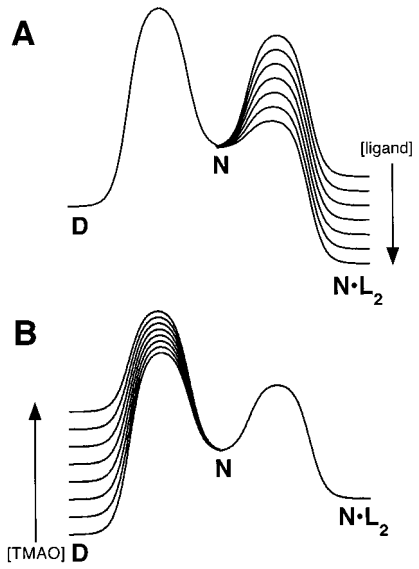


FIGURE 11: Hypothetical free energy reaction diagrams for the P protein depicting the effect of ligand and TMAO concentration on the energetics. N represents the unliganded, native state, which is metastable relative to the two end states, the denatured state and the doubly liganded native state, represented by D and $N{\cdot}L_2$, respectively.

Salt-induced folding of the *E. coli* P protein can also be used to reinterpret the results of Gopalan et al. (*72*), who observed an increased level of immobilization of an EPR spin-label on certain *E. coli* protein variants as a function of acetate concentration. Our results suggest that this change in mobility could be due to the global refolding of the protein induced by the addition of acetate.

Finally, kinetic studies have revealed that the rate constant of association (k_{on}) for formation of the *E. coli* RNase P holoenzyme is 2−3 orders of magnitude below that which is expected for a diffusion-limited process (*71*). This observation may be due to the fact that protein folding represents a rate-limiting step in RNase P assembly.

The Unliganded Native P Protein Is a Metastable State. The reaction diagrams depicted in Figure 11 represent a thermodynamic model for the coupled folding and anion binding reactions of the P protein. The activation barriers in these diagrams are not meant to depict any known kinetics, but merely to separate the three forms of the P protein present in the folding−binding equilibria. Figure 11 shows the effect of addition of ligand or TMAO to the relative energies of D, N, and $N{\cdot}L_2$. As can be seen from these diagrams, the energy of N is usually higher than that of either D or $N{\cdot}L_2$, and thus, it is properly classified as a metastable state. However, the population of N can be increased by selective perturbation of the D or $N{\cdot}L_2$ end states, through the addition of TMAO or the reduction in ligand concentration, respectively. The independence of these two energetic perturbations provides fertile ground for further thermodynamic and kinetic experiments. Such manipulations will allow us to study the structural and energetic properties of a poorly populated but

presumably native-like protein in more detail. The metastable N state of the P protein should prove to be a useful tool for developing methods for studying metastable states in other systems as well.

Relevance of Intrinsic Non-Native Structure of the P Protein. Over the past decade, there have been several proteins that have been identified as natively unfolded or intrinsically unstructured when purified and characterized in vitro (for reviews, see refs *12* and *73*). Folding of these proteins (either locally or globally) is then achieved upon binding of target protein, RNA, or DNA partners. The extent of protein conformational change varies greatly in these systems ranging from a local alteration in the DNA-binding domains (*11*, *74*, *75*) or the helix-induced formation of transactivation domains (*76*−*78*) to the formation of domains upon binding macromolecular targets (*2*, *9*, *10*) and even the complete folding of a protein from an initially unfolded or a molten globule state (*7*, *79*). Proteins that undergo binding-induced structural changes cover a broad range of biological functionality, including, but not exclusively, activation or inhibition of gene transcription (*2*, *8*, *11*, *80*), cell cycle regulation (*3*), cell−cell interaction (*4*), and the formation of the Alzheimer's disease amyloid plaque precursor (*63*). There have been several proposed reasons for maintaining an intrinsically unstructured conformation prior to ligand binding: (1) to improve the protein binding specificity and order of macromolecular assembly, (2) to increase the diversity in protein recognition of various substrate targets, and (3) to allow for an additional means of intracellular protein regulation via proteolysis. However, it is important to note that these suggestions are based on in vitro observations that may not truly mimic physiological conditions. These intrinsically unstructured proteins would be folded if the ligand (either specific or nonspecific) capable of refolding the protein is present at concentrations near or above its dissociation constant in the cell. This is likely the case for the *B. subtilis* P protein because it can fold at physiological concentrations of pyrophosphate alone (*81*). The ability of the P protein to fold in the presence of salt, polyphosphates, or nucleotides may prevent intracellular proteolysis until the protein binds its presumed highest-affinity ligand of P RNA. The mechanism of holoenzyme assembly from the presumed P protein−anion complex is not known, including whether the anion must dissociate during assembly or whether assembly requires partial or complete unfolding of the P protein. However, the ability of the P protein to unfold and refold in the absence and the presence of ligand may ultimately enhance the specificity of the RNase P holoenzyme interaction and/or be important in the assembly process.

ACKNOWLEDGMENT

We thank Dr. Ronald Venters for technical assistance with the NMR experiments, Dr. Niranjanakumari for the gift of pPWT-28b and advice on the preparation of the P protein, Dr. David Christianson and J. David Cox for sharing their insight into potential anion binding sites, and Dr. Gordon Hammes and Dr. Jeffrey Myers for helpful advice.

REFERENCES

1. Wyman, J., and Gill, S. J. (1990) *Binding and Linkage: Functional Chemistry of Biological Macromolecules*, University Science Books, Mill Valley, CA.
2. Daughdrill, G. W., Chadsey, M. S., Karlinsey, J. E., Hughes, K. T., and Dahlquist, F. W. (1997) *Nat. Struct. Biol. 4*, 285−291.
3. Kriwacki, R. W., Hengst, L., Tennant, L., Reed, S. I., and Wright, P. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 11504−11509.
4. Penkett, C. J., Redfield, C., Dodd, I., Hubbard, J., McBay, D. L., Mossakowska, D. E., Smith, R. A., Dobson, C. M., and Smith, L. J. (1997) *J. Mol. Biol. 274*, 152−159.
5. Lydakis-Simantiris, N., Hutchison, R. S., Betts, S. D., Barry, B. A., and Yocum, C. F. (1999) *Biochemistry 38*, 404−414.
6. Liu, D., Ishima, R., Tong, K. I., Bagby, S., Kokubo, T., Muhandiram, D. R., Kay, L. E., Nakatani, Y., and Ikura, M. (1998) *Cell 94*, 573−583.
7. Fiebig, K. M., Rice, L. M., Pollock, E., and Brunger, A. T. (1999) *Nat. Struct. Biol. 6*, 117−123.
8. Van Gilst, M. R., Rees, W. A., Das, A., and von Hippel, P. H. (1997) *Biochemistry 36*, 1514−1524.
9. Zheng, N., and Gierasch, L. M. (1997) *Mol. Cell 1*, 79−87.
10. Mogridge, J., Legault, P., Li, J., Van Oene, M. D., Kay, L. E., and Greenblatt, J. (1998) *Mol. Cell 1*, 265−275.
11. Spolar, R. S., and Record, M. T., Jr. (1994) *Science 263*, 777−784.
12. Wright, P. E., and Dyson, H. J. (1999) *J. Mol. Biol. 293*, 321−331.
13. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) *Nature 395*, 817−819.
14. Altman, S., and Kirsebom, L. (1999) Ribonuclease P, in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 351−380, Cold Spring Harbor Laboratory Press, Plainview, NY.
15. Frank, D. N., and Pace, N. R. (1998) *Annu. Rev. Biochem. 67*, 153−180.
16. Pace, N. R., and Brown, J. W. (1995) *J. Bacteriol. 177*, 1919−1928.
17. Brown, J. W., and Pace, N. R. (1992) *Nucleic Acids Res. 20*, 1451−1456.
18. Altman, S., and Smith, J. D. (1971) *Nat. New Biol. 233*, 35−39.
19. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) *Cell 35*, 849−857.
20. Chamberlain, J. R., Lee, Y., Lane, W. S., and Engelke, D. R. (1998) *Genes Dev. 12*, 1678−1690.
21. Spitzfaden, C., Nicholson, N., Jones, J. J., Guth, S., Lehr, R., Prescott, C. D., Hegg, L. A., and Eggleston, D. S. (2000) *J. Mol. Biol. 295*, 105−115.
22. Kole, R., Baer, M. F., Stark, B. C., and Altman, S. (1980) *Cell 19*, 881−887.
23. Reich, C., Olsen, G. J., Pace, B., and Pace, N. R. (1988) *Science 239*, 178−181.
24. Kurz, J. C., Niranjanakumari, S., and Fierke, C. A. (1998) *Biochemistry 37*, 2393−2400.
25. Crary, S. M., Niranjanakumari, S., and Fierke, C. A. (1998) *Biochemistry 37*, 9409−9416.
26. Niranjanakumari, S., Stams, T., Crary, S. M., Christianson, D. W., and Fierke, C. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 15212−15217.
27. Peck-Miller, K. A., and Altman, S. (1991) *J. Mol. Biol. 221*, 1−5.
28. Guerrier-Takada, C., and Altman, S. (1984) *Science 223*, 285−286.
29. Westhof, E., Wesolowski, D., and Altman, S. (1996) *J. Mol. Biol. 258*, 600−613.
30. Stams, T., Niranjanakumari, S., Fierke, C. A., and Christianson, D. W. (1998) *Science 280*, 752−755.
31. Ramakrishnan, V., and White, S. W. (1992) *Nature 358*, 768−771.
32. Czworkowski, J., Wang, J., Steitz, T. A., and Moore, P. B. (1994) *EMBO J. 13*, 3661−3668.
33. Brown, J. W. (1998) *Nucleic Acids Res. 26*, 351−352.
34. Vioque, A., Arnez, J., and Altman, S. (1988) *J. Mol. Biol. 202*, 835−848.
35. Talbot, S. J., and Altman, S. (1994) *Biochemistry 33*, 1399−1405.

36. Loria, A., Niranjanakumari, S., Fierke, C. A., and Pan, T. (1998) *Biochemistry 37*, 15466−15473.
37. Oubridge, C., Ito, N., Evans, P. R., Teo, C. H., and Nagai, K. (1994) *Nature 372*, 432−438.
38. Biswas, R., Ledman, D. W., Fox, R. O., Altman, S., and Gopalan, V. (2000) *J. Mol. Biol. 296*, 19−31.
39. Gopalan, V., Baxevanis, A. D., Landsman, D., and Altman, S. (1997) *J. Mol. Biol. 267*, 818−829.
40. Baer, M. F., Wesolowski, D., and Altman, S. (1989) *J. Bacteriol. 171*, 6862−6866.
41. Niranjanakumari, S., Kurz, J. C., and Fierke, C. A. (1998) *Nucleic Acids Res. 26*, 3090−3096.
42. Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.
43. Schellman, J. A. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 115−137.
44. Wang, A., and Bolen, D. W. (1997) *Biochemistry 36*, 9101−9108.
45. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
46. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
47. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) *Nat. Struct. Biol. 2*, 663−673.
48. Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) *Protein Eng. 6*, 383−390.
49. Adler, A. J., Greenfield, N. J., and Fasman, G. D. (1973) *Methods Enzymol. 27*, 675−735.
50. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−189.
51. Goto, Y., Takahashi, N., and Fink, A. L. (1990) *Biochemistry 29*, 3480−3488.
52. Baldwin, R. L. (1996) *Biophys. J. 71*, 2056−2063.
53. Fink, A. L., Calciano, L. J., Goto, Y., Kurotsu, T., and Palleros, D. R. (1994) *Biochemistry 33*, 12504−12511.
54. Goto, Y., Calciano, L. J., and Fink, A. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 573−577.
55. Uversky, V. N., Karnoup, A. S., Segel, D. J., Seshadri, S., Doniach, S., and Fink, A. L. (1998) *J. Mol. Biol. 278*, 879−894.
56. Gjerde, D. T., Schmuckler, G., and Fritz, J. S. (1980) *J. Chromatogr. 187*, 35−45.
57. Ginsburg, A., and Carroll, W. R. (1965) *Biochemistry 4*, 2159−2174.
58. Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D., and Somero, G. N. (1982) *Science 217*, 1214−1222.
59. Yancey, P. H., and Somero, G. N. (1980) *J. Exp. Zool. 212*, 205−213.
60. Lin, T. Y., and Timasheff, S. N. (1994) *Biochemistry 33*, 12695−12701.
61. Qu, Y., Bolen, C. L., and Bolen, D. W. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 9268−9273.
62. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
63. Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A., and Lansbury, J. P. T. (1996) *Biochemistry 35*, 13709−13715.
64. Shutova, T., Irrgang, K., Klimov, V. V., and Renger, G. (2000) *FEBS Lett. 467*, 137−140.
65. Arai, M., and Kuwajima, K. (2000) *Adv. Protein Chem. 53*, 209−282.
66. Schmid, F. X. (1989) Spectral methods of characterizing protein conformation and conformational changes, in *Protein structure: a practical approach* (Creighton, T. E., Ed.) pp 251−285, IRL Press, Oxford, U.K.
67. McCrary, B. S., Bedell, J., Edmondson, S. P., and Shriver, J. W. (1998) *J. Mol. Biol. 276*, 203−224.
68. Gursky, O. (1999) *Protein Sci. 8*, 2055−2064.
69. Baskakov, I., and Bolen, D. W. (1998) *J. Biol. Chem. 273*, 4831−4834.
70. Baskakov, I. V., Kumar, R., Srinivasan, G., Ji, Y. S., Bolen, D. W., and Thompson, E. B. (1999) *J. Biol. Chem. 274*, 10693−10696.
71. Talbot, S. J., and Altman, S. (1994) *Biochemistry 33*, 1406−1411.
72. Gopalan, V., Kuhne, H., Biswas, R., Li, H., Brudvig, G. W., and Altman, S. (1999) *Biochemistry 38*, 1705−1714.
73. Frankel, A. D., and Smith, C. A. (1998) *Cell 92*, 149−151.
74. Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., and Struhl, K. (1990) *Nature 347*, 575−578.
75. Holmbeck, S. M. A., Dyson, H. J., and Wright, P. E. (1998) *J. Mol. Biol. 284*, 533−539.
76. Radhakrishnan, I., Perez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) *Cell 91*, 741−752.
77. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., and Pavletich, N. P. (1996) *Science 274*, 948−953.
78. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J., and Verdine, G. L. (1997) *Science 277*, 1310−1313.
79. Bose, H. S., Whittal, R. M., Baldwin, M. A., and Miller, W. L. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 7250−7255.
80. Bracken, C., Carr, P. A., Cavanagh, J., and Palmer, A. G., III (1999) *J. Mol. Biol. 285*, 2133−2146.
81. Wanner, B. L. (1996) Phosphorous assimilation and the control of the phosphate regulon, in *Escherichia coli and Salmonella: cellular and molecular biology* (Neidhardt, F. C., Ed.) pp 1357−1381, ASM Press, Washington, DC.
82. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI002078Y