Solvation of the folding-transition state in *Pseudomonas aeruginosa* azurin is modulated by metal

COREY J. WILSON,^{1,2} DAVID APIYO,¹ AND PERNILLA WITTUNG-STAFSHED^{1,2,3}

¹Department of Biochemistry and Cell Biology, ²Keck Center for Structural Computational Biology, and

³Department of Chemistry, Rice University, Houston, Texas 77251, USA

(RECEIVED September 13, 2005; FINAL REVISION December 14, 2005; ACCEPTED December 18, 2005)

Abstract

The role of water in protein folding, specifically its presence or not in the transition-state structure, is an unsolved question. There are two common classes of folding-transition states: diffuse transition states, in which almost all side chains have similar, rather low ϕ (ϕ) values, and polarized transition states, which instead display distinct substructures with very high ϕ -values. Apo- and zinc-forms of *Pseudomonas aeruginosa* azurin both fold in two-state equilibrium and kinetic reactions; while the apo-form exhibits a polarized transition state, the zinc form entails a diffuse, moving transition state. To examine the presence of water in these two types of folding-transition states, we probed the equilibrium and kinetic consequences of replacing core valines with isosteric threonines at six positions in azurin. In contrast to regular hydrophobic-to-alanine ϕ -value analysis, valine-to-threonine mutations do not disrupt the core packing but stabilize the unfolded state and can be used to assess the degree of solvation in the folding-transition state upon combination with regular ϕ -values. We find that the transition state for folding of apo-azurin appears completely dry, while that for zinc-azurin involves partially formed interactions that engage water molecules. This distinct difference between the apo- and holo-folding nuclei can be rationalized in terms of the shape of the free-energy barrier.

Keywords: protein structure/folding; conformational changes; stability and mutagenesis; circular dichroism; fluorescence; molecular mechanics/dynamics; kinetics

When considering the nature of protein-folding mechanisms in solution, because of the dominant role of the hydrophobic effect, one must address the role of water. Does water merely induce hydrophobic forces or does the discrete nature of water play a structural role in folding? It is common to consider the role of water implicitly in terms of its bulk dielectric property and interactions with hydrophobic groups of the protein.

Implicit solvation methods, such as continuum solvent models, have been widely adopted in computational studies of folding dynamics (Rhee et al. 2004). However, there are important properties of water that are not considered in typical implicit models. In particular, the nonadditive nature of hydrophobicity leads to the so-called “drying effect” in which a layer of vacuum surrounds hydrophobic surfaces and makes hydrophobic collapse cooperative (ten Wolde and Chandler 2002). In addition, continuum models do not account for the discrete nature of water molecules that may cause differences in folding dynamics due to cooperative expulsion of water (Cheung et al. 2002). It is also known that specific coordination of water molecules plays a role in the folded state of many proteins (Teeter 1991; Griffin et al. 2002).

Reprint requests to: Pernilla Wittung-Stafshede, Department of Biochemistry and Cell Biology, Keck Center for Structural Computational Biology, and Department of Chemistry, Rice University, 6100 Main Street, Houston, TX 77251, USA; e-mail: pernilla@rice.edu; fax: (713) 348-5154.

Article published online ahead of print. Article and publication date are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.051838206>.

Small, single-domain proteins often fold by two-state equilibrium and kinetic mechanisms (Jackson 1998; Kamagata et al. 2004). For such proteins, three states are important for defining kinetic and thermodynamic behavior: native, transition, and denatured states. As the transition state never accumulates, details about its properties must be inferred indirectly (Lindorff-Larsen et al. 2004). Experimentally, phi (ϕ) value analysis is the most important strategy to explore the nature of the transition state (Matouschek et al. 1990; Matouschek and Fersht 1991). In such experiments the transition state is probed by measuring the kinetic and the thermodynamic effects of hydrophobic-to-alanine mutations in different regions of the protein. The ϕ -values represent the change in stability of the transition state accompanying the mutation of a residue relative to the effect of the same mutation in the native state, assuming there is no effect on the unfolded state by the mutation. Thus, ϕ equal to 1 suggests that a residue makes interactions that contribute equally to the stability of the transition and the native states (i.e., native-like interactions in the transition state). In contrast, ϕ equal to 0 indicates that the residue forms very few, if any, interactions with other residues in the transition state. Fractional ϕ -values may be interpreted as possessing different degrees of structure in the folding nucleus (Sanchez and Kiefhaber 2003d; Fersht and Sato 2004). Several small proteins folding by two-state kinetic mechanisms have been subjects of mutagenesis to obtain a picture of the folding-transition state with residue-specific resolution. The results of these studies have led to the distinction of two classes of folding-transition states (Itzhaki et al. 1995; Goldenberg 1999). In diffuse transition states all but a few side chains have similar ϕ -values, which are relatively low. This has been shown to indicate a nucleation-condensation mechanism with the folding nucleus located diffusively throughout the protein (e.g., CI2, Arc repressor, CheY, and λ repressor) (Itzhaki et al. 1995; Milla et al. 1995; Lopez-Hernandez and Serrano 1996; Burton et al. 1997). In contrast, polarized transition states display distinct substructures that have very high ϕ -values while residues in the rest of the protein have ϕ -values of approximately zero (e.g., src SH3, α -spectrin SH3, fyn SH3, ACBP, Im7, and protein L) (Grantcharova et al. 1998; Martinez et al. 1998; Kragelund et al. 1999; Kim et al. 2000; Capaldi et al. 2002; Northey et al. 2002).

Recently, another mutational approach was introduced to specifically assess the degree of solvation in the folding-transition state (Fernandez-Escamilla et al. 2004). To our knowledge, this is the only experimental method that directly targets solvation in the folding-transition state. In contrast to hydrophobic-to-alanine exchanges (used in "normal" ϕ -value analysis), hydrophobic-to-polar mutations, specifically valine-to-threo-

nine, will not disrupt the packing of the core of the protein, as these residues (i.e., Val and Thr) are isosteric. However, because the Thr side chain is polar and can interact with water, this substitution will stabilize the unfolded state and thereby destabilize the protein (Serrano et al. 1992; Fernandez-Escamilla et al. 2004). Due to the different energetic consideration, ϕ -values derived from Thr mutations are interpreted in a different manner. If water is absent in the transition state (i.e., a dry core), the ϕ -value for a Val-to-Thr mutation should not introduce much perturbation in the transition state since both residues (i.e., valine and threonine) can make similar contacts. In this case, the ϕ -value approaches 1 as the energy of the unfolded state, and thus also the folding speed, has changed by the introduced Thr. If, on the other hand, there are water molecules in the interior of the protein eliminating contacts in the transition-state structure, a Val-to-Thr mutation should stabilize the transition state by forming hydrogen bonds with the buried waters (Fernandez-Escamilla et al. 2004). In the extreme of this case, the unfolded and transition states will be similarly stabilized and the ϕ -value for the Val-to-Thr mutation will approach 0. If the ϕ -value is between 0 and 1, the particular position is partly solvated in the transition state (Fernandez-Escamilla et al. 2004). To date, this method has only been applied to α -spectrin SH3, revealing that the folding-transition state for this protein is partially solvated (Fernandez-Escamilla et al. 2004). This experimental finding was shown to match computer simulations (Cheung et al. 2002; Guo et al. 2003; Fernandez-Escamilla et al. 2004).

Pseudomonas aeruginosa azurin is a small (128 residues) single-domain protein with a β -barrel structure composed of eight β -strands, which belongs to the sandwich-like protein family (Adman 1991). In vivo, a redox-active copper is coordinated to the protein allowing for electron-transfer activity (Adman 1991). The copper in azurin can be eliminated, creating apo-azurin, or substituted for zinc without change of the overall structure (Nar et al. 1992a,b). Equilibrium and kinetic folding processes for apo- and zinc-forms of azurin are two-state reactions (Pozdnyakova and Wittung-Stafshede 2001b, 2003; Pozdnyakova et al. 2001, 2002; Fuentes et al. 2004; Wittung-Stafshede 2004). The significantly curved Chevron plot for zinc-substituted azurin has been attributed to movement of a highly diffusive folding-transition state and was recently used to investigate the growth of the folding-transition state with residue-specific resolution (Wilson and Wittung-Stafshede 2005b). In contrast, apo-azurin folds via a fixed, highly polarized transition state (Wilson and Wittung-Stafshede 2005a). Thus, azurin is a unique model system that can be used to assess the presence of water in the two most common classes of folding-transition states

within the same structural framework. Here we take advantage of this property and combine Val-to-Ala and Val-to-Thr mutations to probe the solvation of azurin's folding-transition state in the apo- and zinc-forms, respectively.

Results

Selection of azurin mutations

Six hydrophobic positions in *P. aeruginosa* azurin's core were selected for the analysis of solvation in the transition state (i.e., positions 20, 22, 31, 50, 81, and 95; Fig. 1). For positions where the natural residue is not a valine (i.e., Ile20, Leu50, and Ile81), we made the corresponding hydrophobic-to-valine substitution (to be used as "wild type" in data analysis), in addition to the Thr and Ala substitutions. The residues at the selected positions point inward toward the core and introductions of threonines at these positions are presumed not to make any new interactions in the folded state. However, they are expected to stabilize the unfolded state, and thereby decrease overall protein stability, by way of hydrogen bonding with water. The unique emission from Trp48 is a sensitive and distinctive tool that reports on the presence of a native-like core (Turoverov et al. 1985). As expected, all azurin variants studied herein adopt native-like folded structures according to far-UV CD and unique Trp48 fluorescence; moreover, all mutants bind copper like the wild-type protein (data not shown).

To further confirm no conformational change of the folded structure upon mutating a valine to a threonine, computational-energy minimizations in silico were

performed (see Materials and Methods). According to the results, Val-to-Ala mutations lead to cavity formation and minor structural rearrangements, whereas Val-to-Thr mutations do not perturb the native core packing and no new interactions are detected compared to the wild-type structure. Quantitatively, as predicted, the energy corresponding to van der Waals interactions are similar for the Val and the Thr side chains but lower for the Ala side chain at each given position. This result is consistent with another study in which a crystal structure of a Val-to-Thr mutated (i.e., Val44Thr) variant of SH3 revealed no conformational change compared to the wild-type protein (Fernandez-Escamilla et al. 2004).

Equilibrium unfolding of azurin variants

Past biophysical studies have demonstrated that equilibrium unfolding processes of apo- and zinc-forms of azurin are two-state reactions (Mei et al. 1999; Pozdnyakova and Wittung-Stafshede 2001a, 2003; Pozdnyakova et al. 2001, 2002; Fuentes et al. 2004; Marks et al. 2004; Wittung-Stafshede 2004). In the case of unfolded zinc-substituted azurin, the metal remains bound to the native-state ligands Cys112 and His117, while the interactions with His46 and Met121 are lost (Marks et al. 2004). The thermodynamic stability of the variants, in both apo- and zinc-forms, was tested by guanidine hydrochloride (GuHCl) titrations monitored by far-UV CD and fluorescence. All variants unfold in single, reversible transitions (Fig. 2A, apo-forms of Thr-variants; Fig. 2B, zinc-forms of Thr-variants). Curves derived from CD- and fluorescence-data are coincidental for each protein, which is indicative of two-state equilibrium-unfolding processes. Predictably,

Figure 1. Illustration (prepared in PyMol) showing the structure of azurin (1AZU) with the herein targeted residues 20, 22, 31, 50, 81, and 95 highlighted in yellow. Trp48, which unique fluorescence reports on an intact core, is shown in green. Side view, *left*; top view, *right*.

Figure 2. (A) GuHCl-induced equilibrium unfolding of wild type and the apo-forms of the Thr-substituted azurin variants. For each protein, CD- and fluorescence-detected transitions overlap (normalized fluorescence data shown). Solid curves are two-state fits to each set of data. Legend to symbols is given in the figure. (We note that the transition for Val95Thr is steeper than that for wild-type apo-azurin, resulting in a slightly higher $\Delta G_U[H_2O]$ value for the former; Table 1.) (B) GuHCl-induced equilibrium unfolding of wild type and the zinc-forms of the Thr-substituted azurin variants. For each protein, CD- and fluorescence-detected transitions overlap (normalized fluorescence data shown). Solid curves are two-state fits to each set of data (Table 2). Legend to symbols is given in the figure.

the thermodynamic stability is lower for all the variants compared to the stability of wild-type azurin (Table 1, apo-forms; Table 2, zinc-forms), except for Val95Thr apo-azurin (further discussed below). In accord with conservative mutations, the three Ile/Leu-to-Val exchanges result in only minor stability decreases compared to the stability of wild-type azurin.

Effect of mutations on azurin folding dynamics

The kinetic-folding processes for apo- and zinc-forms of azurin are two-state reactions (Pozdnyakova and Wittung-Stafshede 2001b, 2003; Pozdnyakova et al. 2001, 2002; Fuentes et al. 2004; Wilson and Wittung-Stafshede 2005b). Apo-azurin's folding nucleus appears to be highly polarized; that is, a few core residues have ϕ -values of 1, whereas others have low, near-zero ϕ -values (Wilson and Wittung-Stafshede 2005a). In contrast to apo-azurin, the folding dynamics of zinc-substituted azurin exhibits curvature in both folding and unfolding limbs, when the natural logarithms of the observed rate constants are plotted as a function of GuHCl concentration. We have earlier demonstrated that this curvature corresponds to transition-state movement (Pozdnyakova and Wittung-Stafshede 2003; Wittung-Stafshede 2004; Wilson and Wittung-Stafshede 2005b).

Folding and unfolding kinetics of the apo- and zinc-forms of the variants were probed by fluorescence and far-UV CD detection methods using stopped-flow mixing. In support of two-state reactions, all kinetic reactions are single exponential decays and are devoid of missing amplitude. For each variant, CD- and fluorescence-detected kinetic traces overlap at each condition. Moreover, there is no protein-concentration dependence in either unfolding or refolding phases. The Chevron plots for the apo-forms of the variants are V-shaped (Fig. 3A) and were analyzed assuming standard linear dependences on the logarithms of the observed rate constants as a function of denaturant (Fersht 1999). For all apo-variants, the β -values fall in a narrow window around 0.6. For all zinc-variants, the Chevron plots are symmetrically curved (Fig. 3B) and were fit to second-order polynomials (see Materials and Methods) to obtain folding- and unfolding-rate constants in water. The β -values at three selected GuHCl concentrations (calculated from the polynomial-fit parameters b_U and c_U and Equations 2b and 3 in Materials and Methods) are listed in Table 2. It is clear that similar transition-state movements occur for all zinc-substituted azurin variants regardless of the type of mutation (average β -values of 0.46, 0.65, and 0.82, at 0, 2, and 4 M GuHCl, respectively).

Protein engineering analysis

The data in Tables 1 and 2 were used to calculate ϕ -values for the Ala and Thr substitutions, respectively, for azurin in its apo-form (i.e., fixed ϕ -value throughout denaturant range) and zinc-form (i.e., ϕ -values for three different denaturant concentrations). Since there is concern regarding the accuracy of ϕ -values involving low $\Delta\Delta G_U(H_2O)$ values (Sanchez and Kiefhaber 2003d;

Table 1. Thermodynamic and kinetic parameters for wild-type and point-mutated apo-azurin variants in the absence of denaturant (at pH 7, 25°C)

Apo-azurin variant	Mutated β -strand	ΔG_U (H ₂ O) ^a (kJ/mol)	k_F (H ₂ O) (sec ⁻¹)	β	ϕ ^b
Wild type	—	25.0 ± 0.7	160 ± 20	0.60 ± 5	—
Ile20Val	2	23.4 ± 0.5	125 ± 15	0.54 ± 5	ND
Ile20Ala	2	18.5 ± 0.5	79 ± 9	0.58 ± 5	0.24 ± 0.05
Ile20Thr	2	15.0 ± 0.5	110 ± 15	0.69 ± 5	0.04 ± 0.03
Val22Ala	2	19.0 ± 0.5	103 ± 10	0.66 ± 5	0.18 ± 0.05
Val22Thr	2	21.0 ± 0.5	100 ± 10	0.69 ± 5	0.27 ± 0.08
Val31Ala ^c	3	20.0 ± 0.4	23 ± 5	0.60 ± 5	0.93 ± 0.14
Val31Thr	3	20.5 ± 0.5	28 ± 5	0.64 ± 5	1.00 ± 0.15
Leu50Val	4	23.5 ± 0.5	122 ± 15	0.56 ± 5	ND
Leu50Ala ^c	4	18.0 ± 0.2	10.0 ± 3	0.62 ± 5	1.10 ± 0.11
Leu50Thr	4	15.2 ± 0.5	5.4 ± 3	0.68 ± 5	0.93 ± 0.09
Ile81Val	5	24.7 ± 0.5	149 ± 20	0.53 ± 5	ND
Ile81Ala	5	15.0 ± 0.5	19 ± 2	0.65 ± 5	0.53 ± 0.04
Ile81Thr	5	14.5 ± 0.5	26 ± 5	0.67 ± 5	0.42 ± 0.04
Val95Ala ^c	6	22.5 ± 0.2	78 ± 10	0.66 ± 5	ND
Val95Thr	6	29.0 ± 0.5	179 ± 20	0.60 ± 5	(0.07 ± 0.06)

The ϕ -values are calculated as $\phi = \Delta\Delta G^\ddagger(\text{H}_2\text{O})/\Delta\Delta G_U(\text{H}_2\text{O})$, where $\Delta\Delta G_U(\text{H}_2\text{O})$ is calculated as $\Delta G_U(\text{mutant}, \text{H}_2\text{O}) - \Delta G_U(\text{wild type}, \text{H}_2\text{O})$ and $\Delta\Delta G^\ddagger(\text{H}_2\text{O})$ as $RT \ln[k_F(\text{mutant}, \text{H}_2\text{O})/k_F(\text{wild type}, \text{H}_2\text{O})]$. For positions 20, 50, and 81, the respective Val variant was used as “wild type” in the ϕ -calculations. The β -values are calculated according to Equation 3 (see Materials and Methods). Unless noted, errors are derived from the goodness of the appropriate fits.

^a Calculated from the kinetics as $RT \ln(k_F[\text{H}_2\text{O}]/k_U[\text{H}_2\text{O}])$.

^b Standard deviations (σ) for the ϕ -values were calculated as in Sanchez and Kiefhaber (2003d).

^c $\Delta G_U(\text{H}_2\text{O})$ and $k_F(\text{H}_2\text{O})$ values taken from Wilson and Wittung-Stafshede (2005a).

ND, Not determined (see text; Ile20Val, Leu50Val, and Ile81Ala used as “wild types;” $\Delta\Delta G_U(\text{H}_2\text{O})$ is too low for Val95Ala for reliable ϕ calculation).

Fersht and Sato 2004; Settanni et al. 2005) we have here used a cutoff in $\Delta\Delta G_U(\text{H}_2\text{O})$ of 4 kJ/mol. The cutoff results in the exclusion of the apo-form of Val95Ala and the zinc-form of Val95Thr azurin from further analysis (see Tables 1, 2). In addition, the Val95Thr mutation appears to cause new stabilizing interactions in the transition and the native states of the apo-form since the folding speed and the overall stability is increased compared to that of wild type. In support, a slight compaction was detected in the energy-minimized structure of Val95Thr azurin. Because of this, position 95 was omitted from further analysis.

By comparing the coupled ϕ -values for the five remaining mutant pair (Val/Ala and Val/Thr), the degree of solvation was evaluated as a function of residue participation in the transition-state structure. For residues 20, 50, and 81, the respective Val-variant was used as wild type in the ϕ -calculations. For apo-azurin, positions 31 and 50 form native-like interactions in the transition state according to the normal, Ala-based ϕ -values. For these positions, the Thr-based ϕ -values are also close to 1, implying exclusion of water. Positions 20 and 22 do not participate in the folding nucleus (Ala-based ϕ -values <0.25) and, in accord, are almost as solvated as the unfolded state (Thr-based ϕ -values near 0). Position 81 is the only position in apo-azurin that has been found to exhibit a truly intermediate

Ala-based ϕ -value (i.e., 0.53); the corresponding Thr-based ϕ -value for position 81 (i.e., 0.42) indicates partial solvation. Either residue 81 interacts partially with both water and other side chains in the transition state, or this state is heterogeneous and involves two fractions of molecules: one with position 81 not involved at all (and thus the side chain is fully solvated) and another with position 81 involved in fully native-like interactions (i.e., dry).

For zinc-substituted azurin, instead of a small and distinct substructure with high ϕ -values, the transition state involves partial interactions at all the probed positions (and additional positions) (Wilson and Wittung-Stafshede 2005b) that gradually become more native-like according to the normal ϕ -values. The average Ala-based ϕ -value, among the positions probed here, moves from 0.31 to 0.64 to 0.90, in 0, 2, and 4 M GuHCl, respectively. The corresponding Thr-based ϕ -values also start out low and gradually grow higher as a function of GuHCl. The average Thr-based ϕ -value, among the positions probed, goes from 0.15 to 0.60 to 0.91, in 0, 2, and 4 M GuHCl, respectively. This finding implies that water molecules bridge between side chains in zinc-azurin's folding-transition state, resulting in an expanded and diffusive nucleus with only partially formed residue-residue interactions. When the GuHCl concentration is increased, the folding-

Table 2. Thermodynamic and kinetic parameters for wild-type and point-mutated zinc-substituted azurin variants as a function of denaturant (0, 2, and 4 M GuHCl at pH 7, 25°C)

Zinc-azurin variant	$\Delta G_U(\text{H}_2\text{O})^a$ (kJ/mol)	$k_F(\text{H}_2\text{O})$ (sec ⁻¹)	$\beta(\text{H}_2\text{O})$	$\beta(2 \text{ M})$	$\beta(4 \text{ M})$	$\phi(\text{H}_2\text{O})^a$	$\phi(2 \text{ M})^a$	$\phi(4 \text{ M})^a$
Wild type	54 ± 1.5	126 ± 18	0.41 ± 0.08	0.58 ± 0.08	0.76 ± 0.08	—	—	—
Ile20Val	54 ± 1.5	124 ± 10	0.35 ± 0.08	0.57 ± 0.08	0.79 ± 0.08	ND	ND	ND
Ile20Ala ^b	44 ± 1.5	69 ± 6	0.56 ± 0.07	0.74 ± 0.07	0.88 ± 0.07	0.15 ± 0.04	0.46 ± 0.06	0.82 ± 0.09
Ile20Thr	46 ± 1.0	82 ± 5	0.54 ± 0.08	0.72 ± 0.08	0.89 ± 0.08	0.13 ± 0.06	0.59 ± 0.07	0.73 ± 0.09
Val22Ala ^b	47 ± 1.0	84 ± 7	0.47 ± 0.10	0.64 ± 0.10	0.81 ± 0.10	0.14 ± 0.05	0.50 ± 0.07	0.88 ± 0.09
Val22Thr	36 ± 1.0	74 ± 7	0.54 ± 0.07	0.71 ± 0.07	0.87 ± 0.07	0.07 ± 0.02	0.42 ± 0.03	0.98 ± 0.05
Val31Ala ^b	47 ± 1.0	42 ± 3	0.56 ± 0.06	0.73 ± 0.06	0.88 ± 0.06	0.39 ± 0.09	0.67 ± 0.10	0.88 ± 0.09
Val31Thr	37 ± 1.0	65 ± 6	0.48 ± 0.07	0.63 ± 0.07	0.79 ± 0.07	0.09 ± 0.02	0.55 ± 0.04	0.99 ± 0.06
Leu50Val	53 ± 1.5	122 ± 10	0.41 ± 0.07	0.63 ± 0.07	0.85 ± 0.07	ND	ND	ND
Leu50Ala ^b	46.5 ± 1.0	32 ± 3	0.33 ± 0.07	0.55 ± 0.07	0.76 ± 0.07	0.44 ± 0.10	0.97 ± 0.15	1.04 ± 0.12
Leu50Thr	45 ± 1.0	47 ± 5	0.56 ± 0.08	0.71 ± 0.08	0.87 ± 0.08	0.30 ± 0.07	0.85 ± 0.10	1.02 ± 0.13
Ile81Val	53 ± 1.5	122 ± 10	0.41 ± 0.07	0.60 ± 0.07	0.78 ± 0.07	ND	ND	ND
Ile81Ala ^b	48 ± 1.0	55 ± 5	0.58 ± 0.07	0.69 ± 0.07	0.79 ± 0.07	0.33 ± 0.09	0.75 ± 0.12	0.89 ± 0.10
Ile81Thr	40 ± 1.0	58.5 ± 5	0.31 ± 0.07	0.55 ± 0.07	0.78 ± 0.07	0.14 ± 0.04	0.59 ± 0.06	0.85 ± 0.08
Val95Ala ^b	50 ± 1.0	67 ± 8	0.34 ± 0.08	0.57 ± 0.08	0.79 ± 0.08	0.38 ± 0.12	0.48 ± 0.12	0.86 ± 0.15
Val95Thr	53.5 ± 1.0	106 ± 10	0.53 ± 0.08	0.70 ± 0.08	0.86 ± 0.08	ND	ND	ND

The ϕ -values are calculated as $\phi(X) = \Delta\Delta G^\ddagger(X)/\Delta\Delta G_U(X)$ at each specified GuHCl condition ($X = 0, 2$, and 4 M GuHCl). $\Delta\Delta G_U(X)$ is calculated as $\Delta G_U(\text{mutant}; X) - \Delta G_U(\text{wild type}; X)$ and $\Delta\Delta G^\ddagger(X)$ as $R^*T \ln[k_f(\text{mutant}; X)/k_f(\text{wild type}; X)]$. For positions 20, 50, and 81, the respective Val variant was used as wild type in the ϕ calculations. The β -values are calculated according to Equation 3 (see Materials and Methods) at each X value. Unless noted, errors are derived from the goodness of the appropriate fits.

^aStandard deviations (σ) for the ϕ -values were calculated as in Sanchez and Kiefhaber (2003d).

^bData taken from Wilson and Wittung-Stafshede (2005b).

ND, Not determined (see text; Ile20Val, Leu50Val, and Ile81Ala used as “wild types;” $\Delta\Delta G_U(\text{H}_2\text{O})$ is too low for Val95Thr for reliable ϕ calculation).

transition state becomes more structured (Ala-based ϕ -values go toward 1) and this results in decreased solvation. Thus, in zinc-substituted azurin, the water molecules are gradually expelled from the folding nucleus as the side-chain interactions grow more native-like.

Discussion

Several small proteins folding by two-state kinetic mechanisms have been subjected to mutagenesis to obtain a picture of the folding transition state with residue-specific resolution (Itzhaki et al. 1995; Milla et al. 1995; Lopez-Hernandez and Serrano 1996; Burton et al. 1997; Kragelund et al. 1999; Riddle et al. 1999; Kim et al. 2000; Capaldi et al. 2002; Northey et al. 2002). Despite this wealth of information, the role of water in protein folding remains an unresolved question. Are water molecules present in folding-transition states and expelled later, or are they detached before the top of the free-energy barrier is reached? Using computer simulations, this issue has been addressed to some degree, although the conclusions diverge. For example, in one study, implicit and explicit solvent models were found to give very different results in terms of folding dynamics and structures (Zhou 2003), whereas another study reported that although water is trapped in the folding-transition state, it does not have a specific role nor does

it affect the transition-state structure (Rhee et al. 2004). The only protein for which direct comparisons between kinetic experiments and simulations addressing the issue of solvation have been made is the SH3 protein. For this small β -sheet protein, it was found that the highly polarized transition state, which involves formation of one of the two β -sheets, contains water (Cheung et al. 2002; Guo et al. 2003; Fernandez-Escamilla et al. 2004). These water molecules are expelled after the rate-limiting step, when the second β -sheet forms and comes into proximity of the already formed β -sheet.

Here we demonstrate that while the apo-azurin folding-transition state lacks water molecules, zinc-substituted azurin's transition state is solvated (Fig. 4). We speculate that this sharp difference is closely connected to the shape of the free energy barrier for folding. It was recently proposed that apo-azurin's fixed transition state is the result of a small pointed feature projecting from the top of an otherwise broad free-energy profile (Wilson and Wittung-Stafshede 2005b). The presence of zinc in the unfolded state suppresses this local bump and the underlying broad activation barrier becomes accessible. The lack of unique energetically elevated features in this broad free-energy profile results in the moving transition state that is observed for zinc-substituted azurin (Wilson and Wittung-Stafshede 2005b). Based on the current results, we can now assign an energetic explanation for

Figure 3. (A) Semilogarithmic graph of $\ln k_{\text{obs}}$ vs. GuHCl for wild type and the apo-forms of the Thr-substituted azurin variants. Solid lines are two-state fits to the data assuming linear dependencies of the logarithms of the rate constants vs. GuHCl. Legend to symbols is given in the figure. (B) Semilogarithmic graph of $\ln k_{\text{obs}}$ vs. GuHCl for wild type and the zinc-forms of the Thr-substituted azurin variants. Solid lines are polynomial fits to the data (see Materials and Methods). Legend to symbols is given in the figure.

apo-azurin's elevated feature that involves the cost of water removal.

We note that there are many origins of curved Chevron plots (Ternstrom et al. 1999; Bachmann and Kiefhaber 2001; Sanchez and Kiefhaber 2003a,b,c). For zinc-substituted azurin, explanations such as transient aggregation, transient intermediates, ground-state effects, switches between parallel pathways or between a few consecutive barriers on a sequential pathway have all been excluded (Wilson and Wittung-Stafshede 2005b). Although we cannot exclude that variations of zinc ligands in the transition state at different GuHCl concentrations may cause ap-

parent curvature, we disfavor this explanation since positions near the zinc site (zinc coordinates His117 and Cys112 in the unfolded state) have low ϕ -values throughout the denaturant range tested (Wilson and Wittung-Stafshede 2005b). Moreover, if different modes of zinc coordination are responsible for variations in the transition state at different GuHCl concentrations, kinks (instead of a smooth curvature) in the folding arms in combination with straight unfolding arms are expected; however, this is not observed.

In conclusion, this experimental study on *P. aeruginosa* azurin is one of few that address the role of water in the transition state for folding. Our findings demonstrate that there are several possible scenarios: water can be both present and absent in the folding-transition state. Further experimental studies on other proteins, as well as computational studies using explicit solvent models, are required to reveal if water molecules are merely trapped in the transition state or if they play an active role. Nevertheless, the distinct difference in solvation of apo- and holo-azurin's folding-transition states emphasizes the significant effect a small metal ion can have in protein folding.

Materials and methods

Construction of variants

A pUC18 vector containing the wild-type *P. aeruginosa* azurin gene was used as the template to create the point-mutated variants (i.e., Ile20Ala, Ile20Val, Ile20Thr, Val22Ala, Val22Thr, Val31Ala, Val31Thr, Leu50Ala, Leu50Val, Leu50Thr, Ile81Ala, Ile81Val, Ile81Thr, Val95Ala, Val95Thr) via QuikChange site-directed mutagenesis as reported in Marks et al. (2004) and Wilson and Wittung-Stafshede (2005a). Variants were expressed in *Escherichia coli* BL21pLysS cells and purification was performed as in Pozdnyakova et al. (2002). To assure 100% apo- and zinc-substitution, respectively, the apo-form of each variant was first prepared by extensive cyanide dialysis to remove both copper and zinc; to create the zinc form, subsequent dialysis in the presence of excess zinc was performed. Successful apo- and zinc-reconstitution was verified by copper-titration monitoring absorption at 630 nm (Pozdnyakova et al. 2002).

Equilibrium unfolding

GuHCl-induced equilibrium unfolding was performed in 100 mM Tris-HCl (pH 7.0), 25°C using fluorescence (excitation at 285 nm; emission monitored at 308 nm) and far-UV circular dichroism (CD) detection (200–300 nm). Samples were incubated for 2 h before measurements. The equilibrium-unfolding reactions were reversible without any display of protein-concentration dependence (in 5–50 μ M protein range). The equilibrium-unfolding curves were analyzed using a two-state model (Fersht 1999; Pace and Shaw 2000) to determine $\Delta G_{\text{U}}(\text{H}_2\text{O})$ and $m_{\text{F-U}}$ values. For each azurin variant, far-UV CD- and fluorescence-detected transitions overlapped. Errors were derived through multiple experiments.

Figure 4. Illustration of the relation between the presence of structure in the transition state (via Ala-based ϕ) and the presence of water in the transition state (via Thr-based ϕ) for apo- (A) and zinc- (0 M GuHCl, B; 2 M GuHCl, C; 4 M GuHCl, D) forms of azurin. Probed residues (data taken from Tables 1 and 2 with Val95 excluded) are shown in stick representations and color-coded as follows for Ala-based ϕ -values: $\phi = 0$ –0.39, yellow; $\phi = 0.4$ –0.59, orange; $\phi = 0.6$ –0.79, red; $\phi = 0.8$ –1.0, purple. The degree of solvation is indicated by the presence or the absence of water molecules (light blue spheres), where four spheres symbolize complete solvation of that particular residue.

Folding dynamics and data analysis

Time-resolved folding and unfolding was probed by fluorescence (excitation at 285 nm; emission monitored at 308 nm) and far-UV CD (at 220 nm) using an Applied Photophysics Pi-Star stopped-flow mixer. Both detection modes gave identical kinetic traces in all cases. Apo- and zinc-azurin variants were mixed in 1:10 ratio with appropriate GuHCl/buffer solutions. At each condition, six kinetic traces were averaged before fitting. There were no missing amplitudes (within the 2–3 msec dead time) in either fluorescence or far-UV CD kinetic traces, and no protein-concentration dependence was found (5–50 μ M protein range).

For apo-azurin, both unfolding and refolding reactions were fit to monoexponential decay equations. The unfolding and the refolding rate constants at different GuHCl concentrations were then fit assuming standard linear dependence of $\ln k_F$ and $\ln k_U$ on GuHCl concentration (Fersht 1997). For zinc-substituted azurin the kinetic traces were fit to monophasic decay equations for unfolding and biphasic decay equations for refolding (Wilson and Wittung-Stafshede 2005b). For each zinc variant, the minor slow refolding phase (<5% of fluorescence and far-UV CD changes in all cases) was neglected in this study. Unfolding and refolding (i.e., the fast, major phase) rate constants at different GuHCl concentrations were fit to second order polynomials (Otzen et al. 1999; Ternstrom et al. 1999):

$$\ln k_{\text{obs}} = \ln[k_F(\text{H}_2\text{O})\exp(b_F[\text{GuHCl}] + c_F[\text{GuHCl}]^2) + k_U(\text{H}_2\text{O})\exp(b_U[\text{GuHCl}] + c_U[\text{GuHCl}]^2)] \quad (1)$$

In this equation, b_F/b_U is the linear dependence of folding/unfolding, and c_F/c_U is the parameter describing the curvature observed in the folding/unfolding limbs in the Chevron plots; $k_F(\text{H}_2\text{O})$ and $k_U(\text{H}_2\text{O})$ are the folding and unfolding rate constants in the absence of GuHCl. The actual m -values (m_F/m_U) are given by the relationship:

$$m_F = \delta(\ln k_F) / d[\text{GuHCl}] = b_F + 2c_F[\text{GuHCl}] \quad (2a)$$

$$m_U = \delta(\ln k_U) / d[\text{GuHCl}] = b_U + 2c_U[\text{GuHCl}] \quad (2b)$$

The Tanford β -value for folding, which assesses the position of the transition state in relation to the folded and unfolded structures (Tanford 1968, 1970), is calculated as:

$$\beta = -m_F / m_{F-U} = 1 - m_U / m_{F-U} \quad (3)$$

Thus, for curved Chevrans, β becomes a linear function of the denaturant concentration.

Energy minimizations of mutant structures

Energy minimizations were performed with the software Molecular Modeling Pro (Norgwyn Montgomery Software Inc). Using the Swiss PDB Viewer software (Guex and Peitsch 1997; Schwede et al. 2003), the wild-type azurin crystal structure (1AZU) was used as a template in which the targeted side chains were individually replaced with Val, Thr, or Ala side chains to recreate the experimentally studied variants. For each variant, residues within a 5 Å radius around the altered site were selected and saved as a new pdb (to reduce the overall file size). Next, geometry optimization was carried out by minimization of each such substructure using the Molecular Modeling Pro suite, which employs Molecular Mechanics 2 (MM2) force fields to find the local energy minimum (Allinger 1977). In Molecular Modeling Pro, dipole interactions are calculated using modified schemes based on the Del Re Method; their pi-contributions are calculated using the MPEOE method. The overall interaction energies of the minimized structures, as estimated by this program, include contributions from favorable van der Waals and dipole interactions as well as unfavorable interactions, such as bending, Coulomb, and torsion interactions. The favorable van der Waals interactions dominate in all cases.

Acknowledgments

Support for this project was provided by grants from the NIH (GM059663) and the Robert A. Welch Foundation (C-1588).

C.J.W. is supported by the Houston Area Molecular Biophysics Program (GM08280).

References

- Adman, E.T. 1991. Copper protein structures. *Adv. Protein Chem.* **42**: 145–197.
- Allinger, N.L. 1977. Conformational analysis 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *J. Am. Chem. Soc.* **99**: 8127–8134.
- Bachmann, A. and Kiefhaber, T. 2001. Apparent two-state tendamistat folding is a sequential process along a defined route. *J. Mol. Biol.* **306**: 375–386.
- Burton, R.E., Huang, G.S., Daugherty, M.A., Calderone, T.L., and Oas, T.G. 1997. The energy landscape of a fast-folding protein mapped by Ala→Gly substitutions. *Nat. Struct. Biol.* **4**: 305–310.
- Capaldi, A.P., Kleanthous, C., and Radford, S.E. 2002. Im7 folding mechanism: Misfolding on a path to the native state. *Nat. Struct. Biol.* **9**: 209–216.
- Cheung, M.S., Garcia, A.E., and Onuchic, J.N. 2002. Protein folding mediated by solvation: Water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc. Natl. Acad. Sci.* **99**: 685–690.
- Fernandez-Escamilla, A.M., Cheung, M.S., Vega, M.C., Wilmanns, M., Onuchic, J.N., and Serrano, L. 2004. Solvation in protein folding analysis: Combination of theoretical and experimental approaches. *Proc. Natl. Acad. Sci.* **101**: 2834–2839.
- Fersht, A.R. 1997. Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* **7**: 3–9.
- Fersht, A. 1999. *Structure and mechanism in protein science*. W.H. Freeman and Company, New York.
- Fersht, A.R. and Sato, S. 2004. Φ -Value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci.* **101**: 7976–7981.
- Fuentes, L., Oyola, J., Fernandez, M., and Quinones, E. 2004. Conformational changes in azurin from *Pseudomonas aeruginosa* induced through chemical and physical protocols. *Biophys. J.* **87**: 1873–1880.
- Goldenberg, D.P. 1999. Finding the right fold. *Nat. Struct. Biol.* **6**: 987–990.
- Grantcharova, V.P., Riddle, D.S., Santiago, J.V., and Baker, D. 1998. Important role of hydrogen bonds in the structurally polarized transition state for folding of the src SH3 domain. *Nat. Struct. Biol.* **5**: 714–720.
- Griffin, S., Vitello, A., and Wittung-Stafshede, P. 2002. Buried water molecules contribute to cytochrome *c* stability. *Arch. Biochem. Biophys.* **404**: 335–337.
- Guex, N. and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714–2723.
- Guo, W., Lampoudi, S., and Shea, J.E. 2003. Posttransition state desolvation of the hydrophobic core of the src-SH3 protein domain. *Biophys. J.* **85**: 61–69.
- Itzhaki, L.S., Otzen, D.E., and Fersht, A.R. 1995. The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: Evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* **254**: 260–288.
- Jackson, S.E. 1998. How do small single-domain proteins fold? *Fold. Des.* **3**: R81–R91.
- Kamagata, K., Arai, M., and Kuwajima, K. 2004. Unification of the folding mechanisms of non-two-state and two-state proteins. *J. Mol. Biol.* **339**: 951–965.
- Kim, D.E., Fisher, C., and Baker, D. 2000. A breakdown of symmetry in the folding transition state of protein L. *J. Mol. Biol.* **298**: 971–984.
- Kragelund, B.B., Osmark, P., Neergaard, T.B., Schiodt, J., Kristiansen, K., Knudsen, J., and Poulsen, F.M. 1999. The formation of a native-like structure containing eight conserved hydrophobic residues is rate limiting in two-state protein folding of ACBP. *Nat. Struct. Biol.* **6**: 594–601.
- Lindorff-Larsen, K., Vendruscolo, M., Paci, E., and Dobson, C.M. 2004. Transition states for protein folding have native topologies despite high structural variability. *Nat. Struct. Mol. Biol.* **11**: 443–449.
- Lopez-Hernandez, E. and Serrano, L. 1996. Structure of the transition state for folding of the 129 aa protein CheY resembles that of a smaller protein, CI-2. *Fold. Des.* **1**: 43–55.
- Marks, J., Pozdnyakova, I., Guidry, J., and Wittung-Stafshede, P. 2004. Methionine-121 coordination determines metal specificity in unfolded *Pseudomonas aeruginosa* azurin. *J. Biol. Inorg. Chem.* **9**: 281–288.
- Martinez, J.C., Pisabarro, M.T., and Serrano, L. 1998. Obligatory steps in protein folding and the conformational diversity of the transition state. *Nat. Struct. Biol.* **5**: 721–729.
- Matouschek, A. and Fersht, A.R. 1991. Protein engineering in analysis of protein folding pathways and stability. *Methods Enzymol.* **202**: 82–112.
- Matouschek, A., Kellis Jr., J.T., Serrano, L., Bycroft, M., and Fersht, A.R. 1990. Transient folding intermediates characterized by protein engineering. *Nature* **346**: 440–445.
- Mei, G., Di Venere, A., Campeggi, F.M., Gilardi, G., Rosato, N., De Matteis, F., and Finazzi-Agro, A. 1999. The effect of pressure and guanidine hydrochloride on azurins mutated in the hydrophobic core. *Eur. J. Biochem.* **265**: 619–626.
- Milla, M.E., Brown, B.M., Waldburger, C.D., and Sauer, R.T. 1995. P22 Arc repressor: Transition state properties inferred from mutational effects on the rates of protein unfolding and refolding. *Biochemistry* **34**: 13914–13919.
- Nar, H., Huber, R., Messerschmidt, A., Filippou, A.C., Barth, M., Jaquinod, M., van de Kamp, M., and Canters, G.W. 1992a. Characterization and crystal structure of zinc azurin, a by-product of heterologous expression in *Escherichia coli* of *Pseudomonas aeruginosa* copper azurin. *Eur. J. Biochem.* **205**: 1123–1129.
- Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., and Canters, G.W. 1992b. Crystal structure of *Pseudomonas aeruginosa* apo-azurin at 1.85 Å resolution. *FEBS Lett.* **306**: 119–124.
- Northey, J.G., Di Nardo, A.A., and Davidson, A.R. 2002. Hydrophobic core packing in the SH3 domain folding transition state. *Nat. Struct. Biol.* **9**: 126–130.
- Otzen, D.E., Kristensen, O., Proctor, M., and Oliveberg, M. 1999. Structural changes in the transition state of protein folding: Alternative interpretations of curved chevron plots. *Biochemistry* **38**: 6499–6511.
- Pace, C.N. and Shaw, K.L. 2000. Linear extrapolation method of analyzing solvent denaturation curves. *Proteins Suppl.* **4**: 1–7.
- Pozdnyakova, I. and Wittung-Stafshede, P. 2001a. Biological relevance of metal binding before protein folding. *J. Am. Chem. Soc.* **123**: 10135–10136.
- . 2001b. Copper binding before polypeptide folding speeds up formation of active (holo) *Pseudomonas aeruginosa* azurin. *Biochemistry* **40**: 13728–13733.
- . 2003. Approaching the speed limit for Greek Key β -barrel formation: Transition-state movement tunes folding rate of zinc-substituted azurin. *Biochim. Biophys. Acta* **1651**: 1–4.
- Pozdnyakova, I., Guidry, J., and Wittung-Stafshede, P. 2001. Copper stabilizes azurin by decreasing the unfolding rate. *Arch. Biochem. Biophys.* **390**: 146–148.
- . 2002. Studies of *Pseudomonas aeruginosa* azurin mutants: Cavities in β -barrel do not affect refolding speed. *Biophys. J.* **82**: 2645–2651.
- Rhee, Y.M., Sorin, E.J., Jayachandran, G., Lindahl, E., and Pande, V.S. 2004. Simulations of the role of water in the protein-folding mechanism. *Proc. Natl. Acad. Sci.* **101**: 6456–6461.
- Riddle, D.S., Grantcharova, V.P., Santiago, J.V., Alm, E., Ruczinski, I., and Baker, D. 1999. Experiment and theory highlight role of native state topology in SH3 folding. *Nat. Struct. Biol.* **6**: 1016–1024.
- Sanchez, I.E. and Kiefhaber, T. 2003a. Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J. Mol. Biol.* **325**: 367–376.
- . 2003b. Hammond behavior versus ground state effects in protein folding: Evidence for narrow free energy barriers and residual structure in unfolded states. *J. Mol. Biol.* **327**: 867–884.
- . 2003c. Non-linear rate-equilibrium free energy relationships and Hammond behavior in protein folding. *Biophys. Chem.* **100**: 397–407.
- . 2003d. Origin of unusual ϕ -values in protein folding: Evidence against specific nucleation sites. *J. Mol. Biol.* **334**: 1077–1085.
- Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. 2003. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **31**: 3381–3385.
- Serrano, L., Kellis Jr., J.T., Cann, P., Matouschek, A., and Fersht, A.R. 1992. The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. *J. Mol. Biol.* **224**: 783–804.

- Settanni, G., Rao, F., and Caflisch, A. 2005. ϕ -Value analysis by molecular dynamics simulations of reversible folding. *Proc. Natl. Acad. Sci.* **102**: 628–633.
- Tanford, C. 1968. Protein denaturation. *Adv. Protein Chem.* **23**: 121–282.
- . 1970. Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.* **24**: 1–95.
- Teeter, M.M. 1991. Water–protein interactions: Theory and experiment. *Annu. Rev. Biophys. Biophys. Chem.* **20**: 577–600.
- ten Wolde, P.R. and Chandler, D. 2002. Drying-induced hydrophobic polymer collapse. *Proc. Natl. Acad. Sci.* **99**: 6539–6543.
- Ternstrom, T.A., Mayor, U., Akke, M., and Oliveberg, M. 1999. From snapshot to movie: ϕ Analysis of protein folding transition states taken one step further. *Proc. Natl. Acad. Sci.* **96**: 14854–14859.
- Turoverov, K.K., Kuznetsova, I.M., and Zaitsev, V.N. 1985. The environment of the tryptophan residue in *Pseudomonas aeruginosa* azurin and its fluorescence properties. *Biophys. Chem.* **23**: 79–89.
- Wilson, C.J. and Wittung-Stafshede, P. 2005a. Role of structural determinants in folding of the sandwich-like protein *Pseudomonas aeruginosa* azurin. *Proc. Natl. Acad. Sci.* **102**: 3984–3987.
- . 2005b. Snapshots of a dynamic folding nucleus in zinc-substituted *Pseudomonas aeruginosa* azurin. *Biochemistry* **44**: 10054–10062.
- Wittung-Stafshede, P. 2004. Role of cofactors in folding of the blue-copper protein azurin. *Inorg. Chem.* **43**: 7926–7933.
- Zhou, R. 2003. Free energy landscape of protein folding in water: Explicit vs. implicit solvent. *Proteins* **53**: 148–161.