2010

Mutational Studies Uncover Non-Native Structure in the Dimeric Kinetic Intermediate of the H2A–H2B Heterodimer

Matthew R. Stump
George Fox University, mstump@georgefox.edu

Lisa M. Gloss
Washington State University

Mutational Studies Uncover Non-Native Structure in the Dimeric Kinetic Intermediate of the H2A–H2B Heterodimer

Matthew R. Stump and Lisa M. Gloss*

School of Molecular Biosciences, Washington State University, Pullman, WA 99164-7520, USA

Received 6 February 2010; received in revised form 15 June 2010; accepted 16 June 2010
Available online
23 June 2010

The folding pathway of the histone H2A–H2B heterodimer minimally includes an on-pathway, dimeric, burst-phase intermediate, I_2. The partially folded H2A and H2B monomers populated at equilibrium were characterized as potential monomeric kinetic intermediates. Folding kinetics were compared for initiation from isolated, folded monomers and the heterodimer unfolded in 4 M urea. The observed rates were virtually identical above 0.4 M urea, exhibiting a log-linear relationship on the final denaturant concentration. Below ~0.4 M urea (concentrations inaccessible from the 4-M urea unfolded state), a rollover in the rates was observed; this suggests that a component of the I_2 ensemble contains non-native structure that rearranges/isomerizes to a more native-like species. The contribution of helix propensity to the stability of the I_2 ensemble was assessed with a set of H2A–H2B mutants containing Ala and Gly replacements at nine sites, focusing mainly on the long, central α2 helix. Equilibrium and kinetic folding/unfolding data were collected to determine the effects of the mutations on the stability of I_2 and the transition state between I_2 and N_2. This limited mutational study indicated that residues in the α2 helices of H2A and H2B as well as α1 of H2B and both the C-terminus of α3 and the short αC helix of H2A contribute to the stability of the I_2 burst-phase species. Interestingly, at least eight of the nine targeted residues stabilize I_2 by interactions that are non-native to some extent. Given that destabilizing I_2 and these non-native interactions does not accelerate folding, it is concluded that the native and non-native structures present in the I_2 ensemble enable efficient folding of H2A–H2B.

Keywords: kinetic intermediates; oligomeric proteins; circular dichroism; fluorescence; histones

Edited by C. R. Matthews

Introduction

The dimeric structure of eukaryotic histones is an essential feature of their biological function as the protein core of the nucleosome core particle (NCP). The NCP is the fundamental repeating unit in the packaging of DNA in chromatin. In the NCP, ~150 base pairs of DNA are wrapped around a central $(H3–H4)_2$ tetramer flanked by two H2A–H2B dimers. This nucleoprotein macromolecular assembly is a dynamic packaging system that must balance genome compaction with essential DNA processes including transcription, replication, and repair.[1,2] The intrinsic biophysical properties and stabilities of the histones are important components of chromatin dynamics. The cell can modulate these properties through post-translational modifications,[3] incorporation of histone variants,[4] ATP remodeling complexes,[5] and interactions with histone chaperones.[6] Chaperones, such as nucleoplasmin, Asf1, and Nap1, are important in histone deposition onto DNA and exchange of histone variants as well as preventing inappropriate protein–protein and protein–DNA interactions.[6]

**Corresponding author.* E-mail address: lmgloss@wsu.edu.

Present address: M. R. Stump, Division of Cardiovascular Medicine, Oregon Health Sciences University, Portland, OR, USA.

Abbreviations used: FL, fluorescence; KPi, potassium phosphate; NCP, nucleosome core particle; SF, stopped flow; WT, wild type; EDTA, ethylenediaminetetraacetic acid.

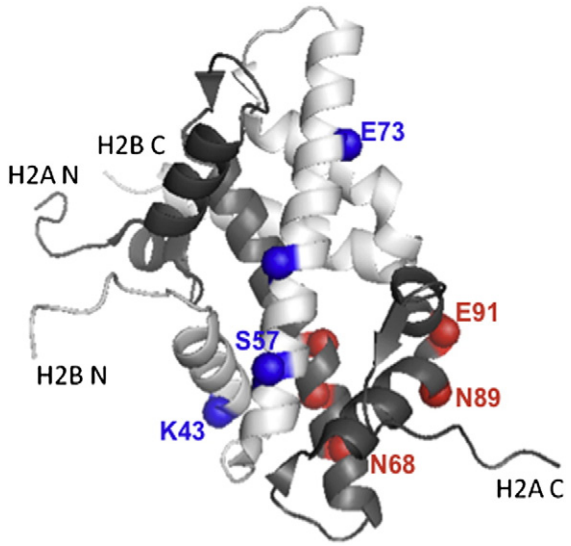
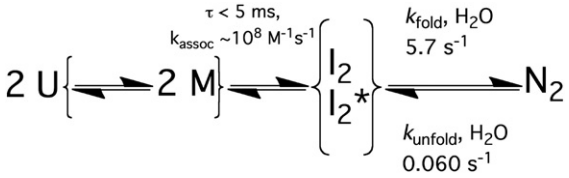


Fig. 1. The ribbon diagram of the H2A–H2B dimer. The H2A chain, showing residues 4–118, is colored gray, and the H2B chain, depicting residues 24–122, is white. The C^{α} atoms of the residues that were mutated to Ala and Gly are indicated by red (H2A) and blue (H2B) spheres. The figure was rendered using PyMOL (Delano Scientific, LLC, San Carlos, CA) using coordinates from the NCP X-ray crystal structure (1kx5.pdb).[42]

Recent studies have shown that some histone chaperones, such as Asf1, can induce significant structural changes in their target histones,[7,8] suggesting that partially folded dimeric conformations, such as those observed during protein folding, are physiologically relevant.

Oligomeric histones from eukaryotes and archaea contain an evolutionarily conserved dimerization motif composed of three helices: a central helix (α2) of ~30 residues, flanked on the N- and C-termini by a β-loop, and an α-helix of ~10 residues (α1, α3). The monomers dimerize in an antiparallel orientation, with extensive intermonomer hydrophobic interactions, particularly along the central α2 helices (Fig. 1). Eukaryotic histones have extended N-terminal tails, which are sites for the post-translational modifications (e.g., acetylation, methylation, and phosphorylation) of the "histone code."[3,9] C-terminal to the canonical histone fold, H2A contains a very short helix and an extended tail. The C-terminus of H2B has an additional α-helix of ~20 residues (αC) that docks on the N-terminal region of the H2A α2 helix and contributes to the hydrophobic dimer interface.

The eukaryotic histone heterodimers, H2A–H2B and H3–H4 [a tetrameric dimer of dimers, $(H3–H4)_2$], fold by a mechanism with at least three states (Scheme 1).[10,11] Unfolded monomers associate to form a dimeric intermediate, I_2, in the 5-ms stopped-flow (SF) mixing time. This obligatory I_2 species contains ~50% of the helical structure and ~50% of the buried surface area of N_2. The observed, first-order kinetic phase represents the conversion of this dimeric intermediate to the native dimer, N_2. Because association occurs with a relaxation time of



Scheme 1. Working mechanism for the kinetic folding of the H2A–H2B heterodimer. 2U, unfolded, dissociated H2A and H2B monomers; 2M, partially folded monomers, not directly observed by SF kinetics; I_2 and I_2*, ensemble of dimeric kinetic intermediates formed in the 5-ms SF dead time, detected by SF-CD burst-phase amplitude; N_2, native H2A–H2B heterodimer.

less than 5 ms at micromolar monomer concentrations, the estimated association rate must exceed 10^7 M^{-1} s^{-1}, approaching the diffusion limit. Since dimerization is not directly observed, it is unknown if monomers fold prior to dimerization (hence, the use of brackets around 2M in Scheme 1). However, the isolated H2A and H2B monomers are partially helical with marginal stability, and upon mixing, these monomers are kinetically competent to fold through I_2 to the native dimer.[12] Thus, the rapid folding of the eukaryotic histones proceeds through dimeric and presumably monomeric transient intermediates. Similar I_2 kinetic intermediates with comparable helical content, buried surface area, and stability have been observed in the folding of other intertwined, obligatorily domain-swapped α-helical dimers with differing topologies, namely, the dimerization core of the *Escherichia coli* Trp repressor[13,14] and the *E. coli* factor for inversion stimulation.[15]

This study examines the importance of helix propensity, particularly in the central α2 helix, on the stability of the I_2 species formed by H2A–H2B. Five residues in H2A and four residues in H2B (shown in Fig. 1 and listed in Tables 1 and 2) were mutated to both Ala and Gly to distinguish between effects of side-chain truncation and altered helix propensity. The mutation sites were chosen because of their high solvent accessibility in the folded dimer, so as to minimize packing or steric effects. Mutations were focused on the α2 helix because of its extensive contribution to the dimer interface. The effects of these mutations on the structure and stability of the isolated H2A and H2B monomers have been described elsewhere.[12]

Results

Kinetic folding from partially folded H2A and H2B monomers: I_2 to N_2

Previous refolding studies characterized SF reactions initiated from H2A and H2B monomers unfolded in 4 M urea.[11] The isolated H2A and H2B monomers can fold to marginally stable species with secondary and tertiary structure.[12] Equilibrium *m* values indicated that the partially folded H2A and

Table 1. Parameters describing the folding and unfolding kinetics of the H2A–H2B variants

Histone	$k_{unf}(H_2O)$ (s^{-1})	$m^{\ddagger}_{unf}$ (kcal mol^{-1} M^{-1})	$k_{fold}(H_2O)$ (s^{-1})	$m^{\ddagger}_{fold}$ (kcal mol^{-1} M^{-1})	$k_{0\ M}$[a] (s^{-1})	$k_{fold}/k_{0\ M}$
WT dimer[b]	0.06	−0.52	6.2	1.1	3.5	1.8
H2A mutants						
E61A α2	0.038	−0.60	13.2	1.19	6.6	2.0
E61G α2	0.054	−0.79	4.9	0.98	4.0	1.2
E64A α2	0.018	−0.81	8.1	0.86	1.6	5.0
E64G α2	0.024	−0.81	4.9	0.94	2.8	1.7
N68A α2	0.034	−0.77	6.1	0.98	4.2	1.4
N68G α2	0.13	−0.65	2.9	0.57	2.4	1.2
N89A α3	0.014	−0.82	4.1	0.82	3.4	1.2
N89G α3	0.012	−0.78	8.8	1.08	3.3	2.7
E91A αC	0.0097	−0.83	6.7	1.17	3.8	1.7
E91G αC	0.042	−0.69	5.9	1.09	3.7	1.6
H2B mutants						
K43A α1	0.018	−0.81	2.7	0.74	1.7	1.5
K43G α1	0.041	−0.81	6.7	0.99	4.1	1.6
S57A α2	0.022	−0.67	5.0	1.02	2.8	1.8
S57G α2	0.036	−0.61	6.0	1.54	2.5	2.4
N64A α2	0.154	−0.59	6.5	0.72	4.0	1.6
N64G α2	0.018	−1.02	3.2	0.71	2.5	1.3
E73A α2	0.023	−0.82	3.6	0.83	3.0	1.2
E73G α2	0.020	−0.98	1.5	0.91	1.5	1.0

Conditions: 200 mM KCl, 20 mM KPi pH 7.2, and 0.1 mM K_2EDTA, 25 °C, with a final monomer concentration of 7.5 μM. Folding kinetics were initiated by mixing the isolated monomers equilibrated separately at various urea concentrations. The $k(H_2O)$ and $m^{\ddagger}$ values are the result of the global fitting of the kinetic data to Eq. (4). Errors at one standard deviation were determined for the fitted parameters but are not shown for brevity. All errors associated with the $k(H_2O)$ values for folding and unfolding were less than 12% of the fitted parameter, with an average error of 4%. For the $m^{\ddagger}$ values, all errors were less than 8% with an average error of 2%.

[a] The $k_{0\ M}$ values are the folding rates determined from semi-global fits of multiple CD and FL kinetic traces for monomers pre-equilibrated in the absence of denaturant. The associated errors are less than 10%. The ratio $k_{fold}/k_{0\ M}$ is a measure of the extent of rollover observed in the folding kinetics.

[b] The WT unfolding data are from Ref. 11.

H2B monomers are overly collapsed relative to the extended structures observed in the native heterodimer and presumably contain non-native structure, that is, interactions that stabilize the partially folded species via contacts that are different from those present in the native state. Nonetheless, upon mixing, the isolated monomers were kinetically competent to proceed to the native dimer through a burst-phase dimeric intermediate, as shown in Scheme 1.[12] This report expands upon the previous study by examining the urea dependence of folding initiated from partially folded H2A and H2B monomers.

Isolated monomers were pre-equilibrated at varying initial urea concentrations, and upon mixing, the monomers were allowed to fold to the heterodimer. Very similar results were obtained when isolated monomers were pre-equilibrated at 0 M urea and refolded to different final urea concentrations (data not shown). The folding reactions were monitored by far-UV circular dichroism (CD) and intrinsic Tyr fluorescence (FL). The rates from local fits of individual SF-FL and SF-CD folding traces were in excellent agreement, demonstrating the concerted formation of helices (secondary structure) and burial of Tyr residues (tertiary and quaternary structure). The CD and FL kinetic data were semi-globally fit, linking the rates across all kinetic traces at a given urea concentration (symbols in Fig. 2). Above ~0.4 M urea, there was a log-linear dependence of the rates on the final urea concentration. Therefore, these CD and FL kinetic responses as a function of final urea concentration were globally fit to Eq. (4) (Methods). The parameters $k_{fold}(H_2O)$ and $m^{\ddagger}_{fold}$ were linked across all kinetic traces (for urea concentrations ≥0.4 M); the results are represented by the continuous line in Fig. 2. The resulting fitted values (Table 1) were in excellent agreement with previous results for refolding of the H2A–H2B dimer unfolded in 4 M urea.[11]

Below 0.4 M urea, the rates for folding from low urea concentrations exhibit a "rollover" or deviation from a log-linear dependence on the urea concentration; these are lower urea concentrations than were accessible by SF dilution from 4 M urea in the previous report.[11] This rollover indicates that there is a change in the rate-determining step below 0.4 M urea. One possibility is that the burst-phase dimerization reaction becomes partially rate-determining at lower urea concentrations. This explanation is discarded because (1) a similar deviation from log-linear dependence is observed at final monomer concentrations of 7.5 and 15 μM, indicating that the cause of the rollover is independent of protein concentration, and (2) the minor protein concentration dependence of the observed rates are similar at 0 M urea (data not shown) and 0.5 M urea.[11] An alternative explanation is that the burst-phase ensemble contains two or more conformations, denoted I_2 and I_2* in Scheme 1, and their relative populations change as a function of urea. The folding rates between 0 and 0.4 M urea exhibit a low apparent $m^{\ddagger}$ value, consistent with a reaction in which there is limited change in solvent-accessible

Table 2. Comparison of mutational effects on equilibrium and kinetic parameters

Histone	$\Delta\Delta G_{equil}$[a]	ΔC_M[a] (M urea)	$\Delta\Delta G^{\ddagger}$ unfold[b]	β-Value[c]	$\Delta\Delta G$ I_2–N_2[d]	$\Delta\Delta G$ 2U–I_2[e]	m 2U–I_2[f]
WT dimer	[11.8]	[1.66]	—	0.18	[2.75]	[9.0]	1.3
H2A mutants							
E61A α2	0.17	0.0	−0.3	0.21	−0.7	0.9	1.0
E61G α2	1.5	0.49	−0.07	0.28	0.07	1.4	1.1
E64A α2	−0.47	−0.32	−0.7	0.30	−0.9	0.4	1.0
E64G α2	0.59	0.22	−0.5	0.28	−0.4	1.0	1.2
N68A α2	0.95	0.25	−0.3	0.28	−0.3	1.3	1.0
N68G α2	1.8	0.57	0.5	0.23	0.9	0.8	1.6
N89A α3	0.58	0.18	−0.8	0.29	−0.6	1.2	1.2
N89G α3	0.43	0.15	−1.0	0.27	−1.2	1.6	1.0
E91A αC	0.38	0.05	−1.1	0.30	−1.1	1.5	0.8
E91G αC	0.78	0.25	−0.2	0.24	−0.2	1.0	1.1
H2B mutants							
K43A α1	0.87	0.12	−0.7	0.32	−0.2	1.1	1.0
K43G α1	0.93	0.27	−0.2	0.29	−0.3	1.2	1.0
S57A α2	0.45	0.0	−0.6	0.26	−0.5	0.9	0.9
S57G α2	0.84	0.23	−0.3	0.22	−0.3	1.1	0.6
N64A α2	0.63	0.18	0.6	0.21	0.5	0.1	1.5
N64G α2	1.8	0.62	−0.7	0.35	−0.3	2.1	1.2
E73A α2	0.48	0.14	−0.6	0.29	−0.2	0.7	1.2
E73G α2	1.5	0.50	−0.7	0.34	0.2	1.3	1.0

Conditions are described in the legend to Table 1. $\Delta\Delta G$ values are expressed as kilocalories per mole. The WT $\Delta G°$ values are given in brackets at the top of the columns, which tabulate the $\Delta\Delta G$ values.

[a] $\Delta\Delta G_{equil} = \Delta G°(H_2O)_{WT} - \Delta G°(H_2O)_{mutant}$; the $\Delta G°(H_2O)$ values and the data for their determination are described in Ref. [12]. A positive value indicates that the mutant is destabilizing. The ΔC_M values ($C_{M\text{-}WT} - C_{M\text{-}mutant}$) were calculated at 7.5 μM monomer, the typical concentration used in the kinetic experiments, using $C_M = \Delta G°(H_2O) + (RT \cdot \ln[\text{monomer}])/m$.

[b] $\Delta\Delta G^{\ddagger}$ for unfolding $= -RT \cdot \ln[(k_{unf}(H_2O)_{WT})/(k_{unf}(H_2O)_{mutant})]$; the $k_{unf}(H_2O)$ values are given in Table 1. Positive $\Delta\Delta G^{\ddagger}$ values correspond to faster unfolding by the mutant.

[c] The unitless β-value $= m^{\ddagger}_{unf}/m_{equil}$, where m_{equil} is the value determined from equilibrium studies.[12]

[d] $\Delta G°(H_2O)$ I_2–N_2 represents the free energy change, in the absence of denaturant, for unfolding of N_2 to I_2 as defined in Eq. (2a). The $\Delta\Delta G$ for I_2–$N_2 = \Delta G°(H_2O)_{WT,\ I2\text{–}N2} - \Delta G°(H_2O)_{mutant,\ I2\text{–}N2}$, so that a negative value indicates that the mutant exhibits a greater free energy difference between I_2 and N_2 than WT.

[e] $\Delta G°(H_2O)$ 2U–I_2 represents the free energy change, in the absence of denaturant, for unfolding of I_2 to 2U as defined in Eq. (2b). The $\Delta\Delta G$ for 2U–$I_2 = \Delta G°(H_2O)_{WT,\ 2U\text{–}I2} - \Delta G°(H_2O)_{mutant,\ 2U\text{–}I2}$, so that a positive value indicates that the I_2 species of the mutant is destabilized relative to WT.

[f] The m 2U–I_2 value was calculated according to Eq. (2c). For consistency, the calculated WT is given, rather than the fitted value from SF-CD burst-phase analyses.

surface area (ΔASA). Thus, the non-native structure present in the isolated monomers may persist in the I_2* species and is resolved by an isomerization-like rearrangement. Rollover at low denaturant concentrations from the formation of overly or prematurely collapsed intermediate states has been observed previously (e.g., see Ref. 16).

Equilibrium effects of mutations that alter helix propensity

The stabilities of the mutant H2A–H2B heterodimers were determined from urea-induced unfolding titrations. The data collection and analysis are described elsewhere[12] and summarized briefly here. All mutant histones exhibited cooperative, two-state, highly reversible equilibrium transitions as observed for wild-type (WT) H2A–H2B.[17] The effects of the mutations were evaluated by two parameters (Table 2): (1) the $\Delta\Delta G_{equil}$ values, that is, $\Delta G°(H_2O)_{WT} - \Delta G°(H_2O)_{mutant}$, where $\Delta G°(H_2O)$ is the free energy of unfolding in the absence of denaturant, and (2) ΔC_M values, that is, the difference between the WT and mutant C_M values, where C_M is the urea concentration at the midpoint of the unfolding transition. H2A-E64A was the only mutation that stabilized the dimer. Ala substitutions at H2A-Glu61, H2A-Glu91, and H2B-S57 had minor effects with WT-like C_M values and $\Delta G°(H_2O)$ values within 5% of WT. The other five Ala mutations were destabilizing by 0.5 to 1.0 kcal mol^{-1}. In general, the Gly mutations were more destabilizing than Ala at the same positions; the exceptions were H2A-N89A/G, a helix-capping residue, and H2B-K43A/G, where the Ala and Gly substitutions were similarly destabilizing.

Unfolding kinetics of the H2A–H2B mutant heterodimers

The unfolding kinetics of the mutant heterodimers were studied by SF-FL and SF-CD. As observed for WT H2A–H2B,[11] there were no detectable burst-phase unfolding reactions, and the observed kinetics were well described by a single, first-order exponential with excellent agreement between CD and FL data. The semi-globally fitted rates exhibited a log-linear dependence on the final urea concentration (Fig. 3 and Supplementary Material). The urea dependence of the unfolding responses was analyzed by global fits of the kinetic traces to Eq. (4), and the fitted parameters are given in Table 1.

Unfolding rates were compared at 3 M urea (in the equilibrium unfolding baseline of all variants) and

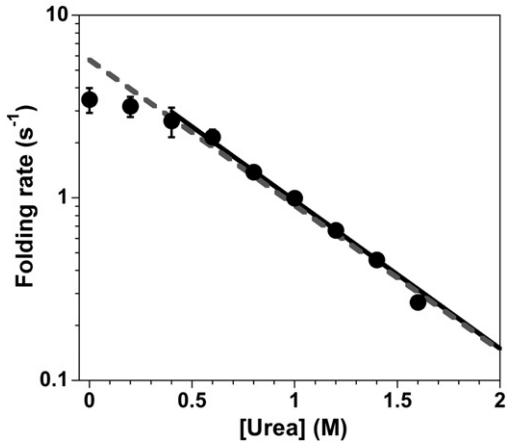


Fig. 2. Rates for the folding of the I_2 kinetic intermediate to the native H2A–H2B dimer. Refolding was initiated from isolated monomers equilibrated at varied initial urea concentrations. Data points represent semi-global fits of multiple SF-CD and SF-FL traces, and the continuous line represents the global fits to Eq. (4) of SF-CD and SF-FL traces from 0.4 to 1.6 M urea. Errors are shown or are smaller than the size of the data points. The gray broken line represents the previously published global fit for refolding of the H2A–H2B dimer from 4 M urea.[11] Conditions: final monomer concentration of 7.5 μM, 200 mM KCl, 20 mM KPi, pH 7.2, and 0.1 mM K_2EDTA, 25 °C.

in the absence of urea, $k_{unf}(H_2O)$, the globally fitted unfolding rate extrapolated to the absence of denaturant. The Ala mutants generally unfold with similar or slower rates than the corresponding Gly mutant. This trend is consistent with the Gly mutations generally being more destabilizing than the Ala mutations. The noteworthy exceptions are H2A-N89G (1.2- to 1.5-fold slower than N89A) and H2B-N64G, which has an 8.5-fold lower $k_{unf}(H_2O)$ value than N64A.

In 3 M urea, 12 mutants exhibited unfolding rates within 2-fold of the WT rate, corresponding to $\Delta\Delta G^{\ddagger}_{unf}$ values $\leq$0.4 kcal mol^{-1} (Fig. 4a). Only six mutants exhibited 3- to 4-fold higher unfolding rates than WT. These mutational effects are in striking contrast to the trends observed for the $k_{unf}(H_2O)$ values (Table 1; Fig. 4a). Only two mutants, H2A-N68G and H2B-N64A, exhibit larger $k_{unf}(H_2O)$ values than WT, while 10 mutants unfold $\geq$2.5-fold more slowly than WT, a finding that is intuitively inconsistent with destabilizing mutations. In a simple two-state kinetic mechanism, one would expect destabilizing mutations to unfold with rates similar to or faster than WT.

The contrasting changes in the unfolding rates at 3 and 0 M urea reflect changes in the $m^{\ddagger}_{unf}$ values, the slope of the log-linear urea dependence of the unfolding rates [Eq. (4)]. The mutants have $m^{\ddagger}_{unf}$ values equal to or greater than WT (Table 1). The kinetic $m^{\ddagger}$ values usually correlate with the ΔASA between the ground state and the transition state. Larger $m^{\ddagger}_{unf}$ values suggest that the transition states traversed by the mutant heterodimers are more unfolded-like, with greater exposure of surface area than observed in the WT dimer. The Tanford β-value (Table 2) describes the position of the unfolding transition state with respect to the burial of the surface area achieved upon folding:

$$\beta = \frac{m^{\ddagger}_{unf}}{m_{equil}} \quad (1)$$

where m_{equil} is the m value determined from equilibrium experiments. The β-value can vary from 0 to 1, reflecting very native-like and unfolded-like transition states, respectively. The WT β-value is 0.18, implying that ~20% of the surface area exposed upon unfolding is solvent-accessible in the rate-determining unfolding transition state between N_2 and I_2. The larger β-values of the mutants imply that the ΔASA of the mutants is 1.16- to 1.9-fold greater than WT.

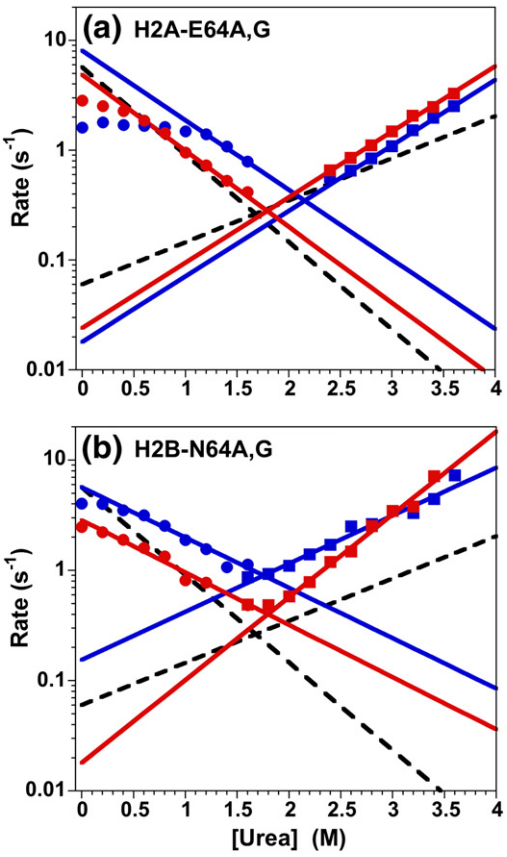


Fig. 3. Representative plots of the urea dependence of the folding (circles) and unfolding (squares) rates for mutant heterodimers with Ala (blue) and Gly (red) substitutions. (a) H2A-E64A/G. (b) H2B-N64A/G. The data points represent semi-global fits of multiple SF-FL and SF-CD kinetic traces at a given final urea concentration, and the associated errors are equal to or less than the size of the data symbols. The global fits of the data to Eq. (4) are shown as lines. The results of the previously published WT fits are shown as black broken lines. Conditions are described in the legend to Fig. 2. Chevron plots for other mutant heterodimers are presented in the Supplementary Data.

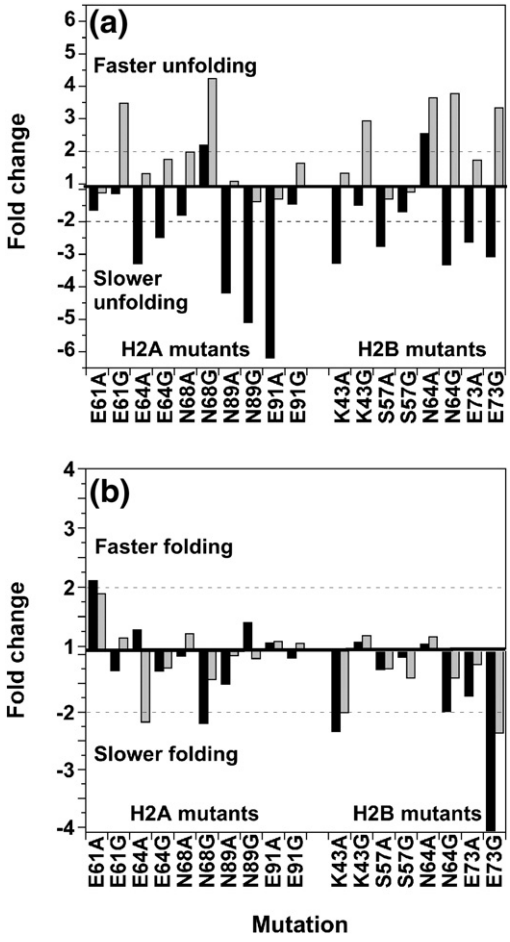


Fig. 4. Effect of mutations on the unfolding and folding rates. (a) Comparison of the fold change in the unfolding rates at 3 M urea (shaded bars) and extrapolated to the absence of denaturant (filled bars). Positive values represent the fold increase in the mutant's unfolding rate (k_{mutant}/k_{WT}), that is, faster unfolding; negative values indicate slower unfolding (k_{WT}/k_{mutant}). (b) Comparison of the fold change in the folding rates at 0 M urea (shaded bars) and extrapolated to the absence of denaturant (filled bars). Positive values represent an increase in the mutant's folding rate (k_{mutant}/k_{WT}), that is, faster folding; negative values indicate slower folding (k_{WT}/k_{mutant}).

Folding kinetics of the H2A–H2B mutant heterodimers

The mutant heterodimers were refolded from isolated monomers, as described above for WT (Fig. 2). Like WT H2A–H2B,[11,12] the folding kinetics of all mutants had the following attributes: (1) the kinetic responses were well fit by a single first-order exponential at all urea concentrations, (2) the fitted rates exhibited little protein concentration dependence, (3) similar rates were measured by SF-FL and SF-CD, (4) there was a substantial SF-CD burst-phase amplitude, (5) a log-linear relationship between the observed rate and the urea concentration at higher urea concentrations, and (6) the folding and unfolding rates appeared to converge in the equilibrium transition region. These attributes demonstrate that the mutants fold by the same mechanism as WT (Scheme 1). The kinetic responses of the mutants were analyzed semi-globally and globally as described for WT, and the fitted parameters are given in Table 1. Data from representative mutants are shown in Fig. 3; the data for the other mutants are presented in the Supplementary Material.

In general, the mutations have much less effect on the folding rates than the unfolding rates. The change in the folding rates measured at 0 M urea and the extrapolated $k_{fold}(H_2O)$ values are shown in Fig. 4b. Thirteen mutations exhibit $k_{fold}(H_2O)$ and $k_{0\ M}$ rates that are within 2-fold of the WT values. Except for H2A-E64A, the trends for $k_{fold}(H_2O)$ and $k_{0\ M}$ are similar (either little effect or change in the same direction); however, the observed effects are usually greater for the $k_{fold}(H_2O)$ value. Destabilization can be manifested as slower folding; however, only four mutations exhibit $k_{fold}(H_2O)$ values that are 2- to 4-fold slower than WT.

The ratio of $k_{fold}(H_2O)/k_{0\ M}$ (Table 1) provides a measure of the extent of rollover at low urea concentrations. WT and most mutants have ratios ≤ 2.0. This is consistent with visual inspection of the chevrons (Fig. 3; Supplementary Figs. S1 and S2), which suggests that most mutations do not significantly enhance the extent of rollover, and many mutations decrease it. Only the stabilizing H2A-E64A mutation significantly enhances the rollover (Fig. 3a).

The ΔΔ*G* values of the H2A–H2B mutant heterodimers

SF-CD burst-phase analyses were used to determine the $\Delta G°(H_2O)$ and m values describing the stability of the WT I_2 ensemble.[11] These fitted values were very similar to the values calculated from the parameters for equilibrium unfolding transitions and kinetic folding and unfolding experiments, using the following equations:

$$\Delta G°(H_2O)_{I_2-N_2} = -RT\ln\left(\frac{k_{unf}(H_2O)}{k_{fold}(H_2O)}\right) \quad (2a)$$

$$\Delta G°(H_2O)_{2U-I_2} = \Delta G°(H_2O)_{equil} - \Delta G°(H_2O)_{I_2-N_2} \quad (2b)$$

$$m_{2U-I_2} = m_{equil} - \left(m^{\ddagger}_{fold} - m^{\ddagger}_{unf}\right) \quad (2c)$$

The stabilities of the I_2 ensembles for the mutants were determined by calculation rather than burst-phase analyses because of the higher precision and greater technical ease. The $k_{fold}(H_2O)$ values were employed in these calculations, rather than the $k_{0\ M}$ values (Table 1), because the reaction described by $k_{fold}(H_2O)$ converges with the unfolding reaction [described by $k_{unf}(H_2O)$] in the transition region. In other words, the principle of microscopic reversibility argues that $k_{fold}(H_2O)$ and $k_{unf}(H_2O)$ are the appropriate parameters to describe the transition state between I_2 and N_2 and the free energy

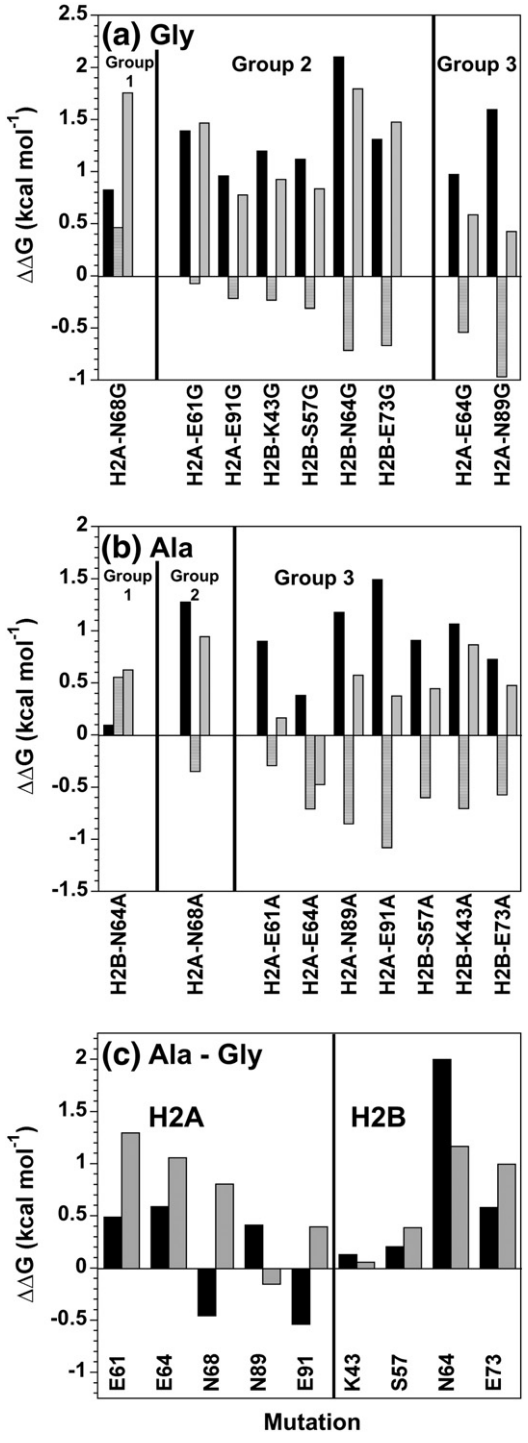


Fig. 5. Comparison of the $\Delta\Delta G$ values for the H2A and H2B mutations. $\Delta\Delta G$ 2U-to-I_2, black, left bars; $\Delta\Delta G^{\ddagger}_{unf}$, hatched, middle bars; $\Delta\Delta G_{equil}$, gray, right bars. The determination of the values is described in the legend to Table 2. Mutants are segregated into three groups as described in the text. (a) WT-Gly values. (b) WT-Ala values. (c) Ala-Gly values. A positive value indicates that the Gly mutation is more destabilizing than the Ala reference state. The span of the *y*-axis (3.5 kcal mol^{-1}) is the same in all panels.

difference between these two species. If $k_{0\ M}$ values are used in Eq. (2a), the conclusions described below are not significantly altered.

Except for H2B-N64A, all mutations destabilize the I_2 species ≥ 0.4 kcal mol^{-1}, with an average $\Delta\Delta G_{2U-I2}$ of ~ 1 kcal mol^{-1} (Table 2). Thus, residues that contribute significantly to the stability of I_2 are distributed across the primary structure of H2A and H2B. The mutant 2U–I_2 *m* values are generally slightly lower than but within 25% of the WT value calculated from Eq. (2c), suggesting that the mutations do not greatly alter the amount of surface area buried in the I_2 ensemble. However, the β-values [Eq. (1)] indicate that the mutations shift the I_2-to-N_2 transition states toward I_2.

The $\Delta\Delta G$ values for different species and states along the folding reaction coordinate (dimer equilibrium, I_2, and unfolding transition state) are compared in Fig. 5. In mutational analyses, the free energy of the unfolded state of the WT and mutants are both typically set to zero (e.g., see the reaction coordinate diagrams in Fig. 6). Thus, the observed $\Delta\Delta G$ values are attributed to removing an interaction that stabilizes a folded or partially folded state, that is, destabilizing N_2, I_2, or the transition state connecting them. However, the $\Delta\Delta G$ values can also reflect introduction of an interaction that stabilizes the unfolded state but is lost during the folding reaction. Whether the equilibrium effect of the mutation is on the N_2 or 2U state, comparing $\Delta\Delta G$ values for different species along the folding coordinate indicates the extent to which the mutated residue participates in stabilizing interactions.

Phi-value analyses are a common quantitative description of the extent of native interactions for an intermediate or transition state *i* where $\phi = \Delta\Delta G_i / \Delta\Delta G_{equil}$. Typically, φ-values range from 0 (indicating that the native-like stabilizing interactions of the residue are not significantly formed in state *i*) to 1 (implying that the native-like interactions are fully formed in state *i*). However, because of the altered β-values, the ϕ_{TS} values calculated with $\Delta\Delta G^{\ddagger}_{unf}$ are predominantly <0. Furthermore, there has been debate about the validity of Phi-value analyses when the $\Delta\Delta G_{equil}$ values are relatively small;[18–20] the various limiting values that have been proposed would exclude 9–16 of the 18 mutations. Therefore, the following analyses focus on the magnitudes of $\Delta\Delta G$ values, rather than specifically on their ratio.

Mutation of a residue to Gly removes side-chain interactions and decreases helix propensity; thus, the Gly $\Delta\Delta G$ values are indicative of the total stabilizing potential of the WT residue. The major effect of Ala mutations is removal of side-chain interactions. Comparison of the effects of Ala and Gly mutations at a given residue, that is, Ala-Gly $\Delta\Delta G$ values, reports on the contribution of helix propensity to stability using a uniform change in helix propensity (Ala *versus* Gly) across the set of mutations.

A key finding of this study is that several residues in H2A and H2B stabilize the burst-phase I_2 ensemble by non-native structure, with mutational effects of 0.5 to 1 kcal mol^{-1}. The basis for describing

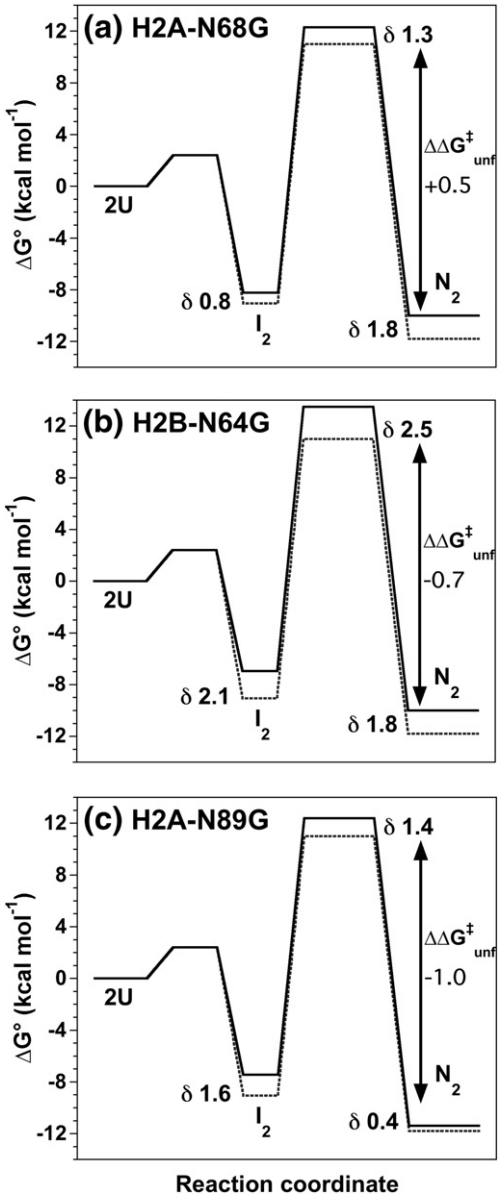


Fig. 6. Reaction coordinate diagrams describing three classes of Gly mutations, grouped as shown in Fig. 5. The WT reaction coordinate is shown as a black dotted line. The $\Delta G°$ value of the unfolded species for both WT and the mutants were arbitrarily set to zero. The energies of the WT transition states were estimated from the Kramers formalism as described previously[14] and should be regarded as illustrative values. The 2U–I_2 dimerization transition state energy was calculated with a pre-exponential factor of 6×10^9 M^{-1} s^{-1} and an estimated rate of 10^8 M^{-1} s^{-1}; since this reaction is too rapid to be directly measured by SF-FL, the barrier is shown as the same for WT and the mutants. The first-order I_2 to N_2 $\Delta G^{‡}$ was calculated for a 5×10^8 s^{-1} prefactor and rate of 6.2 s^{-1}. The changes in ΔG (kcal mol^{-1}) between the WT and mutant states are denoted by δ values. (a) H2A-N68G, equivalent to a ϕ-value of ~0.5; (b) H2B-N64G, representing ϕ-values between 0.9 and 1.3; (c) H2A-N89G, indicative of non-native structure in I_2 and the rate-limiting transition state.

structure as non-native is that the Ala and/or Gly mutants exhibited (1) $\Delta\Delta G_{2U–I2}$ values exceeding the $\Delta\Delta G_{equil}$ values by more than 0.3 kcal mol^{-1} and/or (2) the $\Delta\Delta G^{‡}_{unf}$ values $\ll 0$. These effects are highlighted by the Gly and Ala Group 3 mutants described below (Fig. 5a and b). The presence of non-native interactions is more obvious in the Ala mutations. There is a possible caveat regarding these criteria. It is conceivable that disruption of a partially formed cluster of native interactions, for example, a network of salt bridges, in an intermediate species could potentially result in a $\Delta\Delta G_i$ value of different sign or greater magnitude than disruption of the cluster in the native state, $\Delta\Delta G_{equil}$.

Gly ΔΔG values

The Gly mutations can be segregated into three groups: $\Delta\Delta G_{2U–I2}$ less than, approximately equal to, or greater than the $\Delta\Delta G_{equil}$ values (Fig. 5a). The first condition describes only H2A-N68G, the only Gly mutation with a greater $k_{unf}(H_2O)$ value than WT (Fig. 4a). The stabilizing interactions contributed by Asn68 are only partially realized in the I_2 ensemble and the rate-limiting transition state between I_2 and N_2. These effects are illustrated in the reaction coordinate diagram in Fig. 6a.

The second group of mutations, E61G and E91G in H2A and all four H2B variants, has similar $\Delta\Delta G_{2U–I2}$ and $\Delta\Delta G_{equil}$ values, suggesting that the interactions formed by these residues in the I_2 species contribute comparable stability to those formed in the native state. The small $\Delta\Delta G^{‡}_{unf}$ values (−0.1 to −0.3 kcal mol^{-1}) for E61G and E91G in H2A and K43G and S57G in H2B demonstrate that the interactions present in I_2 are largely maintained in the transition state. In contrast, N64G and E73G in the middle and C-terminal end of the long α2 helix of H2B exhibit significantly negative $\Delta\Delta G^{‡}_{unf}$ values (−0.7 kcal mol^{-1}). These mutations destabilize the I_2 and N_2 species to a similar extent but cause substantially greater destabilization of the transition state. The WT residues may contribute to the stability of I_2 through some non-native structure, and these stabilizing interactions are broken in the transition state leading to N_2. These effects are summarized in the reaction coordinate diagram in Fig. 6b.

The third group of Gly mutations, E64G and N89G in H2A, exhibit $\Delta\Delta G_{2U–I2}$ values significantly greater than the corresponding $\Delta\Delta G_{equil}$ values and $\Delta\Delta G^{‡}_{unf}$ values $\ll 0$ (Fig. 5a). These residues appear to contribute similar non-native stabilization to I_2 and the transition state (e.g., Fig. 6c).

Ala ΔΔG values

The Ala mutations also segregate into three classes: $\Delta\Delta G_{2U–I2} \sim 0$, approximately equal to, or greater than the $\Delta\Delta G_{equil}$ values (Fig. 5b). The first class, namely, H2B-N64A, appears to not contribute to the stability of I_2, in contrast to H2B-N64G (Group 2). The N64A mutation does have a significant effect

on the transition state, with similar $\Delta\Delta G_{equil}$ and $\Delta\Delta G^{\ddagger}_{unf}$ values (Fig. 5b). Taken together, the Ala and Gly data suggest that helix propensity is the major stabilizing facet for I_2 (diminished by N64G, but maintained or enhanced inN64A), and stabilizing side-chain interactions (removed by the N64A mutation) are not realized until the transition state.

The second class of Ala mutants (only H2A-N68A) is similar to the Gly Group 2 mutations. H2A-N68G was the only Gly Group 1 mutation. Of the nine residues studied, only Asn68, at the C-terminal end of the H2A central helix, does not meet the criteria listed above to indicate non-native structure. However, despite the expected increase in helix propensity of replacing Asn with Ala,[21] N64A destabilizes the I_2 ensemble more than N64G (Table 2). This may imply that greater helicity near the end of α2 inhibits formation of I_2-stabilizing non-helical (presumably non-native) interactions formed by neighboring residues.

Like Gly Group 3, the third group of Ala mutations indicates significant non-native structure in I_2($\Delta\Delta G_{2U-I2} \gg \Delta\Delta G_{equil}$) and/or the transition state between I_2 and N_2 ($\Delta\Delta G^{\ddagger}_{unf} \ll 0$), resulting in reaction coordinate diagrams similar to Fig. 6c. These seven variants can be further sub-divided with respect to their contributions to non-native structure in I_2 and transition state.

Like their corresponding Gly variants (Group 3), H2A, E64A, and N89A contribute non-native structure to both I_2 and the transition state. For Glu64, the stabilization of I_2 reflects both side-chain interactions and helix propensity (E64G is more destabilizing than E64A). The E64A mutation drastically increases the extent of rollover at low urea (Fig. 3a), suggesting that the mutation stabilizes the non-native I_2* component of the burst-phase ensemble. The difference between the Asn89 Ala and Gly $\Delta\Delta G_{2U-I2}$ values is small relative to the destabilization of N89G, and therefore, the I_2-stabilizing interactions are largely side chain mediated, as one might expect for a helix-cap residue.

Non-native interactions in I_2 and the transition state are indicated for H2A-E91A, H2A-E61A, and H2B-S57A, but the corresponding Gly mutations are in Group 2, with little indication of non-native structure. Truncation of the H2A-Glu61 side chain significantly destabilizes the I_2 ensemble but has little effect on N_2, a difference that is mitigated by diminished helix propensity (E61G). ForH2A-Glu91, Ala destabilizes I_2 more than Gly, suggesting that decreased helicity may favor non-native structure. For H2B-S57, Ala and Gly destabilize I_2 similarly, demonstrating the importance of side-chain interactions, although helix propensity in the N-terminal section of H2B-α2 may favor non-native structure.

H2B K43A and E73A destabilize I_2 and the transition state, but like their corresponding Gly mutations (Group 2), there is no direct indication of non-native structure in I_2 ($\Delta\Delta G_{2U-I2} \sim \Delta\Delta G_{equil}$). However, Ala $\Delta\Delta G^{\ddagger}_{unf}$ values of −0.6 to −0.7 kcal mol^{-1} imply disruption of non-native structure in the transition state. For H2B-Glu73, Ala and Gly $\Delta\Delta G_{2U-I2}$ values indicate that both helix propensity and side-chain interactions contribute to I_2 stability, but similar Ala and Gly $\Delta\Delta G^{\ddagger}_{unf}$ values suggest that non-native structure in the transition state is mediated by side-chain interactions, presumably electrostatic, given the basic nature of the histone proteins. At H2B-K43, the Ala and Gly effects indicate that the side chain, rather than helix propensity, is the major determinant of stability at this position for I_2 and N_2.

Ala-Gly ΔΔG values

To assess the importance of helix propensity, we compare in Fig. 5c the effects of Ala and Gly mutations on the $\Delta\Delta G_{equil}$ and $\Delta\Delta G_{2U-I2}$ values. Positive values indicate that Ala is the less destabilizing substitution. Helix propensity has a significant impact on equilibrium stability in the α2 helix of H2A (Glu61, Glu64, and Asn68) and at the center and C-terminal end of the H2B α2 helix (Asn64 and Glu73). In contrast, the other sites probed by mutation have Ala-Gly $\Delta\Delta G$ values ≤0.4 kcal mol^{-1}.

For H2A-Glu61 and Glu64 and H2B-Glu73, the stabilization achieved by increased helix propensity is only partially realized in I_2 (Ala-Gly $\Delta\Delta G_{equil} > \Delta\Delta G_{2U-I2}$). Conversely, helix propensity has a greater effect on I_2 than N_2 at Asn64 in the center of H2B's α2 helix. As noted above, at H2A-Asn68 and Glu91, Ala, relative to Gly, stabilizes N_2, but destabilizes I_2, suggesting that helix propensity might inhibit the formation of stabilizing non-native interactions.

Folding–unfolding double-jump experiments

The possible folding mechanisms that could explain the rollover observed at low denaturant concentrations (Figs. 2 and 3) and the relationship between I_2 and I_2* are diagrammed in Supplementary Scheme 1. A parallel mechanism, where I_2 and I_2* are discrete intermediates that both lead to N_2, is inconsistent with the observation that the kinetic responses of WT and all 18 mutants are very well described by a single, first-order exponential at all urea concentrations. For example, two distinct rates should be distinguishable for: (1) H2A-E64A, exhibiting the most pronounced rollover, with a predicted 5-fold difference between the observed $k_{0\ M}$ and extrapolated $k_{fold}(H_2O)$ rates; or (2) WT at ≥1.2 M urea, where the predicted difference between the extrapolated $k_{0\ M}$ and observed folding rate should be ≥5-fold.

To distinguish between other potential mechanisms (Supplementary Scheme 1b–d), we performed SF folding–unfolding double-jump experiments at 0 M urea with WT and H2A-E64A. Overall, the results for both variants were the same as reported previously for WT refolding to 1 M urea.[11] Firstly, no unfolding amplitude (accumulation of N_2) was observed at delay times of ~10 ms, demonstrating that the burst-phase I_2 ensemble is an obligatory intermediate. Secondly, the increase in unfolding

amplitudes was well described by a single exponential with no indication of a lag phase in the accumulation of N_2 (Supplementary Fig. 3). The absence of any lag phase or bi-exponential response casts doubt on potential sequential mechanisms where I_2* must pass through I_2 to reach N_2 (Supplementary Scheme 1b and c). Thirdly, the rates describing the exponential increase in unfolding amplitude for WT and H2A-E64A were in excellent agreement with the rates observed for direct refolding experiments at the same final conditions (details in Supplementary Fig. 3). Furthermore, these double-jump rates were substantially lower than the extrapolated $k_{fold}(H_2O)$ values. These results demonstrate that I_2* can fold directly to the native heterodimer and suggest that there is no significant kinetic barrier between I_2* and I_2. Rather, the data lend themselves to the interpretation that I_2* and I_2 represent alternative populations in a broad ensemble of species. These experiments cannot differentiate between two possibilities (diagrammed in Supplementary Scheme 1d): (1) a broad ensemble of ground-state species formed in the SF burst phase (e.g., I_2 and I_2*) whose relative population shifts toward I_2 as urea concentrations increase, presumably by destabilizing the non-native components of the I_2* end of the ensemble, or (2) a more narrow ground-state ensemble that traverses a broad transition state whose rate-limiting features move as a function of urea concentration.

Discussion

Expanded kinetic mechanism for the folding of the H2A–H2B heterodimer

Our current working mechanism shown in Scheme 1 adds new details to the previously published models.[11,12] First, Fig. 2 verifies that the isolated monomers at equilibrium are kinetically competent to fold to N_2 across a range of urea concentrations. Second, the rollover observed below 0.4 M urea shows that obligatory I_2 burst-phase ensemble contains I_2* species with some degree of non-native structure that is destabilized by low concentrations of urea. An equally plausible but indistinguishable interpretation is that the I_2 ensemble folds to N_2 over a broad, rough transition state that shifts as a function of urea concentration. Third, residues that contribute to the stability of the I_2 ensemble are widespread across the sequences of H2A and H2B. Furthermore, except for H2A-Asn68, the residues targeted for mutation appear to be involved in stabilizing interactions that are non-native to some extent.

Burial of solvent-accessible surface area in I_2 and the transition state to N_2

Of the 18 mutations, 16 exhibit m values associated with the folding of 2U to I_2 that are less than that of WT, although most m_{2U-I2} values are within 0.3 kcal mol^{-1} M^{-1} (~25%) of the WT value. Thus, the solvent-accessible surface area buried upon formation of the I_2 ensemble is only modestly diminished in response to destabilization.

All 18 mutants exhibit $m^{\ddagger}_{unf}$ values (Table 1) with absolute values greater than that of WT and thus higher β-values as well (Table 2). These results indicate that the transition state ensemble between N_2 and I_2 is shifted toward the I_2 species with respect to the amount of buried surface area. This plasticity of the N_2-to-I_2 transition state upon mutation has been observed previously with deletion of the N-terminal tails of H2A and H2B.[11] Transition states that are relatively resistant to mutational effects are the foundation of the Phi-value analyses applied to several small monomeric proteins (for review, see Ref. 22). There are noteworthy exceptions, such as an immunoglobulin domain from human cardiac titin, TI I27,[23] where several destabilizing mutations shift the unfolding transition away from the native state, as observed for the H2A and H2B mutations studied here. This direction of movement of the transition state upon mutation is consistent with the effects observed for the stabilizing ΔN-H2B and destabilizing ΔN-H2A mutations.[11]

Structure in the dimeric kinetic I_2 ensemble

Of the 129 and 122 residues in H2A and H2B, respectively, ~125 residues are in well-structured regions (Fig. 1). From these residues, those targeted for mutation were chosen based on their solvent accessibility, focusing primarily on the long central α2 helices. The choices were further narrowed down by the helix propensity predictions of AGADIR,[24] summarized in Supplementary Fig. 4. In H2A, only the central 17 residues of α2 and the short, seven-residue αC are predicted to have substantial helix propensity. The AGADIR predictions for H2B are generally lower than those for H2A, in contrast to the experimental results where H2B is more helical.[12] For H2B, only α1 has predicted modest helix propensity, with minor propensity predicted for the central to C-terminal segments of α2.

The effects of the Gly mutations, with removal of side-chain interactions and reduction in helix propensity, are the most straightforward indicators of whether a residue contributes to the stability of I_2. All Gly mutations destabilize the I_2 ensemble by ≥0.8 kcal mol^{-1}, demonstrating that I_2 contains elements of structure in the central α2 helices as well as the H2A-α3 and αC and H2B-α1 helices.

In the H2A monomer, only E61G (center of α2) and E91G (N-terminus of αC) were destabilizing, by 0.4 to 0.3 kcal mol^{-1}.[12] This result is consistent with AGADIR predictions of helix propensity for these residues, although in contrast, E64G is predicted to have similar helicity but does not appear to be folded in the H2A monomer. Glu61 and Glu91 are in relatively close proximity in the native dimer as part of an acidic patch, which includes Glu64. Based on their relative $\Delta\Delta G_{equil}$ and $\Delta\Delta G_{2U-I2}$ values, Glu61

and Glu91 residues are fully folded in I_2, while their neighboring residues, Glu64 and Asn89 (C-cap of α3), respectively, are unfolded in the monomer and develop stabilizing non-native interactions in I_2. Asn68 at the C-terminus of α2 becomes partially structured in I_2; however, increased helix propensity at this position (Ala *versus* Gly) as well as at Glu91 destabilizes I_2, presumably through disrupting non-native interactions.

In contrast to the AGADIR predictions, Lys43 in α1 is not folded in the isolated H2B monomer, while S57G, N64G, and E73G destabilize the monomer.[12] The data in this study show that α1 and α2 of H2B become fully folded in the I_2 ensemble, although the central and C-terminal regions (N64G and E73G) may contribute to non-native interactions. For residue 43, helix propensity does not dictate stability in I_2 or N_2 (Fig. 5c), suggesting that the salt bridge between Lys43 and Asp48 is more important and may be formed in the I_2 species.

In summary, residues that contribute to the stability of the isolated H2A and H2B monomers become fully folded upon association to form I_2. The structure developed in the dimeric intermediate ensemble includes residues that span much of α2 in H2B and the C-terminal half of α2 in H2A. Additionally, residues in α1 of H2B as well as α3 and αC of H2A contribute to the stability of I_2.

Non-native structure in the I_2 ensemble and the transition state leading to N_2

Over the past two decades, there has been an ongoing debate regarding whether kinetic intermediates are productive steps in protein folding or misfolded, kinetic traps, including off-pathway species (for review, see Refs. 25–28). Much of the debate has focused on small (<100 residues), single-domain, monomeric proteins, many of which fold rapidly by two-state mechanisms. In contrast, larger proteins with multiple domains, both monomeric and oligomeric, generally fold via kinetic intermediates and more complicated folding mechanisms. Presumably, this greater complexity reflects the enhanced difficulties associated with attaining more complicated native structures composed of multiple domains and/or subunits.

It is instructive to compare the kinetic folding mechanisms of four histone folds: the eukaryotic H2A–H2B heterodimer[11] and the $(H3–H4)_2$ hetero-tetrameric dimer of dimers[10] as well as the archaeal homodimers hMfB and hPyA1.[29] Despite significantly different stabilities,[30] the eukaryotic heterodimers fold by a similar mechanism, as described in Scheme 1, with an association rate that approaches the diffusion limit. However, despite a conserved dimerization motif and stabilities comparable to H2A–H2B, the homodimeric hMfB and hPyA1 fold by simpler mechanisms. Two-state folding is observed for hPyA1 with an association rate of $9\times10^6\ M^{-1}\ s^{-1}$; hMfB folds up to 8 times faster via a burst-phase monomeric intermediate.[29] Based on these examples of faster association and folding in the presence of kinetic intermediates, it was hypothesized that the histone kinetic intermediates are not traps but serve to accelerate folding in a hierarchical manner.

Coupled to the debate regarding the productive or nonproductive nature of kinetic intermediates, there is a question of the contribution of non-native interactions in ensembles of partially folded species. There are several examples of non-native interactions in transient kinetic folding intermediates. In some instances, such as the incorrect residue ligating to the heme in cytochrome *c*,[31] the non-native interactions are clearly off-pathway and inhibit rapid and efficient folding. However, there is a growing body of literature that indicates that the formation of non-native structure is not an off-pathway folding event and, furthermore, that these structures may contribute favorably to rapid protein folding. The hidden folding intermediates of a four-helix bundle, redesigned apocytochrome b_{562}, which exist after the rate-limiting step, have been selectively populated at equilibrium by mutagenesis and structurally characterized by NMR.[32–34] Although these intermediates have native topology in three of the four helices, there are several specific non-native hydrophobic interactions that result in repacking of the hydrophobic core to maximize burial of hydrophobic surface area in the absence of folding of the final helix. However, rearrangement of this non-native structure is rapid and facile such that folding is not slowed; neither are transient kinetic intermediates populated.

Examples of non-native structure in populated kinetic intermediates in helical proteins are the E colicin immunity proteins Im7[35] and SIm9 (a variant of the related Im9 protein)[36] and apomyoglobin.[37] In these examples, a well-formed helix docks in a non-native orientation or register with respect to other helices in the intermediate ensemble, which leads to an enhanced burial of hydrophobic surface area. For the immunity proteins, the general trend seems to be that mutations that destabilize the intermediate exhibit faster rates of folding from the intermediate to the native state, and this is particularly apparent for the Im7 mutants that revealed the presence of non-native interactions.[35] This trend suggests that the population of the intermediates and their non-native structure is a kinetic trap to some extent. The A-state equilibrium intermediate of apomyoglobin, populated at pH 4, is very similar to the on-pathway kinetic intermediate. In the mutational study that identified non-native structure in the apomyoglobin intermediate ensembles, the few mutations that significantly destabilized the equilibrium A-state had little effect or slightly decreased the rate of folding from I to N.[37] In contrast to the immunity protein model, it appears that the apomyoglobin intermediate, with its non-native structure, does not impede folding given that destabilization of the intermediate does not accelerate folding. The apomyoglobin results are similar to those observed for H2A–H2B. Every position mutated resulted in destabilization of I_2 (Table 2), and all mutations

(except H2A-E61A) have little effect or a minor decrease in the rate of folding from I_2 to N_2 (Fig. 4b). These findings demonstrate that non-native structure is not necessarily a significant impediment to folding and may favor efficient folding. This conclusion is supported by computational studies showing that non-native structure in intermediate ensembles, including hydrophobic interactions, can enhance folding.[38,39]

Conclusions

The mutations studied in this report indicate that residues in the α2 helices of H2A and H2B, as well as α1 of H2B and the C-terminus of α3 and the short αC of H2A, contribute to the stability of the I_2 burst-phase species. It is likely that other segments of H2A and H2B are involved in stabilizing the I_2 species, but further mutational studies are necessary to identify these residues. It is a significant result that eight of the nine (if not all) sites targeted by mutation stabilize I_2 by interactions that are non-native to at least some extent. Given that destabilizing I_2 and these non-native interactions does not accelerate folding, it is concluded that the native and non-native structure present in the I_2 ensemble enables efficient folding of the H2A–H2B heterodimer. It is speculated that I_2 stability is achieved by formation of an ensemble of partially folded dimeric structures that maximize burial of hydrophobic surface area and perhaps formation of favorable non-native electrostatic interactions between Glu residues in the acidic patch and the many cationic residues in this highly basic dimer. Furthermore, the interactions that stabilize the I_2 ensemble, including those that are non-native, favor efficient folding by narrowing the manifold of populated conformations to a set that is poised for folding to the native state, with facile rearrangement of non-native interactions to those observed in the native dimer. These speculations can and will be tested by additional mutational studies.

Materials and Methods

Materials

Ultra-pure urea was purchased from ICN Biomedicals (Costa Mesa, CA). All other chemicals were of molecular biology or reagent grade from JT Baker (Phillipsburg, NJ). The construction of the plasmids for expression of the mutant H2A and H2B histones is described elsewhere.[12] The WT and mutant histone monomers were over-expressed as inclusion bodies, purified, and reconstituted into native heterodimers as described previously.[17]

Methods

All equilibrium and kinetic experiments were performed at 25 °C in a standard buffer of 200 mM KCl, 20 mM potassium phosphate (KPi) (pH 7.2), and 0.1 mM ethylenediaminetetraacetic acid (EDTA). The instrumentation, data collection, and analyses for the equilibrium data are described elsewhere.[12] Kinetic folding and unfolding data were collected with an AVIV Instruments SF tower interfaced with an AVIV 202SF CD spectrophotometer. The dead time of the SF experiments was ~5 ms at a 2-ml/s flow rate. SF-CD kinetics were monitored at 222 nm, and 25 kinetic transients were averaged to enhance the signal-to-noise ratio of each kinetic trace. For intrinsic Tyr FL, excitation was at 280 nm, and emission was detected at 90° to the incoming excitation beam, using a 295-nm cutoff filter; for each SF-FL trace, 20 kinetic transients were averaged. At each urea concentration, a data set of four to six kinetic traces (a combination of SF-CD and SF-FL traces of averaged transients) was analyzed.

Unfolding reactions were initiated by SF dilution of folded heterodimer into various final urea concentrations, generally from 2.4 to 3.6 M. The substrates for the folding reactions were the isolated H2A and H2B monomers pre-equilibrated at various urea concentrations from 0 to 1.6 M. Folding was initiated by SF mixing of these partially folded monomers to final urea concentrations equal to the pre-equilibration conditions; that is, $[\text{Urea}]_{\text{initial}} = [\text{Urea}]_{\text{final}}$.

As described previously for WT H2A–H2B,[11] individual folding and unfolding kinetic traces were fit to a single, first-order exponential function:

$$Y(t) = Y_\infty + \Delta Y_i \times \exp(-k_{\text{obs}} t) \tag{3}$$

where Y_∞ is the final equilibrium signal, ΔY_i is the signal change associated with the kinetic phase, and k_{obs} is the observed rate for each kinetic trace. The rates determined from SF-CD and SF-FL were in excellent agreement. Therefore, using Savuka 5.1,[40,41] we semi-globally fit all SF-CD and SF-FL traces at a given final urea concentration with rates linked across all kinetic traces; Y_∞ and ΔY_i were treated as local parameters in all global fits. The resulting fitted rates are represented by the symbols in Figs. 2 and 3 and Supplementary Figs. S1 and S2. Across the concentration regimes where the semi-globally fitted rates exhibited a log-linear dependence on [Urea], all kinetic traces were globally fit to the following equation:

$$k_{\text{obs}} = k(\text{H}_2\text{O}) \times \exp\left(\frac{-m^{\ddagger}[\text{Urea}]}{RT}\right) \tag{4}$$

where the globally fitted parameters $k(\text{H}_2\text{O})$ and $m^{\ddagger}$ are, respectively, the folding or unfolding rate constants in the absence of urea and the dependence of the rates on the final urea concentration.

The folding–unfolding double-jump experiments at 0 M urea employed the Aviv Instruments SF tower interfaced with an AVIV ATF-105 fluorometer. The FL excitation and emission wavelengths were 280 and 308 nm, respectively. Two double mixers with 11 and 198 μl delay lines between the first and second mixers were used, as described previously.[11,15] The shorter mixer allowed folding delays less than 12 ms, while delays of 50 ms to 5 s were achieved by aging the folding protein in the longer delay line. Because of constraints of the three syringe configuration, folding was initiated by diluting pre-mixed H2A and H2B monomers unfolded in 5 mM HCl into the standard folding buffer at 0 M urea. As a control, first-order folding rates were determined by direct SF jumps from 5 mM HCl to urea concentrations between 0 and 1.0 M, and the results were virtually identical for refolding from acid unfolded monomers (data not shown) and isolated, partially folded monomers at pH 7.2 (Fig. 2). After the various refolding delays (22 time points spanning 57 ms to 4.1 s), unfolding

was induced by addition of urea in the second mixer to a final concentration of 3.4 M. The unfolding kinetic responses were fit globally to a single exponential with the rate linked across all delay times. The unfolding rates from the double-jump experiments were in good agreement with the rates determined from direct unfolding SF kinetic methods. The population of N_2 present after a given folding delay time was estimated from the observed unfolding amplitudes.

Acknowledgements

This work was supported by a National Institute of General Medical Sciences grant (GM073787) to L.M.G. M.R.S. was partially supported by a National Institutes of Health Biotechnology training grant (GM08336). We thank Traci Topping for her technical assistance with the double-jump experiments and critical reading of the manuscript.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2010.06.034

References

1. Luger, K. (2006). Dynamic nucleosomes. *Chromosome Res.* **14**, 5–16.
2. Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, **447**, 407–412.
3. Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, **128**, 693–705.
4. Kamakaka, R. T. & Biggins, S. (2005). Histone variants: deviants? *Genes Dev.* **19**, 295–310.
5. Smith, C. L. & Peterson, C. L. (2005). ATP-dependent chromatin remodeling. *Curr. Top. Dev. Biol.* **65**, 115–148.
6. Park, Y. J. & Luger, K. (2008). Histone chaperones in nucleosome eviction and histone exchange. *Curr. Opin. Struct. Biol.* **18**, 282–289.
7. English, C. M., Adkins, M. W., Carson, J. J., Churchill, M. E. & Tyler, J. K. (2006). Structural basis for the histone chaperone activity of Asf1. *Cell*, **127**, 495–508.
8. English, C. M., Maluf, N. K., Tripet, B., Churchill, M. E. & Tyler, J. K. (2005). ASF1 binds to a heterodimer of histones H3 and H4: a two-step mechanism for the assembly of the H3–H4 heterotetramer on DNA. *Biochemistry*, **44**, 13673–13682.
9. Jenuwein, T. & Allis, C. D. (2001). Translating the histone code. *Science*, **293**, 1074–1080.
10. Banks, D. D. & Gloss, L. M. (2004). Folding mechanism of the (H3–H4)2 histone tetramer of the core nucleosome. *Protein Sci.* **13**, 1304–1316.
11. Placek, B. J. & Gloss, L. M. (2005). Three-state kinetic folding mechanism of the H2A/H2B histone heterodimer: the N-terminal tails affect the transition state between a dimeric intermediate and the native dimer. *J. Mol. Biol.* **345**, 827–836.
12. Stump, M. R. & Gloss, L. M. (2008). Mutational analysis of the stability of the H2A and H2B histone monomers. *J. Mol. Biol.* **384**, 1369–1383.
13. Gloss, L. M. & Matthews, C. R. (1998). The barriers in the biomolecular and unimolecular folding reaction of the dimeric core domain of *Escherichia coli* Trp repressor are dominated by enthalpic contributions. *Biochemistry*, **37**, 16000–16010.
14. Gloss, L. M. & Matthews, C. R. (1998). Mechanism of folding of the dimeric core domain of *Escherichia coli* Trp repressor: a nearly diffusion-limited reaction leads to the formation of an on-pathway dimeric intermediate. *Biochemistry*, **37**, 15990–15999.
15. Topping, T. B., Hoch, D. A. & Gloss, L. M. (2004). Folding mechanism of FIS, the intertwined, dimeric factor for inversion stimulation. *J. Mol. Biol.* **335**, 1065–1081.
16. Otzen, D. E. & Oliveberg, M. E. (1999). Salt-induced detour through compact regions of the protein folding landscape. *Proc. Natl Acad. Sci. USA*, **96**, 11746–11751.
17. Gloss, L. M. & Placek, B. J. (2002). The effect of salts on the stability of the H2A–H2B histone dimer. *Biochemistry*, **41**, 14951–14959.
18. Sanchez, I. E. & Kiefhaber, T. (2003). Origin of unusual phi-values in protein folding: evidence against specific nucleation sites. *J. Mol. Biol.* **334**, 1077–1085.
19. Fersht, A. R. & Sato, S. (2004). Phi-value analysis and the nature of protein-folding transition states. *Proc. Natl Acad. Sci. USA*, **101**, 7976–7981.
20. de los Rios, M. A., Muralidhara, B. K., Wildes, D., Sosnick, T. R., Marqusee, S., Wittung-Stafshede, P. *et al.* (2006). On the precision of experimentally determined protein folding rates and phi-values. *Protein Sci.* **15**, 553–563.
21. Pace, C. N. & Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* **75**, 422–427.
22. Oliveberg, M. (2001). Characterisation of the transition states for protein folding: towards a new level of mechanistic detail in protein engineering analysis. *Curr. Opin. Struct. Biol.* **11**, 94–100.
23. Fowler, S. B. & Clarke, J. (2001). Mapping the folding pathway of an immunoglobulin domain: structural detail from Phi value analysis and movement of the transition state. *Structure*, **9**, 355–366.
24. Munoz, V. & Serrano, L. (1997). Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: comparison with Zimm–Bragg and Lifson–Roig formalisms. *Biopolymers*, **41**, 495–509.
25. Jackson, S. E. (1998). How do small single-domain proteins fold? *Fold Des.* **3**, R81–91.
26. Baldwin, R. L. & Rose, G. D. (1999). Is protein folding hierarchic? I. Local structure and peptide folding. *Trends Biochem. Sci.* **24**, 26–33.
27. Krishna, M. M. & Englander, S. W. (2007). A unified mechanism for protein folding: predetermined pathways with optional errors. *Protein Sci.* **16**, 449–464.
28. Brockwell, D. J. & Radford, S. E. (2007). Intermediates: ubiquitous species on folding energy landscapes? *Curr. Opin. Struct. Biol.* **17**, 30–37.
29. Topping, T. B. & Gloss, L. M. (2004). Stability and folding mechanism of mesophilic, thermophilic and hyperthermophilic archaeal histones: the importance of folding intermediates. *J. Mol. Biol.* **342**, 247–260.
30. Banks, D. D. & Gloss, L. M. (2003). Equilibrium folding of the core histones: the H3–H4 tetramer is less

stable than the H2A–H2B dimer. *Biochemistry*, **42**, 6827–6839.
31. Sosnick, T. R., Mayne, L. & Englander, S. W. (1996). Molecular collapse: the rate-limiting step in two-state cytochrome *c* folding. *Proteins*, **24**, 413–426.
32. Feng, H., Takei, J., Lipsitz, R., Tjandra, N. & Bai, Y. (2003). Specific non-native hydrophobic interactions in a hidden folding intermediate: implications for protein folding. *Biochemistry*, **42**, 12461–12465.
33. Feng, H., Vu, N. D. & Bai, Y. (2004). Detection and structure determination of an equilibrium unfolding intermediate of Rd-apocytochrome *b*562: native fold with non-native hydrophobic interactions. *J. Mol. Biol.* **343**, 1477–1485.
34. Feng, H., Zhou, Z. & Bai, Y. (2005). A protein folding pathway with multiple folding intermediates at atomic resolution. *Proc. Natl Acad. Sci. USA*, **102**, 5026–5031.
35. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nat. Struct. Biol.* **9**, 209–216.
36. Morton, V. L., Friel, C. T., Allen, L. R., Paci, E. & Radford, S. E. (2007). The effect of increasing the stability of non-native interactions on the folding landscape of the bacterial immunity protein Im9. *J. Mol. Biol.* **371**, 554–568.
37. Nishimura, C., Dyson, H. J. & Wright, P. E. (2006). Identification of native and non-native structure in kinetic folding intermediates of apomyoglobin. *J. Mol. Biol.* **355**, 139–156.
38. Clementi, C. & Plotkin, S. S. (2004). The effects of nonnative interactions on protein folding rates: theory and simulation. *Protein Sci.* **13**, 1750–1766.
39. Zarrine-Afsar, A., Wallin, S., Neculai, A. M., Neudecker, P., Howell, P. L., Davidson, A. R. & Chan, H. S. (2008). Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding. *Proc. Natl Acad. Sci. USA*, **105**, 9999–10004.
40. Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E. & Matthews, C. R. (1995). Probing the folding mechanism of a leucine zipper peptide by stopped-flow circular dichroism spectroscopy. *Biochemistry*, **34**, 12812–12819.
41. Bilsel, O., Zitzewitz, J. A., Bowers, K. E. & Matthews, C. R. (1999). Folding mechanism of the alpha-subunit of tryptophan synthase, an alpha/beta barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels. *Biochemistry*, **38**, 1018–1029.
42. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–1113.