

Protein Science

Proline in alpha-helical kink is required for folding kinetics but not for kinked structure, function, or stability of heat shock transcription factor

JA Hardy and HCM Nelson

Protein Sci. 2000 9: 2128-2141

Access the most recent version at doi:[10.1110/ps.9.11.2128](https://doi.org/10.1110/ps.9.11.2128)

References

Article cited in:

<http://www.proteinscience.org/cgi/content/abstract/9/11/2128#otherarticles>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

Notes

To subscribe to *Protein Science* go to:
<http://www.proteinscience.org/subscriptions/>

Proline in α -helical kink is required for folding kinetics but not for kinked structure, function, or stability of heat shock transcription factor

JEANNE A. HARDY AND HILLARY C.M. NELSON

Johnson Research Foundation and Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104

(RECEIVED July 18, 2000; FINAL REVISION September 1, 2000; ACCEPTED September 8, 2000)

Abstract

The DNA-binding domain of the yeast heat shock transcription factor (HSF) contains a strictly conserved proline that is at the center of a kink. To define the role of this conserved proline-centered kink, we replaced the proline with a number of other residues. These substitutions did not diminish the ability of the full-length protein to support growth of yeast or to activate transcription, suggesting that the proline at the center of the kink is not conserved for function. The stability of the isolated mutant DNA-binding domains was unaltered from the wild-type, so the proline is not conserved to maintain the stability of the protein. The crystal structures of two of the mutant DNA-binding domains revealed that the helices in the mutant proteins were still kinked after substitution of the proline, suggesting that the proline does not cause the α -helical kink. So why are prolines conserved in this and the majority of other kinked α -helices if not for structure, function, or stability? The mutant DNA-binding domains are less soluble than wild-type when overexpressed. In addition, the folding kinetics, as measured by stopped-flow fluorescence, is faster for the mutant proteins. These two results support the premise that the presence of the proline is critical for the folding pathway of HSF's DNA-binding domain. The finding may also be more general and explain why kinked helices maintain their prolines.

Keywords: α -helical kink; crystallography; folding; heat shock transcription factor

The heat shock transcription factor (HSF) is the principal transcriptional regulator of the heat shock response and is present in all eukaryotes. HSF binds to upstream activating regions called Heat Shock Elements (HSEs). HSEs are composed of multiple inverted repeats of the sequence "nGAAn." These HSEs are found in the promoters of genes involved in the heat shock response. In yeast, HSF is constitutively bound at some heat shock promoters, and upon sensing heat shock or other stresses, HSF directs transcription (Jakobsen & Pelham, 1988; Gross et al., 1990; Liu et al., 1997). The conserved core of HSF consists of a DNA-binding domain (DBD) and a trimerization domain (Wu, 1995). The DBD contains a "winged-helix-turn-helix" motif (Fig. 1), and the trimerization domain contains a three-stranded coiled-coil (Harrison et al., 1994; Peteranderl et al., 1999). This conserved core is flanked by activation domains and regulatory regions (Sorger, 1990,

Jakobsen & Pelham, 1991; Bonner et al., 1992; Chen et al., 1993; Hoj & Jakobsen, 1994; Tamai et al., 1994; Cho et al., 1996; Morano et al., 1999). The regulation of HSF appears to be complex, since mutations in a number of regions of HSF can cause both increases and decreases in HSF activity (Nieto-Sotelo et al., 1990; Sorger, 1990, 1991; Jakobsen & Pelham, 1991; Silar et al., 1991; Yang et al., 1991; Bonner et al., 1992; Hoj & Jakobsen, 1994; Hubl et al., 1994; Tamai et al., 1994; Halladay & Craig, 1995; Sewell et al., 1995; Morano et al., 1999). Moreover, the DBD has been implicated in this regulation (Bonner et al., 1992; Hubl et al., 1994; Hubl, 1995; Hardy et al., 2000).

One of the most unusual aspects of the DBD is a conserved α -helical bulge and kink located in the second helix of the domain, which is observed in all reported structures of HSF DBD from yeast, fruit-fly, and tomato (Damberger et al., 1994; Harrison et al., 1994; Vuister et al., 1994; Schultheiss et al., 1996). The α -helical bulge is caused by the presence of five amino acids in a single helical turn. Immediately following the bulge is a proline, which is at the center of a kink. Because of the strict conservation of this proline, we found it likely to play either a structural role, a functional role, or both.

Prolines are very often found at the N-capping positions of helices, but are rarely found at internal positions in the helix (Chou & Fasman, 1974; Levitt, 1978; Richardson & Richardson, 1988).

Reprint requests to: Hillary C.M. Nelson, University of Pennsylvania School of Medicine, Department of Biochemistry and Biophysics, 814 Stellar-Chance Building/6059, 422 Curie Blvd., Philadelphia, Pennsylvania 19104-6059; e-mail: hnelson@mail.med.upenn.edu.

Abbreviations: CD, circular dichroism; HSF, heat shock transcription factor; HSE, heat shock element; DBD, DNA-binding domain; GuHCl, guanidine hydrochloride; proline *cis/trans* isomerization, *cis/trans* isomerization of the peptide bond preceding proline in an amino acid sequence.

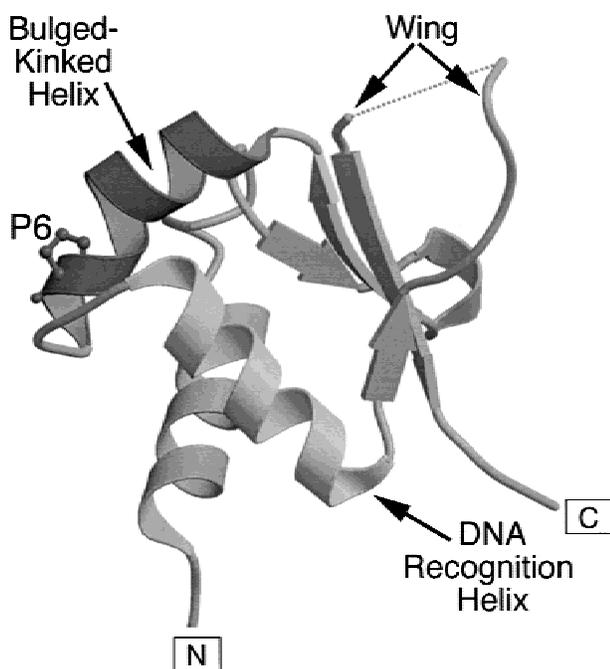


Fig. 1. The crystal structure of the DBD of *K. lactis* HSF (Harrison et al., 1994). The dashed line connects two ends of a loop that was disordered and unobserved in the crystal structure. The residues in the bulged-kinked helix are drawn in dark gray, and the DNA recognition helix is light gray. The proline at the center of the 29° kink is noted as P6.

Because prolines are imino acids rather than amino acids, they are generally considered to be helix breakers due to the lack of the amide proton that normally participates in the hydrogen bonds that stabilize α -helices. Additionally, the cyclic nature of the pyrrolidine side chain of proline restricts the dihedral angles of the preceding amino acid, diminishing flexibility (Barlow & Thornton, 1988). Of the 20 naturally occurring amino acids, proline is the most statistically unlikely of any residue to occupy any of the positions within the body of the helix (Kumar & Bansal, 1998a). Furthermore, the propensity of proline to occur in the middle or C-terminus of a helix is significantly lower than the propensity of any other residue to occur at any position (Williams et al., 1987; Wilmot & Thornton, 1988). When a proline residue is present in a helix, the helix is invariably kinked (Barlow & Thornton, 1988; Kumar & Bansal, 1998a). Despite the absence of the amide proton in the proline-containing helices, all of the hydrogen bond acceptors are satisfied by hydrogen bonds either within the protein or to solvent (Barlow & Thornton, 1988; Woolfson & Williams, 1990; Kumar & Bansal, 1998b). This is also the case in the bulged-kinked helix in HSF, which has all of the hydrogen bonds satisfied, either within the helix or by solvent molecules (Harrison et al., 1994).

Surveys of proteins with kinked helices show that the majority of kinked helices have prolines (Barlow & Thornton, 1988; Kumar & Bansal, 1998b). For example, one study analyzed the structures of all globular proteins that had been solved before 1995 to a resolution of 2.5 Å or better (Kumar & Bansal, 1998b). Of the 1,131 α -helices, 45 (4.0%) were kinked, with an average angle of $33 \pm 11^\circ$. Twenty-five of the 45 kinked helices have a proline at the center of the kink, and these are kinked by $32 \pm 8^\circ$. Proline-

containing helices appear more commonly in transmembrane proteins than in globular proteins and are believed to be essential to the activity of many systems (Konopka et al., 1996; Javitch et al., 1998; Yamaguchi et al., 1999). For both globular and transmembrane proteins, proline-containing helices are always kinked. This correlation has led to the obvious suggestion that prolines might cause the kink or stabilize the kinked conformation. Data presented here and elsewhere suggests that this may not be correct (Alber et al., 1988; Yuan et al., 1994).

To probe the role of the conserved proline in the HSF DBD, we analyzed a series of proline replacement mutants. The proline at position 6 in the bulged-kinked helix was replaced by various residues for genetic, biochemical, and structural studies. The proline substitutions did not lower the activity of HSF, nor did they affect the stability or structure of the domain. However, these substitutions did decrease the solubility of the DBD when over-expressed in *Escherichia coli*, and they also altered the folding kinetics of the isolated domain. We suggest that the role of the conserved proline might be to regulate the folding pathway of the HSF DBD.

Results

Temperature dependent growth of yeast containing mutant HSFs

HSF is an essential gene in yeast (Nieto-Sotelo et al., 1990; Sorger, 1990), so it is possible to measure the effect of mutations in the gene on both viability and growth. Some point mutations in the DBD of HSF are lethal, some render yeast temperature sensitive, and other mutations have no negative effect on growth (Silar et al., 1991; Yang et al., 1991; Bonner et al., 1992; Hubl et al., 1994; Halladay & Craig, 1995; Sewell et al., 1995; Hardy et al., 2000). When conserved residues involved in contacts to DNA are mutated, lethality or temperature sensitivity are usually observed (Hubl et al., 1994; Littlefield & Nelson, 1999). The proline at position 6 within helix 2 (the bulged-kinked helix) is at the center of a kink that is conserved in all known structures of HSF (Damberger et al., 1994, 1995; Harrison et al., 1994; Vuister et al., 1994; Schultheiss et al., 1996; Littlefield & Nelson, 1999). This structural conservation of the proline residue led us to ask whether mutations of this proline would have a similar effect on viability.

Wild-type and mutant versions of full-length *Saccharomyces cerevisiae* HSF, under the control of the wild-type promoter on a low copy CEN/ARS plasmid, were expressed as the only HSF in a tester strain that had been deleted for the wild-type chromosomal copy of HSF. These yeast were then tested for their ability to support growth at a range of temperatures (Fig. 2). The wild-type and the proline substitution mutants P6A and P6D did not support growth at 39°, whereas P6K supported growth at this temperature. Therefore, the substitution of a positively charged lysine at this position was sufficient to render HSF-containing yeast more viable than yeast containing wild-type HSF. Substitution of proline by a neutral residue, alanine, or a negatively charged residue, aspartate, seemed to have no effect on the activity of HSF as it relates to survival of yeast.

Transcriptional activity of proline mutant HSFs

To confirm that the P6K mutation was the only mutation to increase HSF's activity, we compared the transcriptional activity of

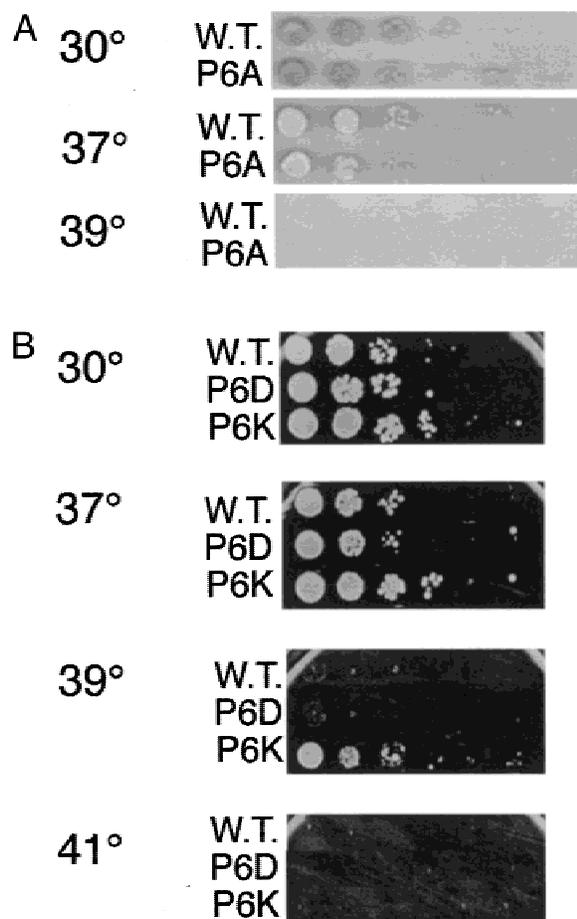


Fig. 2. Growth of yeast strains at various temperatures. **A:** Yeast strains containing plasmid-borne wild-type HSF or P6A as the only source of HSF. **B:** Wild-type, P6D, and P6K. All strains were grown at 30 °C to mid-log phase. These cells were then serially diluted into sterile media (10-fold dilution at each step), and spotted onto agar plates to measure growth at 30, 37, 39, or 41 °C after three days. Plates in **A** were imaged with a white background, accounting for the difference in the color of the image compared to the yeast plates in **B**, which were imaged with a blue background.

the various mutant HSFs. The yeast strains used above were transformed with a reporter construct, pHSE2-lacZ, which has LacZ under the control of an HSE-dependent promoter. The HSE within this promoter is predicted to be fully occupied by HSF under constitutive conditions (Jakobsen & Pelham, 1988, 1991; Gross et al., 1990; Giardina & Lis, 1995). Cells containing mutant or wild-type HSF and the reporter plasmid were grown at 30 °C and either maintained at 30 °C to assay constitutive HSF activity, or heat shocked for 20 min at either 37 or 42 °C to measure inducible activity. Cell extracts were assayed for the amount of expressed β -galactosidase.

Under constitutive conditions (i.e., growth at 30 °C), P6A and P6D had essentially wild-type activity (Fig. 3). The P6K mutant had significantly greater activity than the wild-type DBD or any of the other proline substitution mutants. At heat shock temperatures of 37 or 42 °C, the activity of P6A or P6D were essentially equivalent to wild-type, whereas P6K had greater than wild-type activity. The increase in activity for P6K compared to wild-type was nearly fivefold for constitutive conditions, sixfold at 37°, and fourfold at 42 °C.

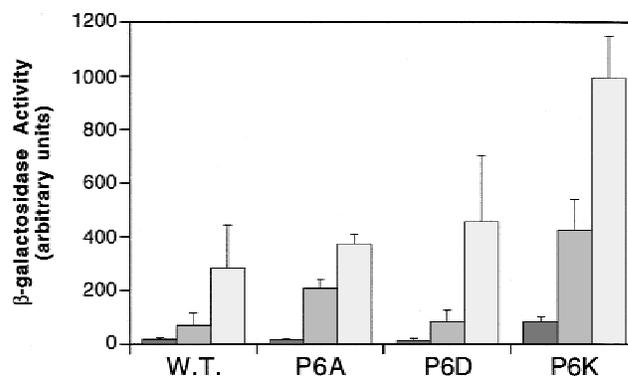


Fig. 3. β -Galactosidase expression from an HSE-lacZ fusion. The constitutive (dark gray), 37 °C heat shock (medium gray), and 42 °C heat shock (light gray) levels of expression are shown.

DNA-binding ability of proline mutant HSFs

To investigate whether the proline substitutions affected DNA binding affinity, truncated versions of *K. lactis* HSF containing the DNA-binding and trimerization domains (DLT) were purified. *K. lactis* HSF functionally substitutes for *S. cerevisiae* HSF in vivo (Jakobsen & Pelham, 1991), and much of our previous biochemical and structural analyses of yeast HSF have used *K. lactis* HSF. As this data correlates well with the genetic data from *S. cerevisiae*, we also chose to use *Kluyveromyces lactis* HSF for the subsequent biochemical and structural analyses (Hubl et al., 1994; Littlefield & Nelson, 1999; Hardy et al., 2000). The DNA binding properties of the proline substitution mutants were determined using the gel mobility shift assay to assess whether their activities correlate with, or can be predicted by, the DNA binding constant (Table 1).

The binding affinities of the proline substitution mutants were altered from the wild-type HSF by no more than threefold. Specifically, the DNA binding of P6A was stronger, although this mutant behaves in the transcriptional and growth assays similarly to wild-type. The binding of P6K was weaker by approximately threefold than the wild-type protein, even though P6K had increased activities in the growth and transcriptional activity assays. This lack of correlation between DNA-binding activities and the transcriptional activities suggested that the differences in DNA-binding activity between the proline substitutions may not be important for HSF function and that the proline substitutions do not dramatically affect DNA-binding affinity. This was not surprising, given that the crystal structure of the HSF DBD in complex with

Table 1. Relative binding constants of wild-type and mutant versions of the DBD on a three-repeat binding site

Protein	Relative K_d^a
Wild-type	1
P6A	3
P6K	0.37

^aDissociation constants were measured by gel-mobility shift assays and were normalized against the wild-type value.

Proline kink

DNA shows that the bulged-kinked helix does not make any contacts to DNA (Littlefield & Nelson, 1999). In addition, HSF is known to be constitutively bound to strong promoters, so only changes in the binding affinity that are significant enough to alter the occupancy at these promoters are likely to affect viability or transcriptional activity. The level of increase or decrease in dissociation constants observed for these mutant DBDs may not account for changes in transcriptional activity or viability, as they are not likely to significantly alter the occupancy of HSF on strong promoters.

Stability of proline mutants

One of the proposed roles of prolines in kinked helices is to stabilize the kink and therefore the protein. To measure whether substitution of the proline in the HSF DBD affected stability, versions of the HSF DBD that contain the mutations P6A, P6K, or P6G were overexpressed and purified. The CD spectra for the various mutant DBDs were similar for the mutant and wild-type proteins and were consistent with the presence of a native, folded protein (data not shown).

The thermal melting temperatures of these mutants were determined by following the loss of CD signal at 222 nm as thermal denaturation was induced by raising the temperature. The T_m for each mutant was assigned as the temperature at which 50% of the protein is unfolded (Fig. 4). None of the substitutions significantly affected the stability of the DBD. This maintenance of the stability of the P6A, P6K, and P6G mutants, taken together with the CD spectra, suggested that the presence of this proline may not be required for achieving the native-state fold or stability, and further prompted the question of why its presence has been conserved.

X-ray crystal structures of the wild-type and proline mutant DBDs

To see the details of the proline substitutions on the DBDs, we determined the structures of both the P6K and P6A mutant versions of the HSF DBD at 2.0 Å resolution, as well as the wild-type

protein at 1.8 Å resolution. All three proteins crystallized in space group P2₁2₁2, with two DBD monomers per asymmetric unit. To determine the structures, we used molecular replacement starting with the previously reported structure of the wild-type DBD (Protein Data Bank (PDB) accession number 1HTS with 1 monomer/asymmetric unit) (Harrison et al., 1994). To avoid bias at position 6 in the kinked helix, the initial electron density maps were calculated from the molecular replacement solutions using a residue other than the actual residue. These initial maps were superimposed with the final models of P6A, P6K, and wild-type to show the unambiguous electron density for the appropriate residue at that position (Fig. 5).

The structures of P6A, P6K, and wild-type HSF DBD in this space group were nearly identical (Fig. 6). The root-mean-square deviations (RMSDs) in backbone coordinates were calculated in CNS and ranged from 0.56 to 0.93 Å (Table 2). The magnitudes of the differences between the two molecules in the asymmetric unit (molecules a and b, respectively) were on par with the deviations observed between the different versions of the protein. No overall structural changes were rendered by the mutations. A superposition of the a molecules of P6K, P6A, and wild-type DBD underscores the fact that P6A, P6K, and wild-type DBDs have nearly identical structures (Fig. 6A).

The N-terminus of the helix and the C-terminal turn also superimposed very closely (Fig. 6B). The residues in the kink (centered around position 6) showed the greatest deviation from the mean structure, not only for this helix, but also for the entire protein. Although these changes in structure of the kink were minor, they may be important, given the tight conformational identity overall. Substitution of the proline at position 6 resulted in a slight shift in the α position in the mutant proteins (Fig. 6C,D). This movement is reflected in the changes in the ϕ - ψ angles upon comparing P6K and P6A to wild-type, which are small in most regions of the protein, on average about 10°. The greatest changes in ϕ - ψ angles are in the immediate vicinity of the kink. The four residues before the proline were 10–30° different than wild-type and one residue following the proline had ϕ - ψ angles that were 110–120° different when comparing each of the structures to one another. These changes in angle are likely to be real differences given that the cross-validated sigma coordinate errors (calculated in CNS; Brünger et al., 1998) for P6K, P6A, and wild-type were 0.16, 0.14, and 0.17 Å, respectively, and the RMSD for angles for P6K, P6A, and wild-type were 1.17, 1.60, and 1.24°, respectively. Although it is formally possible that these subtle structural changes are responsible for the observed changes in activity, we think this situation unlikely.

The mutation of the proline at the center of the kink has not resulted in loss of the kink. In fact, these mutations are close to structurally silent outside the immediate context of the mutation. The motivating question from this work remains: if P6A, P6K, and wild-type are virtually superimposable, and have unperturbed stabilities and undiminished function, then what is the role of the conserved proline in this protein?

Folding kinetics of proline mutant HSFs

The conservation of the proline residue at the center of the helical kink in HSF and other proteins suggests that the prolines in kinked helices may be conserved for a reason. One of the only characteristics of the mutant HSF DBDs that differs from the wild-type DBD is their solubility in *E. coli* after overexpression. The mutant

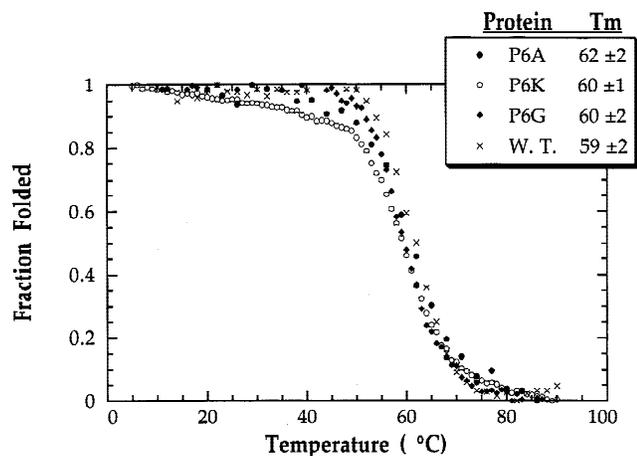


Fig. 4. Thermal denaturation profiles. The thermal melting of wild-type and mutant DBDs was followed by CD at 222 nm. Melting temperature (T_m) is assigned to the temperature at which 50% of the protein is folded and are listed in the box in the upper right-hand corner.

proteins were invariably insoluble upon overexpression, whereas the wild-type protein was fully soluble. Even in experiments designed to purify wild-type DBD from the insoluble portion of the

lysed *E. coli* in which it was overproduced, no wild-type DBD could ever be observed in the insoluble fraction. Aggregation of proteins has been implicated as an indicator of partially folded, unfolded, or misfolded proteins. Misfolded and aggregated proteins have been implicated in several pathological states (Kurada & O'Tousa, 1995; Kuznetsov et al., 1996; Liu et al., 1996; Oksche et al., 1996; Prusiner, 1997; Gow et al., 1998). We hypothesized that the protein-folding pathway for the mutants might be altered relative to wild-type.

A series of stopped flow kinetic refolding assays was performed by rapidly diluting the variants of the DBD denatured in 3 M guanidine to the refolding conditions of 0.6 M guanidine. The refolding was followed by tryptophan fluorescence (Fig. 7). The kinetics of P6K, P6A, and P6G are all clearly increased relative to wild-type. Both P6A and P6K folded with similar kinetics (Fig. 7A,B,E), whereas P6G was different from both wild-type and the other two mutants (Fig. 7D).

Neither the wild-type nor the mutant refolding data could be fitted to a single exponential curve. In all cases, a double exponential expression was required to attain the proper fit of the curve to the data. The replacement of the proline affected both the fast first (k_{f1}) and the slower second folding (k_{f2}) phases for P6A and P6K (Table 3; Fig. 7H). Interestingly, the P6G mutation only dramatically affected the fast folding phase while the slow phase was similar to wild-type (12.5 vs. 13.9 s⁻¹). The requirement for a double exponential equation to adequately fit the folding curves was evidence for the formation of a kinetic intermediate in the folding pathway that accumulates between the first fast and second slower phases of folding (Fig. 7H). It is the accumulation of this type of intermediate that could lead to aggregation.

For each of the proteins, these two phases of the folding reaction occur much too rapidly to be attributable to changes in proline *cis/trans* isomerization, which usually occurs with an uncatalyzed rate on the order of seconds to minutes (Grathwohl & Wüthrich, 1981; Schmid, 1993). Because proline *cis/trans* isomerization is a known determinant in the folding of many proteins, we sought to identify a folding rate controlled by proline *cis/trans* isomerization. A third (much slower) rate constant as would be expected for proline *cis/trans* isomerization could not be identified, as fitting of a three exponent expression resulted in a third rate that was identical to the second rate constant and did not improve the fit. Nevertheless, the slope of the curves at late time points for the wild-type protein were on the order of 0.16, whereas the mutant slopes were always fitted to values of ~0.04 or less. The nonzero slope for the wild-type may actually be masking a third folding phase that makes up very little amplitude of the total signal and occurs on a very slow timescale, as is normally attributed to proline *cis/trans* isomerization (Schmid, 1993). The HSF DBD contains two other prolines (P195, P226) in addition to the one at the

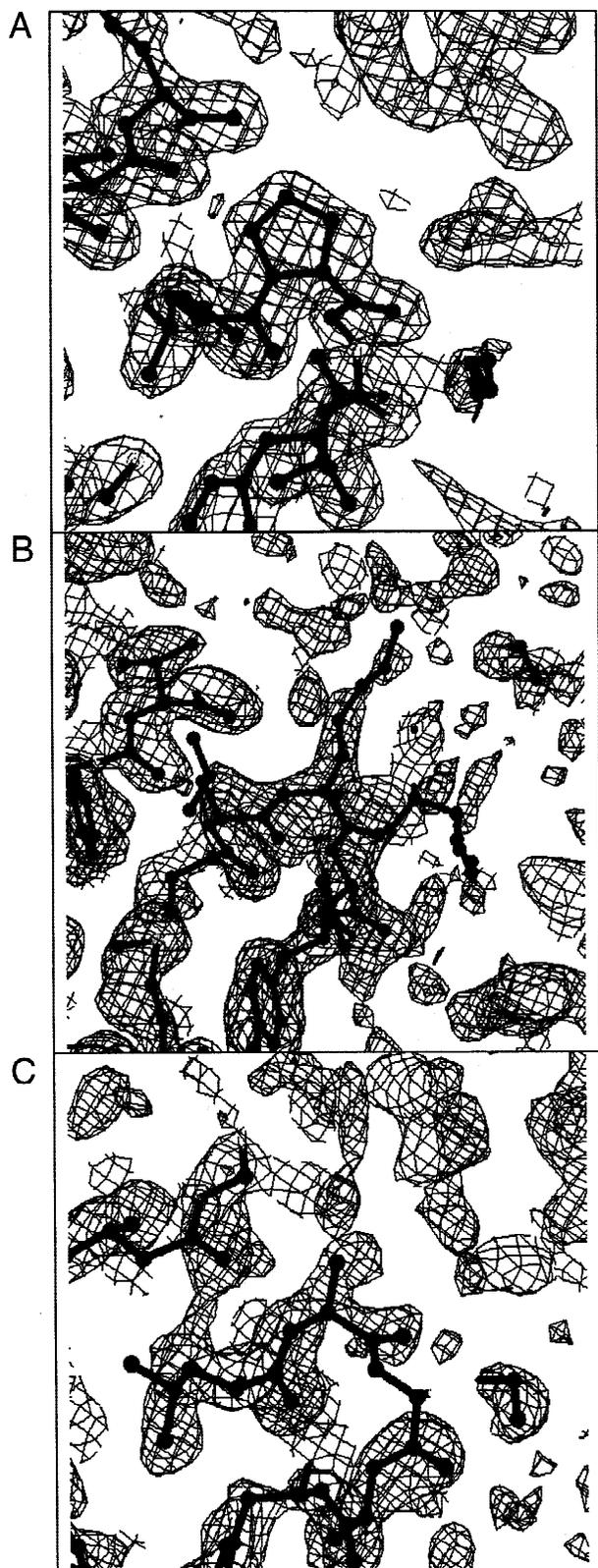


Fig. 5. The initial electron density maps calculated from the molecular replacement solutions are shown above superimposed with the final refined model of the (A) wild-type, (B) P6K, or (C) P6A structures, in ball-and-stick representation. The (A) proline, (B) lysine, or (C) alanine at the center of the image is at position 6 of the DBD. The maps are drawn as a wire mesh contoured at 1.0σ . The map for the P6K DBD was calculated based on the wild-type model with proline truncated to alanine at position 6. The P6A map was calculated with a model containing proline at position 6, and the wild-type map was calculated with the P6K model. This figure was generated using O (Jones et al., 1991) and MOLSCRIPT (Kraulis, 1991).

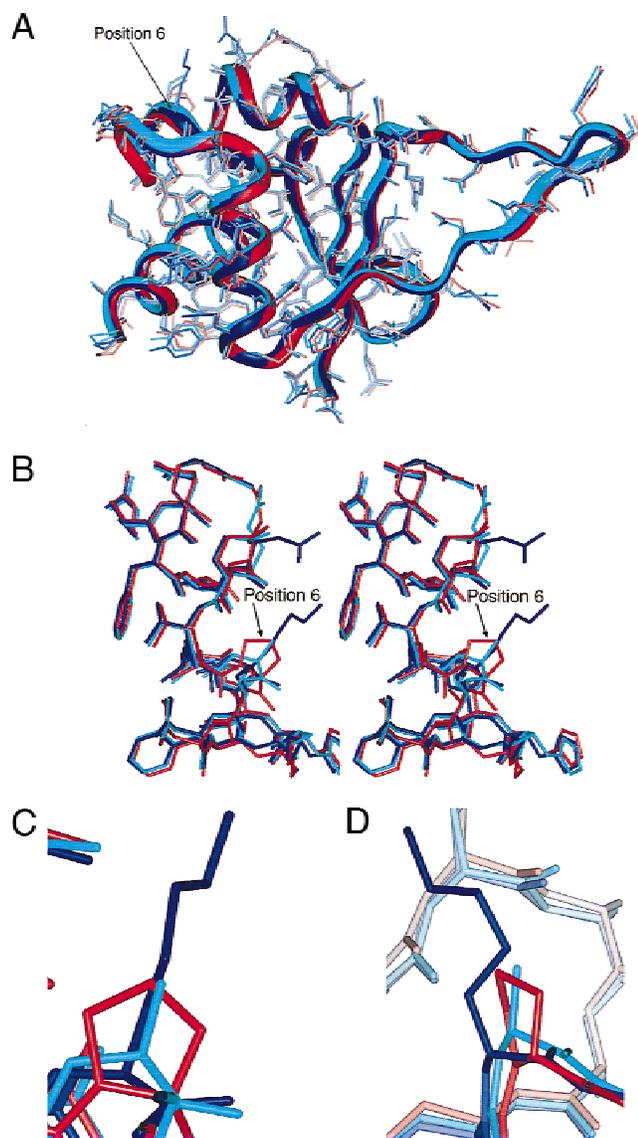


Fig. 6. The superpositions of wild-type, P6K, and P6A DBD structures. The models above show wild-type in red, P6K in blue, and P6A in cyan. All alignments were performed with ALIGN software (Satow et al., 1986; Cohen, 1997), and models were rendered using InsightII version 97.0 (Molecular Simulations Inc., San Diego, California). **A:** Superposition of the wild-type, P6A, and P6K DBDs. **B:** A stereo diagram of the bulged-kinked helix and the turn that follows in the three versions of the HSF DBD. The residue a position 6 is marked to indicate the location of the mutations. **C:** A close-up view of the residues at position 6 in the bulged kinked helix to show the relative orientations of the wild-type proline, or the lysine or alanine that were substituted at that location. **D:** The same view as in C, but rotated by 90°. The residues in the turn are visible behind the position 6 residues in this orientation.

center of the helical kink (P6 is P237); all three of them are in the *trans* isomer in the native state (Harrison et al., 1994). Although P195 is a highly conserved residue, its placement as the third residue from the N-terminus of the domain, outside the core of the protein, makes it unlikely to be important in folding of the DBD. P226 is not a conserved residue and acts as an N-capping residue for the bulged-kinked helix. Since a nearly zero baseline was observed for substitutions at P237 (P6A, P6K, or P6G), *cis/trans*

Table 2. RMSDs^a between P6K, P6A, and wild-type DBDs

	P6K - a	P6K - b	P6A - a	P6A - b	W.T. - a	W.T. - b
P6K - a	—	0.9107	0.7014	0.9581	0.8244	0.8633
P6K - b	—	—	0.7876	0.5581	0.8739	0.6706
P6A - a	—	—	—	0.8424	0.8114	0.9272
P6A - b	—	—	—	—	0.8796	0.7107
W.T. - a	—	—	—	—	—	0.7665
W.T. - b	—	—	—	—	—	—

^aCalculations of RMSD (Å) were performed independently using CNS (Brünger et al., 1998) on molecules a and molecule b from the asymmetric unit of P6K, P6A, and wild-type, respectively.

isomerization of the other two proline residues (P195, P226) is not likely to be involved in the folding pathway, whereas *cis/trans* isomerization of the proline at the center of the helical kink (P237, denoted P6) may be involved in the latest steps of folding.

For proteins that fold with bi-phasic kinetics, the energy barrier that leads to the slower rate constant for the folding reaction should control the reverse (unfolding) reaction (Fig. 7H). Unfolding rates of wild-type, P6A, P6K, and P6G were measured by rapid dilution of the DBD variants from no denaturant to high guanidine concentration (3.8 M), and the reactions were followed by tryptophan fluorescence (Table 3; Fig. 7F,G). All unfolding curves could be unambiguously fitted to a single exponential expression. The unfolding rates of P6K and P6A were similar (0.9 and 1.0 s⁻¹), as were the unfolding rates for P6G and wild-type (4.9 and 4.2 s⁻¹). The unfolding rates for all of the variants were correlated linearly with the slow rate constant for the folding reaction ($R = 0.968$, significant with >95% confidence interval). This correlation was an indication that the unfolding reaction is indeed controlled by the same energy barrier that leads to the slow folding phase. It also independently suggested that the rate constants attained from the fitting to the folding curves are accurate.

Discussion

Role of HSF's helical proline

Many aspects of the proline-substituted mutants in the HSF DBD have been characterized. These mutants have equivalent or increased activity, so this proline has not been conserved for function. Additionally, the stability of the proline-substituted mutants is unaltered from wild-type, so the proline has not been conserved to maintain stability. The crystal structure of the DBD is practically invariant from wild-type in the presence of either the P6A or P6K mutations; therefore, the proline has not been conserved to maintain the native-state kinked structure.

If not for structure, function, or stability, why has proline been conserved, particularly if other residues are equivalent? P6A, P6K, and P6G vary from wild-type in two ways: their solubility upon overexpression and their folding kinetics. Changes in the kinetic folding pathway of various proteins have resulted in their observed insolubility (Ramm et al., 1999). When overexpressed, the proline-substituted DBDs are invariably insoluble. Substitution of alanine, lysine, or glycine for proline alters the protein folding pathway. This change in the folding pathway does not appear to be due to loss of the constraints imposed by proline *cis/trans* isomerization,

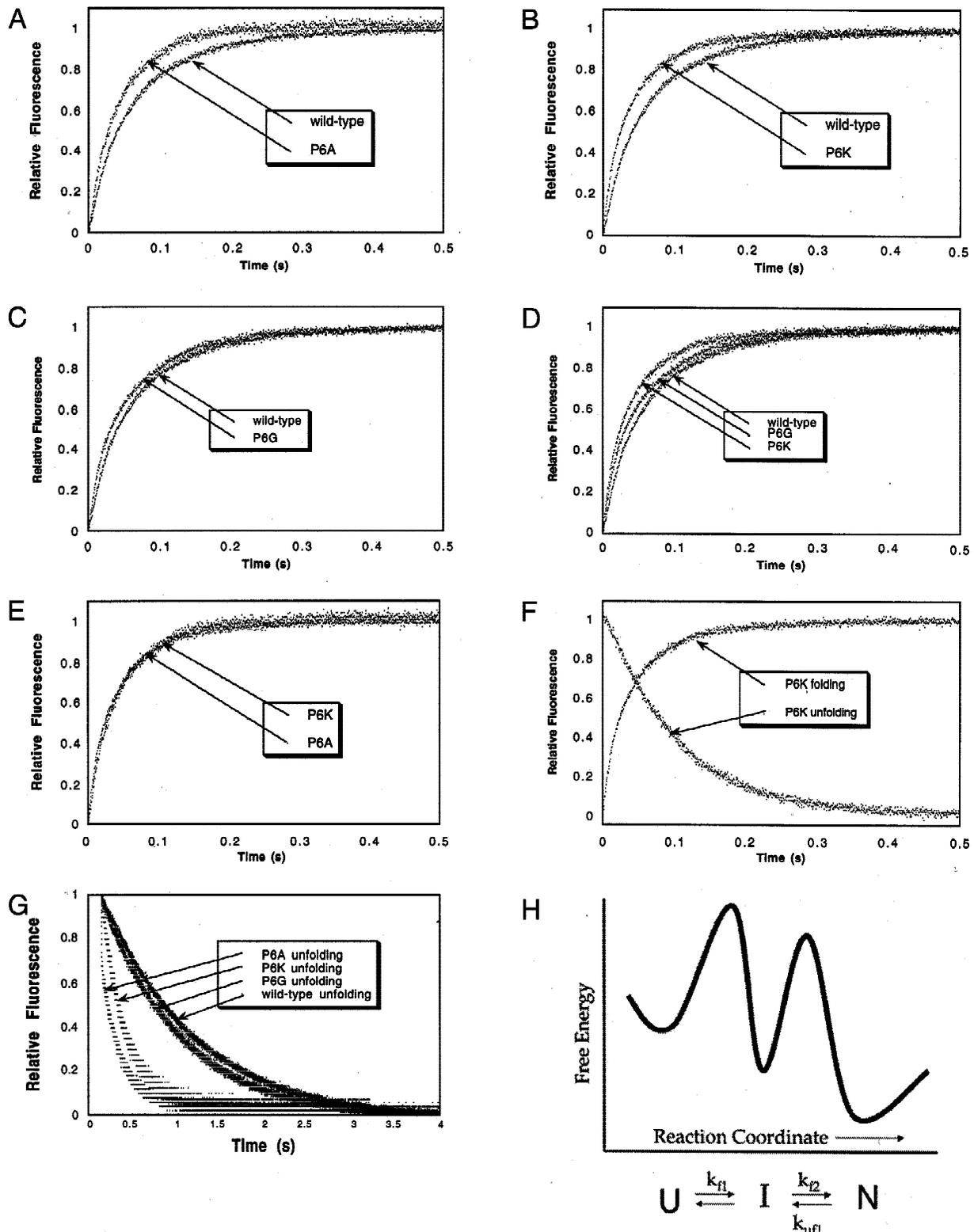


Fig. 7. Folding and unfolding kinetics measured by stopped-flow fluorescence. The refolding of wild-type, P6A, P6K, and P6G were measured by rapid dilution of DBD from 3M GuHCl to 0.6 M GuHCl. The unfolding kinetics were measured by rapid dilution from 0 to 3.8 M GuHCl. Data were normalized to a relative fluorescence scale from 0 to 1 to allow comparison of the data from various mutants. For ease of comparison of folding of (A) wild-type to P6A, (B) wild-type to P6K, (C) wild-type to P6G, (D) wild-type to P6G, and P6K, (E) P6K to P6A are shown. (F) The folding and the unfolding kinetics of P6K. (G) The unfolding kinetics of all three mutants, P6A, P6K, P6G, and wild-type. (H) A model reaction coordinate for the folding and unfolding reactions of the HSF DBD. The energy barriers controlling the formation of the native state (N), intermediate (I), and unfolded (U) states are diagrammed as a function of the free energy of the reaction. The rate constants k_{f1} and k_{f2} are for the first, fast and second, slower folding rates. The unfolding reaction is controlled by the same energy barrier as the second step of the folding reaction and exhibits the unfolding rate constant k_{uf1} .

Table 3. Rate constants for DBD refolding^a

Protein	k_{f1} (s ⁻¹)	$t_{1/2}$ folding (ms)	k_{f2} (s ⁻¹)	$t_{1/2}$ folding (ms)	k_{f1}/k_{f2}	k_{uf1} (s ⁻¹)	$t_{1/2}$ unfolding (s)
Wild-type	42.3 ± 1.5	24	12.5 ± 0.6	80	3.4	0.9 ± 0.1	1.1
P6A	80.8 ± 2.4	13	17.7 ± 1.1	56	4.6	4.9 ± 1.3	0.2
P6K	78.4 ± 5.1	13	18.0 ± 0.9	56	4.4	4.2 ± 0.8	0.2
P6G	83.2 ± 8.4	12	13.9 ± 0.5	72	6.0	1.1 ± 0.3	0.9

^aRate constants and half-times for the folding reaction ($t_{1/2}$) shown above are the average from the data fitted from each of five stopped flow refolding experiments or four stopped flow unfolding experiments.

as the increases in rate constants were present in the fastest phase of folding, which, even for wild-type, is much faster than the rate of proline *cis/trans* isomerization.

The folding rate constants suggest the formation of an intermediate in the folding pathway for both the wild-type and mutants (Fig. 7H). The first, fast phase of folding occurs more rapidly for P6A, P6K, or P6G than it does for wild-type, so the intermediate accumulates to a greater extent for the mutant proteins than for wild-type. The native proline attenuates the fast rate of folding, and its presence prevents accumulation of a kinetic intermediate that is apparently prone to aggregation. Is folding a characteristic that could be selected evolutionarily? Partially folded intermediates are known to be prone to aggregate. Any protein prone to aggregation during expression is equivalent to a hypo-active mutant in cost to the cell, since both require increased expenditure of cellular resources to achieve the same level of HSF activity. Therefore, the kinetic folding pathway may indeed be an evolutionarily selectable trait.

If the proline is conserved to prevent formation of a folding intermediate and its subsequent aggregation, it is interesting that a relatively slight increase in the ratio of the two rate constants can account for such a dramatic change in the ability of these proteins to aggregate. The wild-type k_{f1} rate constant is 3.4 times that of the k_{f2} constant. The ratio of k_{f1} to k_{f2} for the mutants increases to only 4.6, 4.4, or 6.0, but appears to be sufficient to change the solubility properties of these proteins. Approximately 44% of kinked helices do not have prolines. These findings predict that proteins containing nonproline-centered kinks do not need to avoid populating a kinetic intermediate like that formed in the P6K, P6A, and P6G mutants.

The mechanism by which proline substitution leads to faster folding is not clear. One possibility for the faster folding rates is that substituting proline with any other amino acid, all of which have increased allowable torsion angles, expands the productive conformations that are populated during folding. If this is the case, then replacing proline in other elements of secondary structure might likewise increase the folding kinetics in a fast phase (i.e., one not controlled by proline *cis/trans* isomerization).

There are alternative explanations that also provide a rationale for the conservation of the proline in the native state. For example, the structural basis for the kink could be overdetermined, with other tertiary interactions in addition to the presence of the proline providing the necessary stability. If the proline is involved in the folding kinetics, it might be required at early stages of folding to maintain the kink in the absence of these other tertiary interactions. Or, it might be selected for its ability to disfavor alternative struc-

tures that might form under adverse conditions, such as aggregated beta-like structures. These alternatives could be differentiated by a more extensive study of the folding pathway of HSF's DNA-binding domain.

Basis of P6K's increased activity

The mutants P6A, P6K, and P6G have stabilities and structures that are essentially unperturbed from wild-type. The activity of P6G was not characterized, but the activity of P6A is indistinguishable from wild-type in both growth and β -galactosidase assays. In the series of proline-substituted mutants, only one mutant, P6K, has an altered activity either in vitro (β -galactosidase assay) or in vivo (growth at 39°C). Substitution of proline by alanine does very little to change the molecular surface of that face of the protein (Fig. 8). Proline and alanine are both uncharged amino acids, so the surface potentials of P6A and wild-type are virtually identical. The substitution of proline by lysine, on the other hand, incites a much more dramatic effect, and the molecular surface is also significantly altered. It may be a combination of its length (lysine is one of the longest amino acids) and its charge that is responsible for its increase in activity.

Introduction of a negative charge does not have the same effect, as the P6D mutant has wild-type temperature-dependent viability and transcriptional activity. This suggests that insertion of a charge *per se* is not sufficient to increase activity. Given that the positively charged mutant has increased activity, whereas the negatively charged mutant has unaltered activity, it may be reasonable to conclude that this difference is mediated specifically by the positively charged residue. The structure of the HSF DBD in complex with DNA shows that the kinked helix is exposed on the opposite face of the complex from the DNA, in a position available for interaction with an ancillary factor or a regulatory region of HSF (Littlefield & Nelson, 1999). A novel positive charge could physically disrupt or form a salt bridge, or disrupt a hydrophobic interaction. Thus, this heightened activity of P6K might occur via either of two mechanisms: by disrupting an interaction necessary for negative regulation or by enhancing an interaction necessary for activation, making activation more probable and frequent.

Comparisons of proline substitution mutants to bulge mutants in HSF's DBD

We have previously reported the effects of deletion of residues from within the bulge that is adjacent to the proline-centered kink, and which is also strictly conserved amongst all HSFs. These

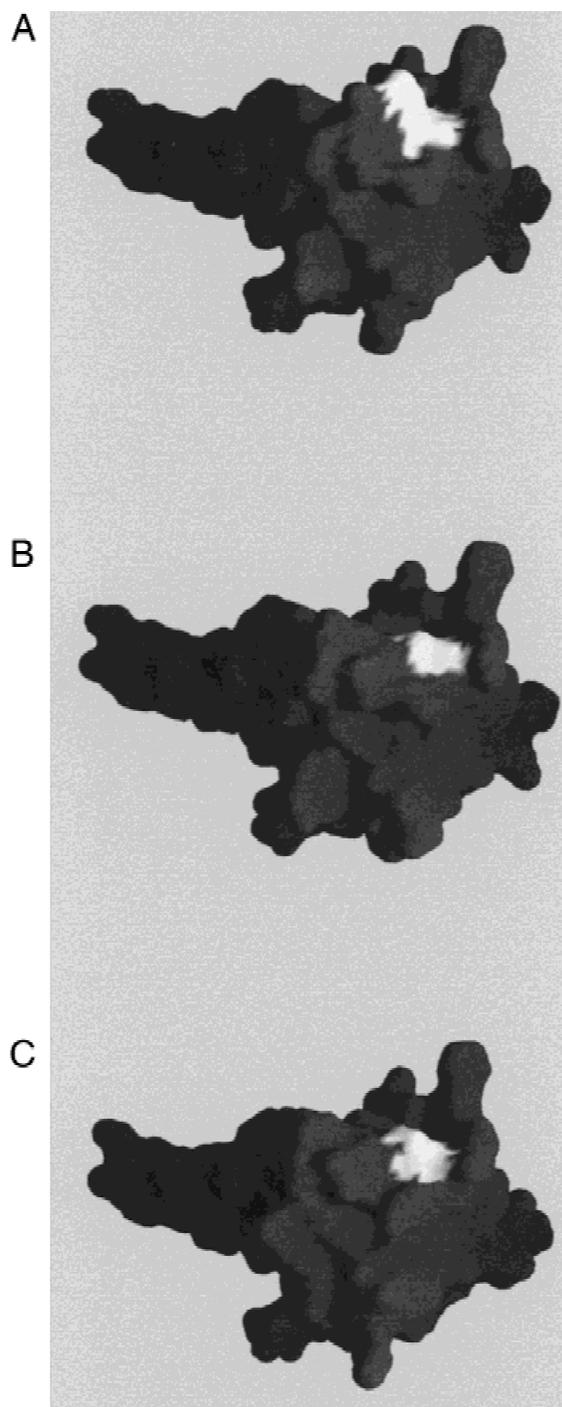


Fig. 8. Molecular surface representations of P6K, P6A, and wild-type. The molecular surface highlighting the change in the surface shape of (A) P6K, (B) P6A, and (C) wild-type molecules are drawn in gray with the residue at position 6 colored white. Surface renderings were made using GRASP (Nicholls et al., 1993). All molecules are shown in the same orientation. The large protrusion on the left side of the protein is the wing, and the bulged-kinked helix runs horizontally along the front of the surface.

deletions result in the thermal destabilization of the mutant proteins. Nevertheless, the deletion mutants are more active than wild-type in both ability to grow at elevated temperatures and in transcriptional activity. We had suggested that the DBD may sense

heat shock directly: the deletion mutants, due to their diminished thermal stability, may be induced to a transcriptionally activated state at a lower temperature (Hardy et al., 2000).

The mutant P6K exhibits a growth phenotype and an *in vivo* activity that is analogous to the increase in activity that we observed for the deletion of residues within the bulge, suggesting that, like the deletion mutant proteins, P6K may have altered thermal stability. This is not the case. Unlike the bulge mutants, the stability of P6K is virtually indistinguishable from wild-type. It is notable that the same phenotype (increased growth and transcriptional activity) can possibly be achieved by two distinct mechanisms. Whereas negative regulatory events that keep HSF in a repressed state in the absence of heat may be prematurely terminated by the bulge mutants in the presence of heat, these or similar interactions may be disabled though the introduction of a positively charged lysine on the surface of the protein.

Substitution of other helical prolines yielded similar effects on structure, function, and stability

Two studies in addition to ours have addressed the role of prolines in helical kinks. In T4 lysozyme, a proline within a 3_{10} helix was replaced by 10 other amino acids (Alber et al., 1988). The structures of the 10 mutants at position 86 shows that all of the substituted residues maintain the kink, even though they also extend the helix by one turn in the N-terminal direction. In addition, all of substitution mutants have thermal stabilities within several degrees of wild-type, and most also have only mildly diminished activity (Alber et al., 1988). Therefore, the proline is not necessary for the structure or stability of kinked conformation.

Fis, an *E. coli* factor that stimulates inversion by the Hin family of recombinases, also contains an α -helical proline that is at the center of a 20° kink. Substitution of proline 61 by alanine maintains a 16° kink in the helix and increases the thermal stability of the protein by 15°C (Yuan et al., 1994). Substitution of proline by serine, leucine, or alanine all result in increased thermal stability. This suggests that this proline was not responsible for forming the helical kink, but may be present to modulate stability to the appropriate level.

The major structural implication of our study of the HSF DBD concurs with these previous studies. Any of the residues that replaced proline were capable of maintaining the kinked conformation, suggesting that proline residues do not cause helical kinks. In addition, any of the residues that replaced proline also maintained, or even increased, the thermal stability of the protein, suggesting that proline residues are not required to stabilize kinked helices. To date, the folding kinetics of the Fis and T4 lysozyme proline-substitution mutants have not been reported, but it would be fascinating to know whether these substitutions alter the folding kinetics similarly to the HSF mutants.

Statistically, prolines are the least likely of any residue to occur in an α -helix, and their propensities to occur in the middle or C-terminus of a helix are significantly lower than the propensity of any other residue to occur at any position (Williams et al., 1987; Wilmot & Thornton, 1988; Kumar & Bansal, 1998a). Nevertheless, when prolines are observed in helices they are strongly correlated with the presence of a kink, as the majority of kinked helices (56%) are centered on a proline residue. If any residue can accomplish the kinked conformation, with native or greater stability, with no negative effect on function, then all residues should have equal representation at the centers of kinked helices. The preponderance

of prolines observed in kinks also lends credence to our suggestion that the prolines have been selected for a reason. We suggest that one reason may be to regulate the folding pathway and prevent the accumulation and subsequent aggregation of a kinetic intermediate.

Materials and methods

Protein nomenclature

HSF from either *K. lactis* or *S. cerevisiae* were used in the different experimental procedures. The core regions of HSF from *K. lactis* and *S. cerevisiae* are highly homologous and available data suggest that HSF from these two yeast are conserved both structurally and functionally. *K. lactis* was chosen for biochemical and biophysical analysis because of the structural data available. *S. cerevisiae* was used for in vivo analysis because of the ease of genetics in that system. An alignment of the sequences of the second helices of the DBDs from *K. lactis* and *S. cerevisiae* is shown here:

228 RERFV Q EVLPKYF 240	<i>K. lactis</i>
206 REEFV H QILPKYF 218	<i>S. cerevisiae</i>

The amino acids in the bulged region, highlighted above, have been numbered sequentially from 1–4 (Hardy et al., 2000). The proline at position 6, the center of kink, is underlined. The following abbreviations were used:

Mutant abbreviation	<i>K. lactis</i> mutation	<i>S. cerevisiae</i> mutation
P6A	P237A	P215A
P6K	P237K	P215K
P6G	P237G	P215G
P6D	P237D	P215D

Construction of yeast strains and test for viability of mutant strains

The construction of pHN1031, the *S. cerevisiae* full-length HSF expressed from its own promoter on a yeast CEN/ARS plasmid, has been previously described (Hubl et al., 1994). Mutations in this plasmid were introduced by a mega-primer polymerase chain reaction (PCR) approach and confirmed by DNA sequencing. These plasmids were introduced into the tester strain PS145, which carries the *HSFΔ2::LEU2* chromosomal deletion, as well as a *URA3*-marked plasmid containing the wild-type HSF gene under the control of a *GALI* promoter (Sorger & Pelham, 1988). Transformants were plated on media-containing dextrose to ensure expression of HSF only from its own promoter (Hubl et al., 1994). Once it was obvious that the deletion mutants would support growth, the *URA3*-marked plasmid containing the *GALI:HSF* fusion gene was removed from the strain by growth in the absence of its selective marker and subsequent testing to ensure its loss. Further studies of viability and temperature dependence were performed on the strains expressing only the wild-type or mutant versions of HSF from the CEN/ARS plasmid. Cultures were grown from a single colony in selective liquid media for 8 h. The OD₆₀₀ was then measured. All strains were diluted to the same cell density, then serially diluted in sterile media, spotted on the agar plates, and grown at 30°

(constitutive temperature), or 37, 39, or 41 °C (heat shock temperatures) for three days to measure growth.

β-Galactosidase assay

The strains derived for testing viability and thermotolerance were transformed with the plasmid pHSE2-lacZ, which has an HSE inserted into a disabled *CYCI* promoter upstream of the *lacZ* gene (Sorger & Pelham, 1987). The HSE sequence in this promoter is AGAAGCTTCTAGAAGCTTCTAGAGGATCCC, where consensus GAA repeats are in bold. Yeast were grown on selective synthetic media at 30 °C to mid-log phase (OD₆₀₀ 0.2–0.6) over a 12 to 18 h period. Three aliquots were dispensed into 15 mL tubes such that each aliquot gave the equivalent of 5 mL of cells at an OD₆₀₀ of 0.2. The control aliquot was maintained at 30 °C. Other aliquots were shocked at 37 or 42 °C for 20 min, then were allowed to recover for 90 min to allow expression of β-galactosidase. All cultures were harvested by centrifugation at 4 °C. β-Galactosidase activity was determined as previously described (Hardy et al., 2000). Each reported activity was an average of four experiments, where all strains were assayed on each of four days.

Overexpression and purification of HSF fragments containing the DNA-binding and trimerization domains

The expression plasmid pHN208, previously described as HSF_{DT} (Rye et al., 1993), is a derivative of pET-3b (Studier et al., 1990). It contains the coding sequence for *K. lactis* HSF amino acids 192–394, which includes the DNA-binding and trimerization domains, generating a 23.7 kD protein. Mutations in the proline were introduced into pHN208 by PCR mutagenesis and confirmed by DNA sequencing.

Wild-type and mutant protein were expressed and purified identically, as previously described (Hardy et al., 2000). The protein was then stored on ice and used within one week of purification. Protein concentrations were determined with an extinction coefficient of 33,700 cm⁻¹ M⁻¹, which was calculated from the tyrosine and tryptophan content of the proteins. Samples of all mutant proteins were also analyzed by analytical reversed-phase high-performance liquid chromatography to more accurately determine purity; the values between 86 and 95% were used to adjust the protein concentrations appropriately.

DNA-binding assays

Binding of mutant HSF DNA-binding and trimerization domains to DNA was performed as previously described, using the 145 bp HSE-containing fragment from the plasmid pBL2 (Drees et al., 1997). The HSE in the DNA probe is AGAATATTCTAGAAA, where GAA repeats are shown in bold. Binding assays were performed at room temperature; however, the results at room temperature and 4 °C were the same for wild-type protein (Hardy et al., 2000). Dried gels were used to expose Molecular Dynamics PhosphorImager plates, which were then scanned on the Molecular Dynamics PhosphorImager. The signals of unbound bands and all bands from bound species were quantified using Molecular Dynamics ImageQuant. To estimate binding affinity, the amount of bound DNA was graphed against protein concentration and fitted to a rectangular hyperbola. The data from four independent experiments were used for the curve fitting. The *K_d* was measured as the protein concentration at which 50% of the DNA was bound. The

analysis did not address cooperativity of binding of additional trimers of HSF to the DNA; only global dissociation constants were obtained.

Overexpression and purification of the DBD

The expression plasmid pHN212 described previously (Harrison et al., 1994) is a derivative of pET-3b (Studier et al., 1990) and contains the coding sequence for *K. lactis* HSF amino acids 194 to 282, which includes the DBD, generating a 10.9 kD protein. pHN212 was mutated to replace arginine codons that are infrequently used in *E. coli*, and that in previous studies were often translated as lysine instead of arginine. Site-directed mutagenesis was used to replace the codons for residues R194, R228, R275, and R282* (a C-terminal artifact of the cloning procedure) to the codon CGC, a common *E. coli* arginine codon. This new plasmid, pHN212R, was subsequently mutated by site-directed mutagenesis to change the proline to other residues. All mutations were confirmed by DNA sequencing.

Mutant and wild-type proteins were overexpressed in *E. coli* strain BL21(DE3) containing the pLysS plasmid (Studier et al., 1990). Purification of the wild-type DBD has been described previously (Harrison et al., 1994). The mutations introduced for this study rendered the mutant DBDs less soluble upon overexpression than wild-type DBD. Therefore, they were purified by an alternative method for insoluble DBDs, which has been previously described (Hardy et al., 2000). Qualitatively, the amount of overexpressed mutant protein seemed to vary based on the mutation present. Specifically, much more P6G was typically expressed than P6A or P6K. We speculate that the expression levels may be related to the solubility and folding kinetics of that version of the protein. All protein concentrations were determined with extinction coefficient of $25,400 \text{ cm}^{-1} \text{ M}^{-1}$, which was calculated from the tyrosine and tryptophan content of the proteins.

Circular dichroism spectroscopy

Lyophilized wild-type and mutant HSF DBD constructs were resuspended in water then diluted to $20 \mu\text{M}$ in 20 mM citrate/phosphate buffer (pH 4). CD spectra were recorded on an AVIV 60 DS spectropolarimeter at 4°C . Measurements were taken at 1 nm intervals with a 1 s time constant and 1.5 nm bandwidth. A path-length of 0.1 cm was used. The thermal melting of these constructs was followed at 222 nm. The temperature was raised by 1°C at each step, using a 5 min equilibration time, 10 s averaging time, and 1.5 nm bandwidth. The temperature at the inflection point of the cooperative unfolding curve was assigned as the T_m .

Crystallization

Crystals of P6A, P6K, or wild-type DBD were grown under similar conditions. P6A, P6K, or wild-type DBD that had been purified and lyophilized were suspended at concentrations of 5 to 12 mg/mL in water, then incubated for 25 min at room temperature before the protein was added to buffer-containing solutions. P6K and P6A each crystallized in 22–30% PEG 4K, 50–200 mM ammonium acetate, and 100 mM citrate buffer (pH 5.6 to 6.2). The crystals form as clustered or single thin square plates of up to $0.2 \times 0.2 \times 0.05$ in 4° hanging drop vapor diffusion drops in 1–2 weeks.

Electrospray ionization mass spectrometry of P6A after purification indicated that the molecular weight, as expected, was 10,907

Da. Following crystallization of P6A, a crystal grown in the same tray as the crystal from which the data were collected was dissolved and analyzed by mass spectroscopy. Three distinct species were present in the sample: a 10,907 Da species, which is the calculated molecular weight of P6A, a 10,922.4 species, which corresponds to a singly oxidized form of P6A ($10,907 + 16 = 10,923$), and a 10,938.7 Da species, which is believed to be a doubly oxidized P6A. The most likely candidate for oxidation was methionine, but this cannot be determined from mass spectrometry. These molecular weight species confirmed the identity the P6A mutant. The majority of the protein in the crystal was oxidized, suggesting that oxidation may be a prerequisite for crystallization in this space group.

Structure determination

Crystals were soaked for 1 min in an artificial mother liquor solution that was freshly prepared of the same components as had produced the drop plus 10% methylpentanediol, and then frozen in liquid nitrogen for data collection. The two mutant versions of the DBD, P6K, and P6A, both crystallized in nearly identical unit cells to wild-type, in the space group $P2_12_12$ (see table below).

Protein	a (Å)	b (Å)	c (Å)	$\alpha = \beta = \gamma$
P6K	57.12	63.16	50.19	90°
P6A	57.22	63.05	50.16	90°
Wild-type	56.8	64.10	50.90	90°

Data for P6A and wild-type that were used in refinement were collected at the Stanford Synchrotron Radiation Laboratory beamline 7–1, using a wavelength of 1.08 Å. Data for P6K that were used for refinement were collected at beamline X12B at the National Synchrotron Light Source at Brookhaven using a wavelength of 1.06 Å. The data were processed using DENZO and SCALEPACK (Otwinowski & Minor, 1996). Data reduction statistics are shown in Table 4.

Based on the unit cell dimensions listed in the table above, and the calculated Matthews coefficient (V_m), two molecules are likely to comprise the asymmetric unit. The V_m for one molecule in the asymmetric unit is 4.1, indicating a 70% solvent content, whereas the V_m for two molecules in the asymmetric unit is 2.05, which suggests a solvent content of $\sim 40\%$. Given that these crystals diffracted to relatively high resolution, it appeared likely that two molecules were present in the asymmetric unit.

The P6K structure was solved by molecular replacement using the aMoRe package (Navaza, 1994). The search model was based on the structure of the HSF DBD (PDB accession number 1HTS), previously determined in the space-group $P2_12_12_1$ with one molecule in the asymmetric unit (Harrison et al., 1994). The model that yielded the proper solution contained residues 193–262 and 272–284, which includes all of the residues that are present in the construct except a portion of the extended wing, which is known to be flexible (Damberger et al., 1994, 1995; Harrison et al., 1994). All side chains were contained in the model, except residue 237, where the proline side chain was truncated to an alanine.

Molecular replacement resulted in two clear solutions, corresponding to the two molecules in the asymmetric unit. The two molecules were denoted molecules a and b, respectively. The re-

Table 4. Data and refinement statistics

	P6K	P6A	Wild-type
Data statistics			
Resolution	50–2.0	20–2.0	50–1.83
Observed reflections	42,447	N/A	53,481
Unique reflections	11,753	16,534	13,067
Completeness	91.6	90.8	85.4
R_{sym}^a	3.9	5.3	8.0
I/σ	15.6	40.9	17.1
Refinement statistics			
Resolution	50–2.0	20–2.0	20–1.8
Reflections (working set)	10,363	10,396	13,686
Reflections (test set)	1,190	1,201	1,505
Residues in model molecule a	195–282	195–281	195–282
Residues in model molecule b	195–284	195–284	195–284
Total solvent molecules	172	197	128
R_{cryst}^b	21.1%	23.8%	23.3%
R_{free}^c	26.2%	28.3%	26.9%
RMS bond length	0.014	0.005	0.005
RMS bond angle	1.61	1.72	1.25
RMS dihedral	23.3	21.9	22.3
RMS improper	0.86	0.55	0.61
RMS B bonded main chain	1.00	0.97	1.06
RMS B bonded main chain	1.25	1.19	1.22

$$^a R_{sym} = \sum_h (I_h - \langle I_h \rangle) / \sum_h (\langle I_h \rangle)$$

$$^b R_{factor} = \sum (|F_{obs}| - |F_{calc}|)^2 / \sum (|F_{obs}|)^2$$

^c R_{free} is calculated based on 10% of the total reflections.

sulting electron density map, calculated at 2.0 Å resolution, was easily interpretable. Density was observed for all residues except the two amino terminal residues (193–194) in both the a and b molecules, and the carboxyl-terminal three residues (an artifact of the cloning procedure) in the a molecule. The missing residues were not included in the subsequent models. Refinement consisted of iterative rounds of model rebuilding in O (Jones et al., 1991) and subsequent simulated annealing, restrained isotropic temperature factor refinement, bulk solvent correction and monitoring the free R -factor with 10% of the data omitted using CNS (Brünger et al., 1998). Statistics for the refinement process and the final models are listed in Table 4.

As P6A and wild-type DBD each crystallized in a space group that was isomorphous to that of P6K, it was possible to solve the structures directly using the final P6K model without having to perform either molecular replacement or calculate a difference Fourier map. Refinement for both P6A and wild-type was carried out using iterative rounds of O and CNS, in similar fashion to the P6K refinement.

The final models of all of the proteins show good geometry with no outliers in the Ramachandran plot. For P6K molecules a and b, P6A molecules a and b and wild-type a and b, respectively, 90.2, 90.5, 86.4, 90.5, 91.4, and 91.6% of residues were in the most favored regions of the Ramachandran plot. Models shown in the manuscript were rendered with MOLSCRIPT (Kraulis, 1991), or InsightII (Molecular Simulations, Inc., San Diego, California). Coordinates have been deposited in the Protein Data Bank and have been assigned the accession codes 1FBU for wild-type, 1FBQ for P6K, and 1FBS for P6A.

Tryptophan fluorescence measurements

Steady state fluorescence was measured on an F-2000 fluorescence spectrophotometer (Hitachi, Inc.) at 25° using excitation and emission wavelengths of 295 and 346 nm, respectively. Samples were prepared as 5 μM DBD in 20 mM Tris (pH 7.5), in the presence or absence of various concentrations of the denaturant guanidine hydrochloride. The C_m for the unfolding of all versions of the HSF DBD was ~1.95 M guanidine hydrochloride. Both wild-type and mutant protein were fully unfolded in concentrations of guanidine hydrochloride greater than 3 M, and fully folded at concentrations less than 1 M as indicated by fluorescence intensity. This indicated that the DBD could be rapidly diluted from 3 M or greater guanidine to concentrations less than 1 M to measure refolding kinetics, and vice versa for unfolding kinetic measurements.

Kinetic measurements were performed on a Bio-logic SFM4 stopped flow module (Molecular Kinetics, Inc., Pullman, Washington). For refolding kinetics, samples were prepared as 50 μM DBD in the presence of 20 mM Tris (pH 7.5) and 3 M guanidine hydrochloride. This solution was rapidly diluted by a 1:5 dilution into buffer free of guanidine hydrochloride. The calculated dead time of the instrument was 2.3 ms. Kinetics of refolding were followed by fluorescence that was excited with 280 nm and 1 mm slit width and emission with a 300 nm cutoff filter. Curve fitting was performed Bio-kine software (Bio-logic Co., Echiroles, France). Data were fit with either single exponential functions or a sum of two exponential functions for bi-phasic kinetics using the following equations: $Y = at + b + c_1 e^{-k_1 t}$ or $Y = at + b + c_1 e^{-k_1 t} + c_2 e^{-k_2 t}$. Y is the measured fluorescence, a is the slope, b is the offset, c_1 and c_2 are the amplitudes, and k_1 and k_2 are the rate constants for the two phases of folding. As is the case with most curve fitting routines, curves with a wide range of rate constants could be fitted to the same curve with equivalent values of the statistical validation, in this case χ^2 values. To confirm that the appropriate expression was selected, it was imperative to closely monitor the residual fitting values. The expression with the lowest rate constant with unbiased residual values was selected.

Unfolding kinetics were measured using the same instrument. Wild-type or mutant DBD at 50 μM in 20 mM Tris (pH 7.5) were rapidly diluted with guanidine hydrochloride by a 1:6.67 dilution to a final guanidine concentration of 3.8 M. Fitting of the unfolding curves was done with a single exponential expression and was completely unambiguous.

Accession number

The coordinates have been deposited with the Protein Data Bank. The PDB ID codes are 1FBQ, 1FBS, and 1FBU.

Acknowledgments

We gratefully acknowledge the previous work of Celia J. Harrison in growing the crystals of the wild-type protein and collecting and reducing the data. We thank Charles E. Bell, Steve Huang, Mitchell Lewis, Jamie Schlessman, and Steven Stayrook for assistance with crystallographic procedures. We also thank Leland Mayne for assistance with stopped-flow fluorescence measurements, A. Josh Wand for used of his stopped-flow fluorimeter, and Walter Englander and Andrew Lee for insightful discussions about the manuscript. This work was supported by an NIH grant GM44086 to HCMN. This work is based on research conducted at the NSLS beamline X12B and at the SSRL beamline 7-1.

References

- Alber T, Bell JA, Sun DP, Nicholson H, Wozniak JA, Matthews BW. 1988. Replacements of Pro86 in phage T4 lysozyme extend an alpha-helix but do not alter protein stability. *Science* 239:631–635.
- Barlow DJ, Thornton JM. 1988. Helix geometry in proteins. *J Mol Biol* 201:601–619.
- Bonner JJ, Heyward S, Fackenthal DL. 1992. Temperature-dependent regulation of a heterologous transcriptional activation domain fused to yeast heat shock transcription factor. *Mol Cell Biol* 12:1021–1030.
- Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse KR, Jiang JS, Kuszewski J, Nilges M, Pannu NS, et al. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* 54:905–921.
- Chen Y, Barlev NA, Westergaard O, Jakobsen BK. 1993. Identification of the C-terminal activator domain in yeast heat shock factor: Independent control of transient and sustained transcriptional activity. *EMBO J* 12:5007–5018.
- Cho HS, Liu CW, Damberger FF, Pelton JG, Nelson HC, Wemmer DE. 1996. Yeast heat shock transcription factor N-terminal activation domains are unstructured as probed by heteronuclear NMR spectroscopy. *Protein Sci* 5:262–269.
- Chou PY, Fasman GD. 1974. Prediction of protein conformation. *Biochem* 13:222–245.
- Cohen GH. 1997. Align. *J Appl Crystallogr* 30:1160–1161.
- Damberger FF, Pelton JG, Harrison CJ, Nelson HC, Wemmer DE. 1994. Solution structure of the DNA-binding domain of the heat shock transcription factor determined by multidimensional heteronuclear magnetic resonance spectroscopy. *Protein Sci* 3:1806–1821.
- Damberger FF, Pelton JG, Liu C, Cho H, Harrison CJ, Nelson HC, Wemmer DE. 1995. Refined solution structure and dynamics of the DNA-binding domain of the heat shock factor from *Kluyveromyces lactis*. *J Mol Biol* 254:704–719.
- Drees BL, Grotkopp EK, Nelson HC. 1997. The GCN4 leucine zipper can functionally substitute for the heat shock transcription factors trimerization domain. *J Mol Biol* 273:61–74.
- Giardina C, Lis JT. 1995. Sodium salicylate and yeast heat shock gene transcription. *J Biol Chem* 270:10369–10372.
- Gow A, Southwood CM, Lazzarini RA. 1998. Disrupted proteolipid protein trafficking results in oligodendrocyte apoptosis in an animal model of Pelizaeus-Merzbacher disease. *J Cell Biol* 140:925–934.
- Grathwohl C, Wüthrich K. 1981. The X-Pro peptide bond as an NMR probe for conformational studies of flexible linear peptides. *Biopolymers* 15:2025–2041.
- Gross DS, English KE, Collins KW, Lee SW. 1990. Genomic footprinting of the yeast HSP82 promoter reveals marked distortion of the DNA helix and constitutive occupancy of heat shock and TATA elements. *J Mol Biol* 216:611–631.
- Halladay JT, Craig EA. 1995. A heat shock transcription factor with reduced activity suppresses a yeast HSP70 mutant. *Mol Cell Biol* 15:4890–4897.
- Hardy JA, Walsh STR, Nelson HCM. 2000. Role of an alpha-helical bulge in the yeast heat shock transcription factor. *J Mol Biol* 295:393–409.
- Harrison CJ, Bohm AA, Nelson HC. 1994. Crystal structure of the DNA binding domain of the heat shock transcription factor. *Science* 263:224–227.
- Hoj A, Jakobsen BK. 1994. A short element required for turning off heat shock transcription factor: Evidence that phosphorylation enhances deactivation. *EMBO J* 13:2617–2624.
- Hubl S. 1995. Structure-function studies of the DNA binding domain of heat shock transcription factor [PhD Thesis]. University of California, Berkeley.
- Hubl ST, Owens JC, Nelson HCM. 1994. Mutational analysis of the DNA-binding domain of yeast heat shock transcription factor. *Nat Struct Biol* 1:615–620.
- Jakobsen BK, Pelham HR. 1988. Constitutive binding of yeast heat shock factor to DNA in vivo. *Mol Cell Biol* 8:5040–5042.
- Jakobsen BK, Pelham HR. 1991. A conserved heptapeptide restrains the activity of the yeast heat shock transcription factor. *EMBO J* 10:369–375.
- Javitch JA, Ballesteros JA, Weinstein H, Chen J. 1998. A cluster of aromatic residues in the sixth membrane-spanning segment of the dopamine D2 receptor is accessible in the binding-site crevice. *Biochemistry* 37:998–1006.
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
- Konopka JB, Margarit SM, Dube P. 1996. Mutation of Pro-258 in transmembrane domain 6 constitutively activates the G protein-coupled alpha-factor receptor. *Proc Natl Acad Sci USA* 93:6764–6769.
- Kraulis PJ. 1991. MOLSCRIPT—A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946–950.
- Kumar S, Bansal M. 1998a. Dissecting alpha-helices: Position-specific analysis of alpha-helices in globular proteins. *Proteins* 31:460–476.
- Kumar S, Bansal M. 1998b. Geometrical and sequence characteristics of alpha-helices in globular proteins. *Biophys J* 75:1935–1944.
- Kurada P, O'Tousa JE. 1995. Retinal degeneration caused by dominant rhodopsin mutations in *Drosophila*. *Neuron* 14:571–579.
- Kuznetsov G, Bush KT, Zhang PL, Nigam SK. 1996. Perturbations in maturation of secretory proteins and their association with endoplasmic reticulum chaperones in a cell culture model for epithelial ischemia. *Proc Natl Acad Sci USA* 93:8584–8589.
- Levitt M. 1978. Conformational preferences of amino acids in globular proteins. *Biochemistry* 17:4277–4285.
- Littlefield O, Nelson HCM. 1999. A new use for the “wing” of the “winged” helix-turn-helix motif in the HSF-DNA cocrystal. *Nat Struct Biol* 6:464–470.
- Liu X, Garriga P, Khorana HG. 1996. Structure and function in rhodopsin: Correct folding and misfolding in two point mutants in the intradiscal domain of rhodopsin identified in retinitis pigmentosa. *Proc Natl Acad Sci USA* 93:4554–4559.
- Liu XD, Liu PC, Santoro N, Thiele DJ. 1997. Conservation of a stress response: Human heat shock transcription factors functionally substitute for yeast HSF. *EMBO J* 16:6466–6477.
- Morano KA, Santoro N, Koch KA, Thiele DJ. 1999. A trans-activation domain in yeast heat shock transcription factor is essential for cell cycle progression during stress. *Mol Cell Biol* 19:402–411.
- Navaza J. 1994. Amore—An automated package for molecular replacement. *Acta Crystallogr A* 50:157–163.
- Nicholls A, Bharadwaj R, Honig B. 1993. Grasp graphical representation and analysis of surface properties. *Biophys J* 64:A166.
- Nieto-Sotelo J, Wiederrecht G, Okuda A, Parker CS. 1990. The yeast heat shock transcription factor contains a transcriptional activation domain whose activity is repressed under nonshock conditions. *Cell* 62:807–817.
- Oksche A, Schulein R, Rutz C, Liebenhoff U, Dickson J, Muller H, Birnbaumer M, Rosenthal W. 1996. Vasopressin V2 receptor mutants that cause X-linked nephrogenic diabetes insipidus: analysis of expression, processing, and function. *Mol Pharm* 50:820–828.
- Otwinowski Z, Minor W. 1996. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
- Peteranderl R, Rabenstein M, Shin YK, Liu CW, Wemmer DE, King DS, Nelson HCM. 1999. Biochemical and biophysical characterization of the trimerization domain from the heat shock transcription factor. *Biochemistry* 38:3559–3569.
- Prusiner SB. 1997. Prion diseases and the BSE crisis. *Science* 278:245–251.
- Ramm K, Gehrig P, Pluckthun A. 1999. Removal of the conserved disulfide bridges from the scFv fragment of an antibody: Effects on folding kinetics and aggregation. *J Mol Biol* 290:535–546.
- Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations at the ends of alpha helices. *Science* 240:1648–1652.
- Rye HS, Drees BL, Nelson HC, Glazer AN. 1993. Stable fluorescent dye-DNA complexes in high sensitivity detection of protein-DNA interactions. Application to heat shock transcription factor. *J Biol Chem* 268:25229–25238.
- Satow Y, Cohen GH, Padlan EA, Davies DR. 1986. Phosphocholine binding immunoglobulin Fab McPC603. An X-ray diffraction study at 2.7 Å. *J Mol Biol* 190:593–604.
- Schmid FX. 1993. Prolyl isomerase: Enzymatic catalysis of slow protein-folding reactions. *Annu Rev Biophys Biomolec Struct* 22:123–142.
- Schultheiss J, Kunert O, Gase U, Scharf KD, Nover L, Ruterjans H. 1996. Solution structure of the DNA-binding domain of the tomato heat-stress transcription factor HSF24. *Eur J Biochem* 236:911–921.
- Sewell AK, Yokoya F, Yu W, Miyagawa T, Murayama T, Winge DR. 1995. Mutated yeast heat shock transcription factor exhibits elevated basal transcriptional activation and confers metal resistance. *J Biol Chem* 270:25079–25086.
- Silar P, Butler G, Thiele DJ. 1991. Heat shock transcription factor activates transcription of the yeast metallothionein gene. *Mol Cell Biol* 11:1232–1238.
- Sorger PK. 1990. Yeast heat shock factor contains separable transient and sustained response transcriptional activators. *Cell* 62:793–805.
- Sorger PK. 1991. Heat shock factor and the heat shock response. *Cell* 65:363–366.
- Sorger PK, Pelham HR. 1987. Purification and characterization of a heat-shock element binding protein from yeast. *EMBO J* 6:3035–3041.
- Sorger PK, Pelham HR. 1988. Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. *Cell* 54:855–864.
- Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185:60–89.
- Tamai KT, Liu X, Silar P, Sosinowski T, Thiele DJ. 1994. Heat shock transcription factor activates yeast metallothionein gene expression in response to heat and glucose starvation via distinct signalling pathways. *Mol Cell Biol* 14:8155–8165.
- Vuister GW, Kim SJ, Orosz A, Marquardt J, Wu C, Bax A. 1994. Solution structure of the DNA binding domain of *Drosophila* heat shock transcription factor. *Nat Struct Biol* 1:605–613.

- Williams RW, Chang A, Juretic D, Loughran S. 1987. Secondary structure predictions and medium range interactions. *Biochim Biophys Acta* 916:200–204.
- Wilmot CM, Thornton JM. 1988. Analysis and prediction of the different types of beta-turn in proteins. *J Mol Biol* 203:221–232.
- Woolfson DN, Williams DH. 1990. The influence of proline residues on alpha-helical structure. *FEBS Lett* 277:185–188.
- Wu C. 1995. Heat shock transcription factors: Structure and regulation. *Annu Rev Cell Dev Biol* 11:441.
- Yamaguchi H, Muth JN, Varadi M, Schwartz A, Varadi G. 1999. Critical role of conserved proline residues in the transmembrane segment 4 voltage sensor function and in the gating of L-type calcium channels. *Proc Natl Acad Sci USA* 96:1357–1362.
- Yang WM, Gahl W, Hamer D. 1991. Role of heat shock transcription factor in yeast metallothionein gene expression. *Mol Cell Biol* 11:3676–3681.
- Yuan HS, Wang SS, Yang WZ, Finkel SE, Johnson RC. 1994. The structure of Fis mutant Pro61Ala illustrates that the kink within the long alpha-helix is not due to the presence of the proline residue. *J Biol Chem* 269:28947–28954.