Exploring the Energy Surface of Protein Folding by Structure−Reactivity Relationships and Engineered Proteins: Observation of Hammond Behavior for the Gross Structure of the Transition State and Anti-Hammond Behavior for Structural Elements for Unfolding/Folding of Barnase†

Jacqueline M. Matthews‡ and Alan R. Fersht*

MRC Unit for Protein Function and Design, Cambridge Centre for Protein Engineering, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.

Received January 9, 1995; Revised Manuscript Received March 8, 1995⊗

ABSTRACT: The structure of α-helix$_1$ (residues 6−18) in the transition state for the unfolding of barnase has been previously characterized by comparing the kinetics and thermodynamics of folding of wild-type protein with those of mutants whose side chains have been cut back, in the main, to that of alanine. The structure of the transition state has now been explored further by comparing the kinetics and thermodynamics of folding of glycine mutants with those of the alanine mutants at solvent-exposed positions in the α-helices of barnase. Such "Ala→Gly scanning" provides a general procedure for examining the structure of solvent-exposed regions in the transition state. A gradual change of structure of the transition state was detected as helix$_1$ becomes increasingly destabilized on mutation. The extent of change of structure of helix$_1$ in the transition state for the mutant proteins was probed by a further round of Ala→Gly scanning of those mutants. Destabilization of the helix$_1$ was found to cause the overall transition state for unfolding to become closer in structure to that of the folded protein. This is analogous to the conventional Hammond effect in physical-organic chemistry whereby the transition state moves parallel to the reaction coordinate with change in structure. But, paradoxically, the structure of helix$_1$ itself becomes less folded in the transition state as helix$_1$ becomes destabilized. This is analogous, however, to the rarer anti-Hammond effect in which there is movement perpendicular to the reaction coordinate. These observations are rationalized by plotting correlation diagrams of degree of formation of individual elements of structure against the degree of formation of overall structure in the transition state. There is a relatively smooth movement of the degree of compactness in the transition state against changes in activation energy on mutation that suggests a smooth movement of the transition state along the energy surface on mutation rather than a switch between two different parallel pathways. The results are consistent with the transition state having closely spaced energy levels. Helix$_1$, which appears to be an initiation point and forms early in the folding of wild-type protein, may be radically destabilized to the extent that it forms late in the folding of mutants. The order of events in folding may thus not be crucial.

Some current questions in protein folding are, What is the order of events during folding, what are the structures of intermediates and transition states, are there compulsory initiation events, can folding proceed via different pathways, and what is the nature of the energy surface during folding? Information on the order of events and on the structures of intermediates and transition states may be gleaned from analysis of rate and equilibrium measurements on the folding and unfolding of engineered proteins using structure−activity relationships (Fersht, 1993). A parameter ϕ is obtained that is an index of the degree of formation of a particular element of structure in a transition state (or intermediate). ϕ generally varies between 0 or 1: with 0 meaning that the structure is fully unfolded at the site of mutation, 1 meaning fully folded, and fractional values meaning fractional folding or a mixture of states (Fersht et al., 1992; Matouschek et al., 1989, 1990, 1992a; Fersht, 1995). By applying a ϕ-value analysis, we have obtained a picture at the level of individual residues of the major folding intermediate and the major folding/unfolding transition state of barnase and the order of formation of different elements of secondary structure and the hydrophobic core (Fersht, 1993; Serrano et al., 1992b).

Second-order effects, termed structure−reactivity relationships, are sometimes found in structure−activity relationships in simple organic reactions (Jencks, 1985). These are caused by the structure of the transition state changing with changing reactivity. The most famous of these is the Hammond effect (Hammond, 1955) whereby the structure of a transition state becomes more like that of a neighboring intermediate on the reaction profile as the stability of the intermediate is decreased. There is a lesser known behavior, "anti-Hammond", in which the structure of a transition state becomes less like that of a neighboring intermediate on the reaction profile as the stability of the intermediate is decreased.

Hammond behavior has recently been discovered to occur in protein folding, specifically for mutations in the major α-helix of barnase (α-helix$_1$) (Matouschek & Fersht, 1993).

† This paper is dedicated to Dr. W. P. Jencks for his inspirational work on structure−activity and structure−reactivity relationships.

* To whom correspondence should be addressed.

‡ Present address: Joint Protein Structure Laboratory, Melbourne Tumour Biology Branch, Ludwig Institute For Cancer Research, PO Royal Melbourne Hospital, Victoria 3050, Australia.

⊗ Abstract published in *Advance ACS Abstracts,* May 15, 1995.

0006-2960/95/0434-6805$09.00/0

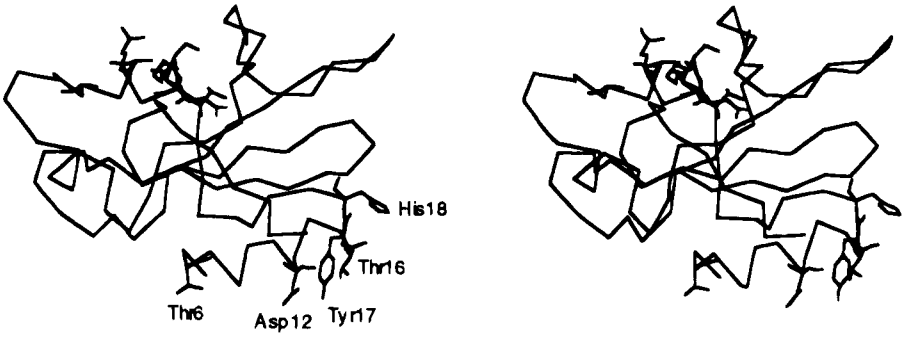

FIGURE 1: Stereodiagram of barnase illustrating residues mutated in $helix_1$ (bottom right) and $helix_2$ (top left).

Structure−reactivity effects may be readily detected in protein folding because the average overall position of a transition state on the reaction coordinate may be determined by comparing the kinetics and equilibria of urea-mediated unfolding as a function of urea concentration (Tanford, 1970). The importance of studying structure−reactivity effects is that they can potentially give information on the energy surface of protein folding and whether one can perturb the folding pathway.

Barnase, a 12 382 Da RNase from *Bacillus amyloliquefaciens* folds from the unfolded state, U, via a folding intermediate, I, and a major transition state for folding/unfolding (‡) between I and the folded state F (eq 1). There

$$U \rightleftarrows \ldots \rightleftarrows I \underset{k_u}{\overset{k_{-u}}{\rightleftharpoons}} F \qquad (1)$$

are probably earlier intermediates which have not, as yet, been detected directly. The major α-helix of barnase (Figure 1) is of interest because it is a putative initiation site and has been shown previously by ϕ-value analysis to be basically intact, apart from the first turn, in the major transition state for unfolding and the major folding intermediate. We now explore this helix further by a second layer of mutations, comparing alanine with glycine at solvent-exposed positions ("Ala→Gly scanning"). The mutations are seen to slide the major transition state for unfolding/folding smoothly around its energy surface, as revealed by the presence of Hammond (1955) and now anti-Hammond effects. The results are consistent with the transition state having closely spaced energy levels. The major α-helix, which appears to be an initiation point and forms early in the folding of wild-type protein, may be radically destabilized to the extent that it forms late in mutants.

MATERIALS AND METHODS

Materials. The buffer used in the denaturation experiments was 2-(*N*-morpholino)ethanesulfonic acid (Mes) from Sigma, St Louis, MO, and the urea was enzyme grade from Bethesda Research Laboratories, MD. Radiochemicals for sequencing were purchased from Amersham International plc, Buckinghamshire, U.K. SP-Trisacryl was from IBF, Villeneuve La Garenne, France. Enzymes were purchased as indicated. Water was purified to 18 MΩ using either an Elgastat Prima reverse osmosis water purifier system or a Millipore polishing unit. All other materials were of analytical grade and were obtained from BDH Ltd, Poole, England, FSA Laboratory Supplies, Loughborough, U.K., Sigma, Poole, U.K., or Aldrich, Gillingham, U.K.

Mutagenesis. The mutations to Ala and Gly at positions 6, 12, 16, 17, and 18 on $helix_1$ of barnase have been described previously: Thr-6→Ala, Thr-6→Gly, Thr-26→Ala, Thr-26→Gly Glu-29→Gly (Matouschek et al., 1989; Serrano & Fersht, 1989); Asp-12→Ala (Serrano et al., 1990); Tyr-17→Ala (Serrano et al., 1991); Lys-27→Ala (Meiering et al., 1992); Ala-32→Gly (Horovitz et al., 1992); Thr-16→Ala, Thr-16→Gly, Tyr-17→Gly, His-18→Ala, His-18→Gly, Lys-27→Gly, Ser-28→Ala, Ser-28→Gly, Gln-31→Ala, Gln-31→Gly (Serrano et al., 1992c); Asp-12→Gly and Gly-34→Ala (Serrano et al., 1992d).

The double mutants Asp-12→Ala/Tyr-17→Gly and Asp-12→Ala/Tyr-17→Gly were constructed using inverse polymerase chain reaction (IPCR) mutagenesis. The oligonucleotides used to direct the mutagenesis were designed in pairs, to bind to the coding strand and complementary strand of the plasmid containing the barnase gene, in a blunt-ended fashion. Mismatched bases are underlined. The oligonucleotides were synthesised using an Applied Biosystems 380B DNA synthesiser. Double-stranded wild-type barnase DNA was used as the template in all cases. Oligonucleotides (1) and (2) were used to direct the mutagenesis of Asp-12→Ala/Tyr-17→Gly and (1) and (3) were used to produce Asp-12→Gly/Tyr-17→Gly.

(1) 5′-CAG ACA GGT AAG CTA CCT GAT AAT TAC-3′

(2) 5′-AAG ATA AGC CGC AAC CCC GTC AAA CGT GTT-3′

(3) 5′-AAG ATA ACC CGC AAC CCC GTC AAA CGT GTT-3′

The IPCR reaction mixture (oligonucleotides, 10 μM, 2 μL; dNTPs, 5 μM, 1 μL; DNA, 1:100, 5−10 μL; 10× *Pfu* buffer, 2.5 μL, and Exo⁻*Pfu* polymerase, from Stratagene, in 25 μL) was heated to 94 °C for 5 min, followed by 25 cycles of 1 min at 94 °C, 1 min at 55 °C, and 7 min at 72 °C. The mixture was then held at 72 °C for 30 min. IPCR products were visualized and separated on TAE agarose gels containing ethidium bromide. The IPCR products were cleaned using a USBioclean Kit and ligated using T4 DNA Ligase (Boeringer Mannheim). The ligation product was transformed into competent TG2 cells for single-stranded template preparations and expression. The mutations was screened for and the gene fully sequenced using dideoxy nucleotide sequencing.

Expression and Purification of Proteins. The standard expression and purification of barnase has been described

in detail (Horovitz et al., 1990; Kellis et al., 1989; Serrano et al., 1990). For severely destabilized mutants, urea (7 M) was included in a second elution step from SP Trisacryl, and in the high salt FPLC buffer, to allow more complete recovery of protein from the resin and columns (Vuilleumier et al., 1993).

Urea Denaturation Experiments. The buffer stock solution was 1 M MES, containing 387 mM acid form and 613 mM sodium salt, which gives a pH of 6.3 at 25 °C on dilution to 50 mM. All experiments were at 25.0 ± 0.1 °C. Temperature was carefully monitored using an Edale Model C thermometer. The extent of unfolding was monitored by fluorescence spectroscopy with excitation at 290 nm and emission at 315 nm.

(a) Equilibrium Unfolding. The equilibrium urea denaturation experiments for many of the described mutants have been performed previously (Horovitz et al., 1992; Matouschek et al., 1989; Meiering et al., 1992; Serrano & Fersht, 1989; Serrano et al., 1990, 1991, 1992c,d). Other proteins were denatured as described by Horovitz et al. (1992). Denaturation is assumed to follow the standard equation of $\Delta G_{U\text{-}F} = \Delta G^{H_2O}_{U\text{-}F} - m_{U\text{-}F}[\text{urea}]$, where $\Delta G_{U\text{-}F}$ is the free energy of unfolding in urea, $\Delta G^{H_2O}_{U\text{-}F}$, the energy in water, and $m_{U\text{-}F}$ is a constant. This equation be transformed to eq 2 as explained by Clarke and Fersht (1993):

$$F = [(a_N + b_N[\text{urea}]) + (a_D + b_D[\text{urea}]) \exp\{m_{U\text{-}F}([\text{urea}] - [\text{urea}]_{50\%})/RT\}]/[1 + \exp\{m_{U\text{-}F}([\text{urea}] - [\text{urea}]_{50\%})/RT\}] \quad (2)$$

where F is the observed fluorescence, a_N is the fluorescence of the native form and a_D that of the denatured form in water, and b_N and b_D are the slopes of the linear changes of a_N and a_D with [urea], T is temperature in Kelvin, and R is the gas constant. This transformation provides values of $[\text{urea}]_{50\%}$, the concentration of urea at which 50% of the protein is denatured, and $m_{U\text{-}F}$ and their standard errors using conventional curve fitting with the program Kaleidagraph (Abelbeck Software). Since $m_{U\text{-}F}$ does not change significantly for individual mutants of barnase (Serrano et al., 1992a), $\Delta\Delta G_{U\text{-}F}$, the difference in stability between two mutant proteins, is calculated by eq 3 where $\langle m \rangle$ is an average value

$$\Delta\Delta G_{U\text{-}F} = \langle m \rangle [\text{urea}]_{50\%} \quad (3)$$

of $m_{U\text{-}F}$ (= 1.96 ± 0.01) and $\Delta[\text{urea}]_{50\%}$ is the difference in $[\text{urea}]_{50\%}$ between the two mutants. This gives values of $\Delta\Delta G_{U\text{-}F}$ that are consistent with those found from differential scanning calorimetry. Although eq 2 is based on the standard linear equation $\Delta G_{U\text{-}F} = \Delta G^{H_2O}_{U\text{-}F} - m_{U\text{-}F}[\text{urea}]$, eq 3 still gives precise results for more complex equations where there is a constant degree of curvature in plots of $\Delta G_{U\text{-}F}$ of wild type and mutants versus [urea] (Matouschek et al., 1994). It has now been shown that there is curvature in such plots for barnase (Johnson & Fersht, 1995), but eq 2 does hold over a narrow range of concentrations of urea. The use of the precise form of eq 2 is important since the enables the calculation of errors in $\Delta[\text{urea}]_{50\%}$ and hence the precision of $\Delta\Delta G_{U\text{-}F}$ from eq 3.

(b) Kinetics of Unfolding. Kinetic unfolding experiments were performed as described previously (Matouschek et al., 1989; Serrano et al., 1992a). Both protein (≈0.25 mg mL^{-1}) and urea solutions contained 50 mM MES. Concentrations of urea on dilution were usually between 8.5 and 1.5 M units above the value of $[\text{urea}]_{50\%}$ determined in equilibrium studies (3−6.5 M).

The rate constants at different [urea] were fitted to eq 4, which has been found experimentally to describe the unfolding data of a very large number of mutants:

$$\log k_u = \log k_u^{H_2O} + m^{\ddagger}[\text{urea}] - 0.014[\text{urea}]^2 \quad (4)$$

where k_u is the rate constant of unfolding and $m^{\ddagger}$ varies with mutation (Matouschek et al., 1994). The standard error of the second-order term (−0.014) is ±1 0.001 (Matouschek et al., 1994).

Folding Kinetics. Folding experiments were performed as described previously (Matouschek et al., 1990, 1992a). The folding buffer was 7 mM acid form and 93 mM sodium salt of MES. The kinetic traces for the changes of fluorescence on folding were fitted to a triple-exponential equation. The fast and major phase represents the folding of the bulk of the unfolded protein (75%) that has all three peptidyl−prolyl bonds in the *trans* conformation. The slow phases involve *cis*−*trans* isomerization of the unfolded species containing *cis* peptidyl−prolyl bonds. Only the rate constants of the major folding phase are used in this analysis.

Analysis of the Energetics. The free energy changes of the transition state ($\Delta\Delta G_{\ddagger\text{-}F}$) and intermediate ($\Delta\Delta G_{I\text{-}F}$) upon mutation from Ala to Gly at any given position were calculated by eqs 5 and 6.

$$\Delta\Delta G_{\ddagger\text{-}F} = -RT \ln(k_u/k'_u) \quad (5)$$

$$\Delta\Delta G_{I\text{-}F} = -RT \ln[(k_u/k_{-u})/(k'_u/k'_{-u})] \quad (6)$$

where k_u and k'_u are the rate constants of unfolding (and k_{-u} and k'_{-u} are the rate constants of folding) for the Ala and Gly proteins, respectively. ϕ values for the transition state ($\phi_{\ddagger\text{-}U}$) are calculated from the equation $\phi_{\ddagger\text{-}U} = \Delta\Delta G_{\ddagger\text{-}U}/\Delta\Delta G_{F\text{-}U}$ and for the intermediate by $\phi_{I\text{-}U} = \Delta\Delta G_{I\text{-}U}/\Delta\Delta G_{F\text{-}U}$.

RESULTS

Tables 1−3 refer to single Ala and Gly mutations along the solvent-exposed faces of the two main α-helices. A further generation of mutants was constructed in which Asp-12 was mutated to Ala and Gly in the mutant Tyr-17→Gly (Tables 4−6). Results are quoted with ± standard errors of the mean. Measurements of $[\text{urea}]_{50\%}$, $m_{U\text{-}F}$, k_{-u}, log k_u, and $m^{\ddagger}$ in H_2O, 4.0 M urea (the approximate $[\text{urea}]_{50\%}$ for barnase mutants), and 7.25 M urea (the midpoint of the unfolding range of urea concentrations for wild-type barnase) for the mutants described are shown in Table 1. The standard errors for log k_u are largest for the unfolding reaction in H_2O because of the longest extrapolation. The errors are the smallest at 7.25 M urea, which is in the middle of the range of concentrations of urea used in the experimental determination and so no extrapolation is required. Also listed in Table 1 are the apparent values of $m^{\ddagger}$ (termed $m^{\ddagger}_{app}$) at 4 and 7.25 M urea. ($m^{\ddagger} = \partial \log k_u/\partial [\text{urea}]$, i.e., the tangents of the curve of log k_u versus [urea] at those concentrations.) Over a short range of concentrations of urea the plots of log k_u versus [urea] are not perceptively curved. The values of $m^{\ddagger}_{app}$ at 4 M urea are obtained from the fit of eq 4 to the

Table 1: Equilibrium and Kinetic Parameters for the Unfolding and Refolding of Ala and Gly Mutants Along the Surface-Exposed Faces of the Two Main α-Helices in Barnase at 25 °C and pH 6.3[a]

mutant	$[urea]_{50\%}$ (M)	$m_{U\text{-}F}$ (kcal mol^{-1} M^{-1})	k_u (s^{-1})	log k_u			$m^{\ddagger}$	$m^{\ddagger}_{app}$	$m^{\ddagger}_{app}$
				H_2O	4 M urea	7.25 M urea	H_2O	4 M urea	7.25 M urea
WT	4.58 ± 0.01	1.94 ± 0.06	12.7 ± 0.1[b]	−4.60 ± 0.01	−2.21 ± 0.01	−0.604 ± 0.003	0.653 ± 0.001	0.541	0.450
Thr-6→Ala	3.47 ± 0.02[b]	1.97 ± 0.08	13.5 ± 1.0[b]	−3.24 ± 0.05[f]	−0.98 ± 0.02[f]	0.527 ± 0.007[f]	0.605 ± 0.003[f]	0.509[f]	0.418[f]
Thr-6→Gly	3.95 ± 0.02[b]	2.08 ± 0.08	12.4 ± 0.3[b]	−3.66 ± 0.04[f]	−1.52 ± 0.01[f]	−0.11 ± 0.01[f]	0.591 ± 0.007[f]	0.479[f]	0.388[f]
Asp-12→Ala	4.42 ± 0.02[c]	1.91 ± 0.08	15.3 ± 0.2	−4.33 ± 0.05	−2.02 ± 0.02	−0.469 ± 0.006	0.634 ± 0.007	0.521	0.431
Asp-12→Gly	3.97 ± 0.02[d]	1.87 ± 0.08	11.2 ± 0.1	−4.26 ± 0.02	−1.96 ± 0.01	−0.410 ± 0.002	0.633 ± 0.003	0.521	0.430
Thr-16→Ala	4.43 ± 0.02[d]	1.88 ± 0.08	12.0 ± 0.1	−4.56 ± 0.07	−2.19 ± 0.03	−0.602 ± 0.008	0.647 ± 0.010	0.535	0.444
Thr-16→Gly	3.72 ± 0.02[e]	2.01 ± 0.08	8.8 ± 0.1	−4.22 ± 0.07	−1.97 ± 0.03	−0.460 ± 0.010	0.620 ± 0.010	0.507	0.417
Tyr-17→Ala	3.53 ± 0.02[g]	2.02 ± 0.08	8.6 ± 0.1	−3.87 ± 0.04	−1.70 ± 0.02	−0.276 ± 0.007	0.597 ± 0.006	0.464	0.394
Tyr-17→Gly	2.50 ± 0.02[d]	1.91 ± 0.08	6.9 ± 0.1	−2.88 ± 0.02	−0.91 ± 0.01	0.351 ± 0.005	0.546 ± 0.003	0.434	0.344
His-18→Ala	3.53 ± 0.01	1.91 ± 0.05	8.19 ± 0.02	−3.93 ± 0.07	−1.85 ± 0.03	−0.487 ± 0.008	0.576 ± 0.010	0.464	0.373
His-18→Gly	4.11 ± 0.02	1.95 ± 0.08	9.16 ± 0.02	−4.33 ± 0.05	−2.01 ± 0.03	−0.430 ± 0.010	0.633 ± 0.007	0.521	0.430
Thr-26→Ala	3.58 ± 0.02[b]	2.00 ± 0.08	14.0 ± 0.3[b]	−2.89 ± 0.04[f]	−0.78 ± 0.02[f]	0.600 ± 0.007[f]	0.591 ± 0.006[f]	0.479[f]	0.388[f]
Thr-26→Gly	3.84 ± 0.02[b]	2.03 ± 0.08	12.9 ± 0.1	−3.18 ± 0.03[f]	−1.17 ± 0.01[f]	0.130 ± 0.010[f]	0.591 ± 0.007[f]	0.449[f]	0.358[f]
Lys-27→Ala	4.88 ± 0.02	1.93 ± 0.12	13.4 ± 0.3	−4.81 ± 0.03	−2.51 ± 0.01	−0.964 ± 0.003	0.632 ± 0.004	0.520	0.428
Lys-27→Gly	4.29 ± 0.02[e]	2.11 ± 0.08	17.1 ± 0.7	−3.80 ± 0.02	−1.69 ± 0.01	−0.300 ± 0.003	0.585 ± 0.003	0.472	0.382
Ser-28→Ala	4.78 ± 0.02[e]	1.98 ± 0.08	12.5 ± 0.1	−4.98 ± 0.07	−2.53 ± 0.03	−0.868 ± 0.008	0.668 ± 0.009	0.556	0.465
Ser-28→Gly	4.34 ± 0.02[d]	1.88 ± 0.08	13.4 ± 0.1	−4.32 ± 0.01	−1.98 ± 0.01	−0.403 ± 0.001	0.641 ± 0.002	0.530	0.439
Glu-29→Ala	4.11 ± 0.02	1.83 ± 0.10	12.8 ± 0.3	−4.07 ± 0.04	−1.73 ± 0.02	−0.147 ± 0.006	0.643 ± 0.006	0.531	0.440
Glu-29→Gly	3.64 ± 0.02[e]	1.81 ± 0.08	12.7 ± 0.1	−3.20 ± 0.10	−0.93 ± 0.05	0.580 ± 0.020	0.622 ± 0.015	0.510	0.419
Gln-31→Ala	4.62 ± 0.02[d]	1.83 ± 0.08	15.5 ± 0.3	−4.33 ± 0.05	−2.11 ± 0.02	−0.640 ± 0.006	0.611 ± 0.007	0.499	0.408
Gln-31→Gly	4.07 ± 0.02[d]	1.89 ± 0.08	19.1 ± 0.7	−3.59 ± 0.05	−1.37 ± 0.02	0.100 ± 0.008	0.610 ± 0.008	0.498	0.407
Ala-32→Gly	4.12 ± 0.01	1.94 ± 0.06	13.3 ± 0.1	−3.76 ± 0.08	−1.46 ± 0.04	0.090 ± 0.010	0.631 ± 0.010	0.520	0.429
Gly-34→Ala	2.97 ± 0.02[d]	2.06 ± 0.08	15.0 ± 0.3	−2.23 ± 0.08	0.30 ± 0.01	2.03 ± 0.070	0.689 ± 0.020	0.577	0.486

[a] Results are quoted with ±1 standard errors of the mean. The value of the free energy of unfolding in water can be calculated from columns 2 and 3 using the equation $\Delta G_{U\text{-}F} = \Delta G^{H_2O}_{U\text{-}F} - m_{U\text{-}F}[urea]$; that is, $\Delta G^{H_2O}_{U\text{-}F} = m_{U\text{-}F}[urea]_{50\%}$. The standard errors for log k_u are largest for H_2O because of the longest extrapolation and smallest at 7.25 M urea, which is in the middle of the range of concentrations of urea used in the experimental determination. [b] Data from Matouschek et al. (1989). [c] From Horovitz et al. (1990). [d] From Serrano et al. (1992). [e] Data from Matouschek et al. (1992). [f] Recalculated from original data of Matouschek et al. (1989). [g] Data from Horovitz et al. (1991).

Table 2: Changes in the Equilibrium and Activation Energies of Unfolding for the Mutations Ala→Gly in the Surface-Exposed Faces of the Two Major Helices of Barnase[a]

position of Ala→Gly mutation[b]	$\Delta\Delta G_{U\text{-}F}$	$\Delta\Delta G_{I\text{-}F}$	$\Delta\Delta G_{\ddagger\text{-}F}$		
			H_2O	4 M urea	7.25 M urea
Thr-6	−0.94 ± 0.05	−0.52 ± 0.36	−0.57 ± 0.08	−0.73 ± 0.03	−0.87 ± 0.02
Asp-12	0.88 ± 0.06	0.27 ± 0.13	0.09 ± 0.10	0.08 ± 0.03	0.08 ± 0.01
Thr-16	1.39 ± 0.06	0.64 ± 0.16	0.46 ± 0.14	0.31 ± 0.06	0.19 ± 0.02
Tyr-17	2.02 ± 0.06	1.48 ± 0.09	1.35 ± 0.06	1.08 ± 0.03	0.86 ± 0.01
His-18	−1.13 ± 0.04	−0.61 ± 0.12	−0.54 ± 0.12	−0.22 ± 0.06	0.01 ± 0.02
Thr-26	−0.51 ± 0.05	−0.16 ± 0.15	−0.22 ± 0.11	−0.53 ± 0.02	−0.65 ± 0.02
Lys-27	1.14 ± 0.06	1.22 ± 0.24	1.37 ± 0.05	1.12 ± 0.02	0.91 ± 0.01
Ser-28	0.86 ± 0.06	0.85 ± 0.10	0.89 ± 0.09	0.75 ± 0.04	0.64 ± 0.01
Glu-29	1.08 ± 0.01	1.45 ± 0.10	1.45 ± 0.09	1.10 ± 0.07	0.98 ± 0.03
Gln-31	1.08 ± 0.06	0.89 ± 0.22	1.01 ± 0.10	1.01 ± 0.05	1.01 ± 0.01
Ala-32	0.90 ± 0.03	1.12 ± 0.13	1.14 ± 0.11	1.03 ± 0.05	0.95 ± 0.02
Gly-34	−3.15 ± 0.04	−3.13 ± 0.14	−3.23 ± 0.12	−3.43 ± 0.02	−3.59 ± 0.09

[a] All values are in kcal mol^{-1}. Energies are relative to Ala at the position: a positive value of $\Delta\Delta G_{U\text{-}F}$ or $\Delta\Delta G_{I\text{-}F}$ means that Ala is stabilizing compared with Gly; a positive value of $\Delta\Delta G_{U\text{-}\ddagger}$ means that the Gly mutant unfolds faster. [b] Wild-type residue which is mutated to Ala and Gly.

experimental data. The values at 7.25 M are measured directly from the experimental data by fitting the data points to a simple linear regression curve. The standard errors of these directly measured values of $m^{\ddagger}_{app}$ at 7.25 M are all in the range ±0.5 to ±2%, the majority being ±1%.

Differences in free energy between the Ala and Gly mutants, $\Delta\Delta G_{I\text{-}F}$ (in H_2O) and $\Delta\Delta G_{\ddagger\text{-}F}$ (in H_2O, 4 M urea, and 7.25 M urea) are listed in Table 2. Again, the values are most precise at 7.25 M urea. The energies quoted are relative to the folded state, as these are the way the raw data are measured. The values of ϕ may be calculated using the relationships $\phi_{\ddagger\text{-}U} = 1 - \Delta\Delta G_{\ddagger\text{-}F}/\Delta\Delta G_{U\text{-}F}$ and $\phi_{I\text{-}U} = 1 - \Delta\Delta G_{I\text{-}F}/\Delta\Delta G_{U\text{-}F}$. Values of $\phi_{I\text{-}U}$ in H_2O and $\phi_{\ddagger\text{-}U}$ in H_2O, 4 M urea, and 7.25 M urea for Ala→Gly mutations are listed in Table 3. The values of $\phi_{I\text{-}U}$ in H_2O and $\phi_{\ddagger\text{-}U}$ in H_2O require extrapolation of log k_u to zero molar urea using eq 4. It has been shown elsewhere that the values of $\Delta\Delta G_{I\text{-}F}$ and $\Delta\Delta G_{\ddagger\text{-}F}$ are reasonably tolerant of errors in the precise form of eq 4 as long as the curvature is constant, but the values of ϕ are the least accurate at 0 M urea because of extrapolation (Matouschek et al., 1994). The values of $\phi_{\ddagger\text{-}U}$ at 7.25 M are the most accurate and the most reliable. As 7.25 M urea is in the experimental range for the determination of $\Delta\Delta G_{\ddagger\text{-}F}$, there is no extrapolation of data and direct experimental data are used. Further, the values of $\Delta\Delta G_{U\text{-}F}$ used are also determined at high concentrations of urea so that kinetic and equilibrium data are measured under similar conditions.

Timing of Formation of Helices: Formation of α*-Helix*$_2$. $Helix_2$ is the only significant region whose timing of formation was not well documented previously for the

Table 3: ϕ Values for Formation of the Intermediate and Transition States for the Mutations Ala→Gly in the Solvent-Exposed Faces of the Two Major α-Helices in Barnase

position of Ala→Gly mutation	ϕ_{I-U} H_2O	$\phi_{\ddagger-U}$ H_2O	4 M urea	7.25 M urea
$helix_1$				
Thr-6	0.45 ± 0.38	0.39 ± 0.09	0.22 ± 0.06	0.08 ± 0.06
Asp-12	0.69 ± 0.15	0.90 ± 0.11	0.91 ± 0.04	0.91 ± 0.01
Thr-16	0.54 ± 0.13	0.67 ± 0.10	0.78 ± 0.05	0.86 ± 0.01
Tyr-17	0.26 ± 0.05	0.33 ± 0.04	0.46 ± 0.02	0.58 ± 0.01
His-18	0.46 ± 0.11	0.52 ± 0.11	0.81 ± 0.05	1.01 ± 0.02
$helix_2$				
Thr-26	0.68 ± 0.30	0.58 ± 0.22	−0.05 ± 0.12	−0.27 ± 0.14
Lys-27	−0.07 ± 0.21	−0.20 ± 0.07	0.03 ± 0.05	0.21 ± 0.04
Ser-28	0.01 ± 0.13	−0.04 ± 0.13	0.13 ± 0.07	0.26 ± 0.05
Glu-29	−0.35 ± 0.09	−0.34 ± 0.09	−0.021 ± 0.06	0.09 ± 0.02
Gln-31	0.17 ± 0.21	0.05 ± 0.11	0.06 ± 0.07	0.06 ± 0.05
Ala-32	−0.24 ± 0.13	−0.27 ± 0.14	−0.15 ± 0.07	−0.05 ± 0.04
Gly-34	0.01 ± 0.05	−0.03 ± 0.04	−0.09 ± 0.02	−0.14 ± 0.03

[a] $\phi_{\ddagger-U}$ values are calculated for H_2O, 4 M urea, and 7.25 M urea. ϕ_{I-U} values are calculated for H_2O. $\phi_{\ddagger-U}$ is defined by $\phi_{\ddagger-U} = \Delta\Delta G_{\ddagger-U}/\Delta\Delta G_{F-U}$, where $\Delta\Delta G_{F-U}$ is the change in equilibrium free energy of unfolding on mutation and $\Delta\Delta G_{Q-U}$ is the change in free energy of the transition state.

folding pathway. There were only two protein engineering probes for the folding of $helix_2$ used earlier: at positions Thr-26 and Glu-29 (Matouschek et al., 1990, 1992a). These showed that the helix is mainly not formed until after the rate-determining transition state. Initial NMR H/^{2}H exchange studies were able to trace the progress of only two hydrogen bonds in $helix_2$ involving the backbone NH groups of Ala-32 and Leu-33 (Bycroft et al., 1990; Fersht, 1993; Horovitz et al., 1991; Matouschek et al., 1989, 1990). These protons appeared to be protected rapidly, implying that at least the second half of $helix_2$ forms early. However, the NMR data for this region were of poor resolution. The new data in Table 3 using the Ala→Gly probes show clearly that the whole of $helix_2$ forms after the rate-determining transition step since the values of ϕ_{I-U} and $\phi_{\ddagger-U}$ are all close to 0. The apparent exception is Ala→Gly-26 in water which has large errors associated with the small values of $\Delta\Delta G_{U-F}$ and extrapolation to 0 M urea. The more accurate data at higher [urea] indicate that $helix_2$ is also disordered at residue 26 in the transition state for folding and hence also in the earlier folding intermediate.

Formation of α-Helix$_1$. All earlier evidence showed clearly that the major helix is present from residues 8−18 in the transition state for folding/unfolding and in the folding intermediate (Horovitz et al., 1990; Matouschek et al., 1989, 1990: Serrano et al., 1992). The data for the Ala→Gly probes Asp-12, Thr-16, and His-18 in $helix_1$ agree well with this, but there are some significant differences that provide clues to the behavior of the transition state on mutation which are discussed below.

DISCUSSION

Ala→Gly Scanning. The structures of the major transition state for the folding of barnase and the intermediate preceding it have been previously analyzed by measuring the equilibria and rates of folding and unfolding of its mutants (Matouschek et al., 1989; Fersht et al., 1992; Fersht, 1993). Each mutation was designed to remove a defined interaction in the folded protein and so destabilize it by an energy $\Delta\Delta G_{F-U}$ (where subscript F = folded, and U = unfolded). The energy of destabilization of, say, the transition state (‡) $\Delta\Delta G_{\ddagger-U}$, is calculated from the kinetics. The ratio of the two energies, $\phi_{\ddagger-U} = \Delta\Delta G_{\ddagger-U}/\Delta\Delta G_{F-U}$, provides a measure of the degree of formation of structure in the transition state: $\phi_{\ddagger-U} = 0$ means the structure has the same response to mutation as does the unfolded state and so implies it is unfolded in the transition state. Conversely, $\phi_{\ddagger-U} = 1$ implies that the transition state responds to mutation in the same way as the fully folded structure and so implies the structure is fully folded in the transition state. Fractional values of ϕ imply an intermediate degree of formation of structure. The caveats are discussed in detail by Fersht et al. (1992) and more generally by Fersht (1995).

The so-called ϕ-value analysis works best when defined and simple interactions can be removed by mutagenesis. Each mutation is a "kinetic probe". These do not always exist, however. Sometimes too many interactions are made by a side chain or there is an insufficient change in $\Delta\Delta G_{F-U}$ to be measured accurately. A new, and general, approach for introducing kinetic probes in solvent-exposed regions of α-helices is applied here, the comparison of Ala and Gly substitutions at solvent-exposed positions. We have shown that the relative stability of Ala→Gly mutants in α-helices can be directly proportional to the difference in solvent exposure of the surrounding groups (Serrano et al., 1992c,d). This suggests that, in practice, the mutation of Ala→Gly is a relatively benign substitution and the relative helix-forming propensities depend on the burial of solvent-accessible surface areas. We now make a second generation of mutants by truncating the larger wild-type side chains to Ala and Gly and then use the energetics of the mutation Ala→Gly for the ϕ-value analysis. The mutation Ala→Gly fits in with our general philosophy of making a small deletion mutation. There are very few good probes for the structure of α-$helix_2$ in barnase, and so the only convenient probe is Ala→Gly. The data here clearly show that the formation of $helix_2$ in barnase is a late event in folding. The use of Ala→Gly scanning has thus filled in the picture of the transition state for barnase. In addition, the additional Ala→Gly mutations in α-$helix_1$ analyzed in this study perturbs the structure of the transition state and leads to the discovery of a range of structure−reactivity effects, corresponding to both Hammond and anti-Hammond behavior.

Gross Position of the Transition State. Proteins unfold in the presence of urea because it stabilizes the side chains and backbone of proteins—the more a protein is unfolded, the more the buried groups are exposed to urea. To a fair approximation, the free energy of transfer of those components of proteins to urea solutions is linearly proportional to [urea], hence the usual equation (eq 7) used for analyzing

Table 4: Kinetic Parameters for the Unfolding and Folding of the Barnase Mutant Tyr-17→Gly in Which Asp-12 Is Mutated to Ala and Gly

mutant	$[urea]_{50\%}$ (M)	k_f (s^{-1})	log k_u H_2O	4 M urea	7.25 M urea
Asp-12→Ala/Tyr17→Gly	2.49 ± 0.01	6.4 ± 0.1	−2.74 ± 0.02	−0.73 ± 0.01	0.58 ± 0.01
Asp-12→Gly/Tyr17→Gly	2.08 ± 0.26	4.7 ± 0.1	−2.28 ± 0.04	−0.32 ± 0.01	0.94 ± 0.02

Table 5: m Values for Equilibrium and Kinetic Parameters for the Unfolding and Folding for Wild-Type Barnase and Mutants at Asp-12 and Tyr-17

mutant	$m_{U\text{-}F}$	$m^{\ddagger}$ H_2O	$m^{\ddagger}_{app}$ 4 M urea	$m^{\ddagger}_{app}$ 7.25 M urea
Asp-12→Ala/Tyr-17→Gly	2.19 ± 0.08	0.559 ± 0.002	0.447	0.356
Asp-12→Gly/Tyr-17→Gly	1.99 ± 0.11	0.544 ± 0.007	0.433	0.342

urea denaturation curves (Tanford, 1970). Similarly, the kinetics of unfolding are usually assumed to follow eq 8.

$$\Delta G_{U\text{-}F} = \Delta G^{H_2O}_{U\text{-}F} - m_{U\text{-}F}[\text{urea}] \quad (7)$$

$$\Delta G_{\ddagger\text{-}F} = \Delta G^{H_2O}_{\ddagger\text{-}F} - m_{\ddagger\text{-}F}[\text{urea}] \quad (8)$$

The constant $m_{U\text{-}F}$ [$= -\partial(\Delta G_{U\text{-}F})/\partial[\text{urea}]$] may be calculated from free energy of transfer data for the side chains and backbone from water to solutions of urea: the full transfer free energy is attenuated by a factor which is the average fractional change of exposure to solvent of the side chains and backbone on going from the folded to the unfolded state (Tanford, 1970). The constant $m_{\ddagger\text{-}F}$ has the same transfer free energy terms as in $m_{U\text{-}F}$ but a numerically smaller term for the fractional change of exposure of the protein on going from the folded to the transition state for unfolding. The ratio $m_{\ddagger\text{-}F}/m_{U\text{-}F}$ is thus traditionally used as an index for the relative degree of unfolding of the transition state compared with the unfolded state (Tanford, 1970). As such, the ratio $m_{\ddagger\text{-}F}/m_{U\text{-}F}$ can be used an index of the position of the transition state on the reaction coordinate, r. We have complicated the analysis slightly by finding that eq 4 for barnase is a simplification and there is a small term in [urea]^2. The rate constant for unfolding follows eq 4. We obtain an apparent value of $m_{\ddagger\text{-}F}$ at any concentration of urea from the kinetics of unfolding: $m^{\ddagger}_{app} = \partial \log k_u/\partial[\text{urea}] = (\partial\Delta G_{\ddagger\text{-}F}/\partial[\text{urea}])/RT$, where $m^{\ddagger}_{app}$ varies according to [urea]. Thus, $RTm^{\ddagger}_{app}/m_{U\text{-}F}$ is a measure of the position of the transition state on the reaction pathway.

A Radical Mutation Retards Formation of Helix$_1$. All earlier protein engineering and NMR evidence showed clearly that the major helix is close to being completely formed from residues 12−18 in the transition state for folding and in the folding intermediate (Fersht, 1993; Horovitz & Fersht, 1992; Horovitz et al., 1991; Matouschek et al., 1989, Matouschek et al., 1992a,b). The data for the Ala→Gly probes at positions 12, 16, and 18 in helix$_1$ agree well with this. The ϕ value for Ala-17→Gly, however, is clearly fractional and indicates incomplete formation of this part of the helix in the intermediate and transition states. Thus, the final turn of the helix$_1$ may be less formed during the folding pathway of barnase than for the wild-type protein when Tyr-17 is mutated to Ala and to Gly. The mutation of Tyr-17→Gly is by far the most radical performed in this study, destabilising the protein by 4 kcal mol^{-1}. Most other mutations destabilise the protein by only 0.5−2 kcal mol^{-1}.

Probing the Effects of the Mutation Tyr-17→Ala→Gly on Helix Formation by Further Mutations of Asp-12→Ala and Asp-12→Gly. To test whether the mutation of Tyr-17→Gly has indeed altered the extent of formation of much of the helix or the results are just the consequence of a local disorganisation of structure at position 17, we constructed a third generation of mutants by mutating Asp-12 to Ala and Gly in the Tyr→Gly-17 mutant (i.e., the double mutants Asp-12→Ala/Tyr-17→Gly and Asp-12→Gly/Tyr-17→Gly). It is seen in Table 7 that whereas the Ala/Gly probe in wild-type protein shows that position 12 is nearly fully folded in the transition state ($\phi_{\ddagger\text{-}U}$ = 0.9) and highly folded in the intermediate in water ($\phi_{I\text{-}U}$ = 0.9), the Ala/Gly probe in the Tyr-17→Gly mutant shows that the transition state is less folded ($\phi_{\ddagger\text{-}U}$ = 0.2−0.4) and the intermediate is unfolded at position 12 ($\phi_{I\text{-}U}$ = 0). Thus, radical destabilization of the helix at position 17 close to the C-terminus retards not only the folding of this turn but also the middle turn: the mutation Tyr-17→Gly leads to the complete formation of helix$_1$ occurring after the major transition state.

There is not a linear relationship between a fractional value of ϕ and the degree of formation of structure; a fractional value just means that there is partial formation of structure (Matouschek et al., 1989; Fersht et al., 1992; Fersht, 1995). However, an increase in $\phi_{folding}$ certainly does mean that there is greater formation of structure, and a decrease in ϕ that there is a decrease in extent of formation [cf. Fersht (1988)].

Paradox of Simultaneous Hammond and Anti-Hammond Effects. According to the Hammond Postulate, the structure of the transition state for protein unfolding may become closer to the structure of the folded state as the protein becomes destabilized on mutation. Hammond behavior is

Table 6: Changes in Free Energies on Mutation of Asp-12 and Tyr-17[a]

mutation	$\Delta\Delta G_{U\text{-}F}$	$\Delta\Delta G_{I\text{-}F}$	$\Delta\Delta G_{\ddagger\text{-}F}$ H_2O	4 M urea	7.25 M urea
Asp-12→Ala	0.31 ± 0.06	0.26 ± 0.07	0.37 ± 0.07	0.27 ± 0.03	0.18 ± 0.01
Asp-12→Gly	1.19 ± 0.06	0.53 ± 0.03	0.46 ± 0.03	0.35 ± 0.02	0.26 ± 0.01
Tyr-17→Ala	2.05 ± 0.06	1.23 ± 0.06	1.00 ± 0.06	0.69 ± 0.03	0.45 ± 0.01
Tyr-17→Gly	4.07 ± 0.07	2.71 ± 0.03	2.35 ± 0.03	1.77 ± 0.02	1.30 ± 0.01
Asp-12(Ala→Gly)	0.88 ± 0.06	0.27 ± 0.13	0.09 ± 0.10	0.08 ± 0.03	0.08 ± 0.01
Tyr-17(Ala→Gly)	2.02 ± 0.06	1.48 ± 0.09	1.35 ± 0.06	1.08 ± 0.03	0.86 ± 0.01
Asp-12(Ala→Gly)/Gly-17	0.80 ± 0.04	0.81 ± 0.05	0.63 ± 0.05	0.55 ± 0.02	0.49 ± 0.02

[a] All energies are in kcal mol^{-1}.

Table 7: ϕ Values for the Mutation of Ala→Gly at Position 12 in Wild-Type Protein and at Position 12 in the Mutant Tyr-17→Gly

position	$\phi_{I\text{-}U}$ H_2O	$\phi_{\ddagger\text{-}U}$ H_2O	4 M urea	7.25 M urea
Asp-12(Ala→Gly)	0.69 ± 0.15	0.90 ± 0.11	0.91 ± 0.04	0.91 ± 0.01
Asp-12(Ala→Gly):Gly-17	−0.01 ± 0.08	0.21 ± 0.07	0.31 ± 0.04	0.39 ± 0.04

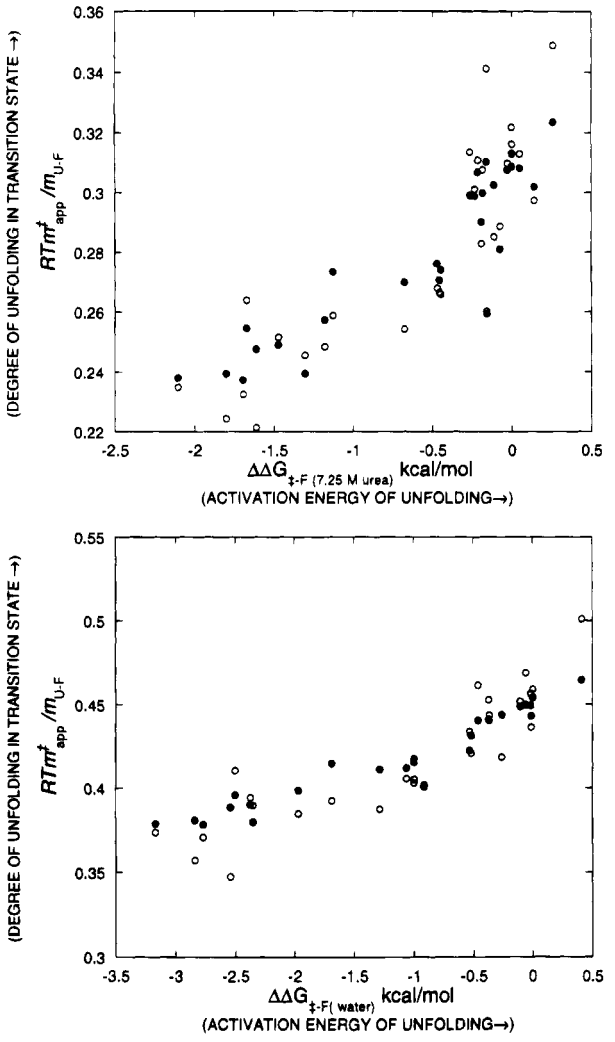

FIGURE 2: Demonstration of Hammond behavior for the response of the position of the transition state for unfolding versus the change in activation energy for mutations in the major helix$_1$ of barnase. (Top) the observed values of $RTm^{\ddagger}_{app}$ measured in 7.25 M urea divided by either the individually measured values of m_{U-F} for each mutant (open circles) or the average value of m_{U-F} (= 1.96, closed circles) are plotted against the change in free energy of activation relative to wild type in 7.25 M urea. (Bottom) As top panel, but values calculated for water, extrapolating to the absence of urea. Data are for wild type, DA8, DA12, DG12, YA13, QI15, TR16, YA17, HK18, HQ18, TA16, TG16, TS16, YA17, YG17, HA18, HG18, DA12/YG17, DG12/YG17, YA13/TS16, YA13/YA17, TS16/YA17, and YA13/TS16/YA17 (Table 1 and unpublished data, this laboratory). [It should be noted that it is not necessary to divide $RTm^{\ddagger}_{app}$ by m_{U-F} to demonstrate the Hammond behavior but that a plot of $m^{\ddagger}_{app}$ *versus* $\Delta\Delta G_{\ddagger-F}$ is adequate since m_{U-F} is close to being constant (Serrano et al., 1992). The value of $m^{\ddagger}_{app}$ is obtained with high precision (typically ± 1%), but individual values of m_{U-F} do have high errors. The use of the average value of m_{U-F} of 1.96 is preferable because the scatter in m_{U-F} is large. We have included the plots using the individually measured values of m_{U-F} to show that the changes in $RTm^{\ddagger}_{app}/m_{U-F}$ are so high that they are not obscured by scatter in the individual values of m_{U-F}.]

clearly demonstrated in Figure 2: plots of the degree of unfolding against changes in activation energy show that as the activation energy for unfolding decreases, the transition state moves closer in structure to that of the folded state, the data shown reinforcing the earlier demonstration of this principle (Matouschek & Fersht, 1993). It is seen in Table 1 that there are significant decreases in $m^{\ddagger}_{app}$ for all mutants from 0→4→7.25 M urea. That is, the structure of the transition state of unfolding moves closer toward that of the folded state with increasing [urea]. This is also entirely in accord with the Hammond postulate: increasing [urea] destabilizes the folded state with respect to the unfolded state, and so there is movement of the transition state toward the destabilized state.

The values of $m^{\ddagger}$ and $m^{\ddagger}_{app}$ for the mutants Tyr-17→Ala and Tyr-17→Gly and hence the ratios $RTm^{\ddagger}_{app}/m_{U-F}$ are the lowest in the Table 1. Thus, the overall transition state for the folding of barnase with Tyr-17 mutated to Ala or Gly occurs later on the reaction pathway than normal and is closer to the folded structure, in accord with Hammond. But this leads to an apparent paradox: the gross structure of the transition state follows Hammond postulate behavior by moving toward the structure of the folded state as it is destabilized by mutations in α-helix$_1$, but the ϕ-value analysis (Table 3) shows that helix$_1$ becomes less folded on destabilization; that is anti-Hammond behavior—movement of the transition state away from the folded state as a result of destabilizing the protein!

The apparent paradox may be resolved by arguments analogous to those conventionally used on classical structure–reactivity relationships, as reviewed by Jencks (1985). In Figure 3 (left) is sketched [cf. Jencks (1985)] the causes of Hammond and Anti-Hammond behavior. Hammond behavior results from movement parallel to the reaction coordinate (Figure 3, left, top). Here, the transition state is at a maximum in the trajectory. If one of the states is destabilized, then the transition state moves toward it. Energy surfaces are multidimensional, however, and so movement can occur in other directions. The energy surface around a transition state is, in three dimensions, saddle shaped with the energy going through a maximum along the reaction coordinate and being in a minimum perpendicular to it. For example, Figure 3, left, bottom, is a slice through a three-dimensional surface at the transition state saddle point but perpendicular to the reaction coordinate. If one of the states is destabilized perpendicular to the reaction coordinate, then the transition state moves away from it—anti-Hammond behavior. On the right-hand side of Figure 3 are diagrams of the average degree of folding against ϕ for the equivalent processes on the left. These correlation diagrams were inspired by, but are different from, the More O'Ferrall-Jencks plots (Jencks, 1985). The top represents the effect of a destabilizing mutation parallel to the reaction coordinate. The transition state moves toward the structure of the folded state. The bottom represents the effect of a destabilizing mutation perpendicular to the transition state. A destabilizing mutation in the helix causes a movement toward the unfolded state. The net result of the destabilizing mutation is that the gross structure of the transition state moves toward the structure of the folded state but the individual element of structure becomes less folded.

Single versus Parallel Pathways. We know that the transition state moves on progressive destabilization of α-helix$_1$ from a structure in which the helix is mainly formed to one in which the helix is mainly unformed. The movement from one transition state to another could be progressive as for the explanation by the classical Hammond and anti-Hammond model, or it could be a jump from one mechanism to another (Figure 4). In other words, there is an alternative explanation for the structure–reactivity behavior that involves the existence of parallel pathways, one of which has the helix predominantly formed in the transition

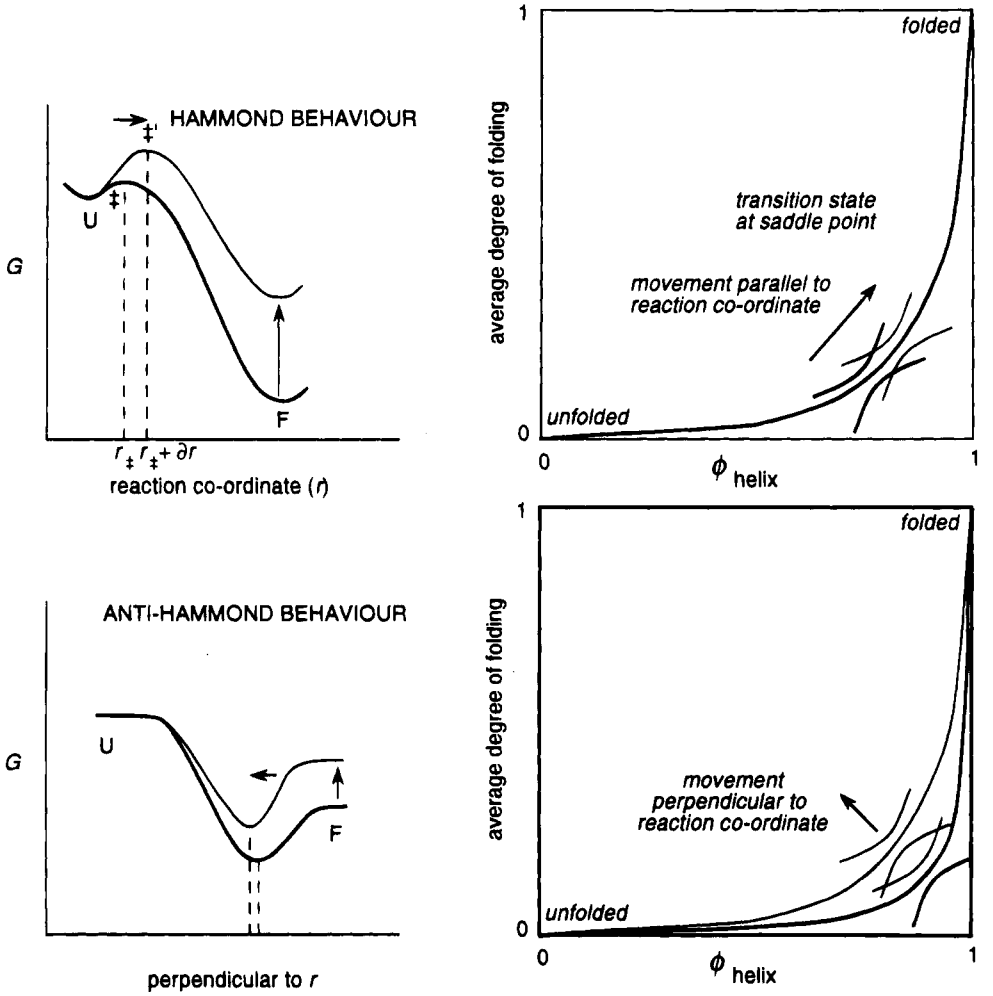

FIGURE 3: Hammond and anti-Hammond behavior for the folding of a protein. (Left top) Free energy profile along the reaction coordinate. There is conventional Hammond behavior as the transition state moves closer to the folded state along the reaction coordinate as a result of destabilization of that state. (Left bottom) Cross-section of the energy profile perpendicular to the reaction coordinate at the transition state. There is anti-Hammond behavior as the transition state moves closer to the unfolded state in a direction perpendicular to the reaction coordinate as a result of destabilization of the folded state [see Jencks (1985)]. (Right) Correlation diagrams of the average degree of folding for the whole protein and ϕ, the degree of formation of the helix, in the transition state for folding. The average degree of folding is measured experimentally from $RTm^{\ddagger}_{app}/m_{U\text{-}F}$ and is 0 in the unfolded state and 1 in the folded state. (Top right) Average degree of folding in the transition state increases as the transition state moves along the reaction coordinate closer to the folded protein as the protein is destabilized by a mutation. (Bottom right) Concurrent with the movement of the transition state along the reaction coordinate in the direction of the folded protein as the protein is destabilized by a mutation, there is a movement perpendicular to the reaction coordinate by the anti-Hammond effect that leads to the helix becoming less folded (ϕ decreases).

state (pathway 1 with $\phi_{\ddagger\text{-}U} \sim 1$ and a lower value of $m_{\ddagger\text{-}F}$) and the other in which the helix is predominantly unformed (pathway 2 with $\phi_{\ddagger\text{-}U} < 1$, and a higher value of $m_{\ddagger\text{-}F}$). There is, of course, a continuum of possibilities for the movement, varying from the smooth movement of the structure of the transition state, to a series of discretely separated states to the extreme of two distinctly different states.

We can analyze the behavior of $m_{\ddagger\text{-}F}$ with mutation of the parallel-pathway model, using for simplicity a fixed concentration of urea (e.g., 7.25 M) as a reference point. The rate constant for unfolding of wild-type protein is given by

$$k_u = k_u^1 + k_u^2 \tag{9}$$

where k_u^1 is the rate constant for pathway 1 and k_u^2 for pathway 2. It may be shown formally that the apparent value of $m_{\ddagger\text{-}F}$ ($m^{\ddagger}_{app}$) at the particular concentration of urea is given by

$$m^{\ddagger}_{app} = (m_1 k_u^1 + m_2 k_u^2)/(k_u^1 + k_u^2) \tag{10}$$

where $m_1 = RT(\partial \ln k_u^1/\partial[\text{urea}])[\text{urea}] = 7.25$ M, etc.

Suppose a mutation is made that lowers the activation energy of unfolding by $\Delta\Delta G^1_{\ddagger\text{-}F}$ for pathway 1 and by $\Delta\Delta G^2_{\ddagger\text{-}F}$ for pathway 2 at the particular concentration of urea. Then

$$k_{u,mut} = k_u^1 \exp(\Delta\Delta G^1_{\ddagger\text{-}F}/RT) + k_u^2 \exp(\Delta\Delta G^2_{\ddagger\text{-}F}/RT) \tag{11}$$

Then

$$m^{\ddagger}_{app} = \frac{m_1 k_u^1 \exp(\Delta\Delta G^1_{\ddagger\text{-}F}/RT) + m_2 k_u^2 \exp(\Delta\Delta G^2_{\ddagger\text{-}F}/RT)}{k_u^1 \exp(\Delta\Delta G^1_{\ddagger\text{-}F}/RT) + k_u^2 \exp(\Delta\Delta G^2_{\ddagger\text{-}F}/RT)} \tag{12}$$

or

$$m^{\ddagger}_{app} = \frac{m_1 k_u^1 \exp(\Delta\Delta G^1_{\ddagger\text{-}F}/RT) + m_2 k_u^2 \exp(\Delta\Delta G^2_{\ddagger\text{-}F}/RT)}{k_{u,mut}} \tag{13}$$

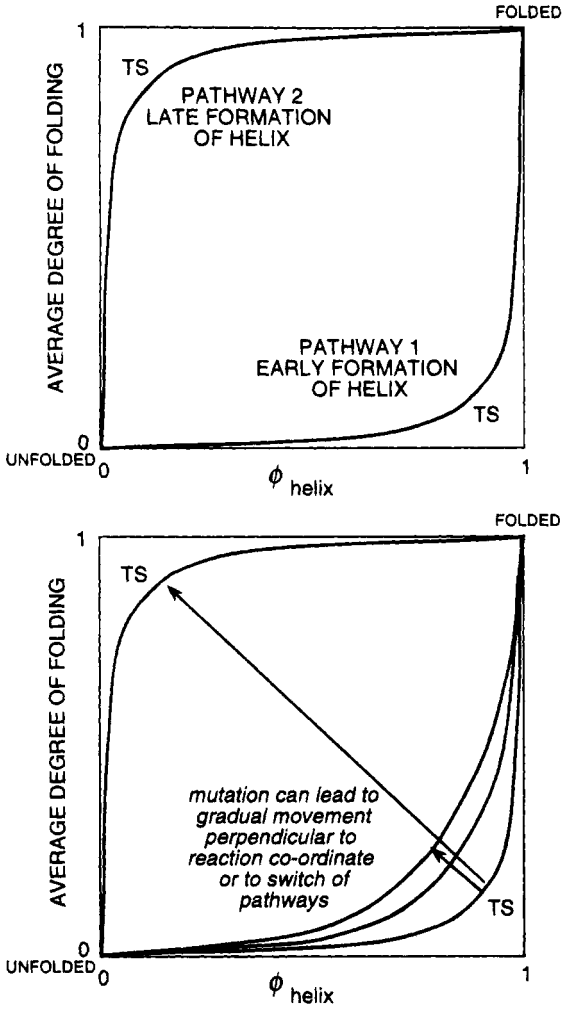

FIGURE 4: Alternative mechanism for Hammond and anti-Hammond behavior arising from the potential energy surface (bottom left) having a double well; that is, there are parallel pathways of unfolding (e.g., one where the helix is nearly fully folded in the transition state and another where it is more unfolded). The destabilization of the folded helix would favor the pathway involving the unfolded helix. (It is shown in the text that this mechanism is not consistent with the data for $helix_1$ of barnase.)

It is inferred from eqs 12 and 13 that a plot of $m^{\ddagger}_{app}$ versus $1/k_{u,mut}$ should be sigmoid with $m^{\ddagger}_{app}$ tending to m_1 for the most stable proteins and to m_2 for the most destabilized. Such a plot for all known mutants from residues 8−18 in $helix_1$ of barnase (Figure 5) shows a progressive rather than a sigmoid variation over a range of 50-fold in k_u. This does not appear consistent with a switch between two extreme pathways but is more consistent with a classical smooth progression of the transition state or the movement along a series of changing transition states.

Why Are Structure−Reactivity Effects Observed in $Helix_1$? So far, we have detected significant Hammond and anti-Hammond behavior on mutation of the major $helix_1$ and a weak Hammond effect in the hydrophobic core (Matouschek & Fersht, 1993). The plot (Figure 2) of $RTm^{\ddagger}_{app}/m_{U\text{-}F}$ versus $\Delta\Delta G_{\ddagger\text{-}F}$ at 7.25 M urea has an intercept at 0.3 for wild type, a slope of 0.036 ± 0.03, and correlation coefficient 0.9. A similar plot for all 110 mutants of barnase on which measurements of $m_{\ddagger\text{-}F}/m_{U\text{-}F}$ have been made in this laboratory does show an overall weaker Hammond behavior with intercept 0.30, slope 0.013 ± 0.001, and correlation coefficient 0.54. But these are strongly biased by the data on

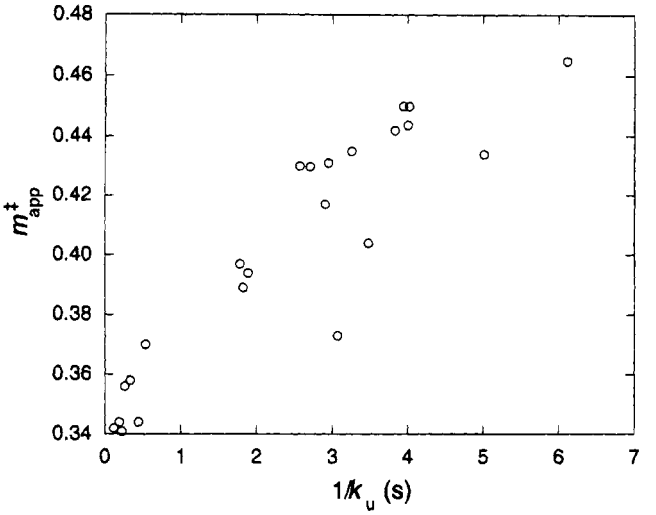

FIGURE 5: Plot of $m^{\ddagger}_{app}$ versus $1/k_{u,mut}$. Data are for wild type, DA8, DA12, DG12, YA13, QI15, TR16, YA17, HK18, HQ18, TA16, TG16, TS16, YA17, YG17, HA18, HG18, DA12/YG17, DG12/YG17, YA13/TS16, YA13/YA17, TS16/YA17, and YA13/TS16/YA17 (Table 1 and unpublished data, this laboratory).

the helix and the core, and the data elsewhere do not show significant Hammond effects. $Helix_1$ is formed as a highly cooperative unit from residues 10−18 in the folding of barnase as shown by multimutant cycle (Horovitz & Fersht, 1992; Horovitz et al., 1991) and H/^{2}H-exchange analysis (Bycroft et al., 1990; Matouschek et al., 1992b). Further, NMR studies of the synthetic peptide corresponding to residues 1−22 of barnase show that it is helical from residues 6−18 in 35% trifluoroethanol solution, but the mutation Tyr-17→Ala destabilizes the helix so that it is seen from residues 7−14 only (Kippen et al., 1994). The hydrophobic core of the protein is formed by the docking of $helix_1$ on the sheet, and, further, the docking is part of the rate determining step (Bycroft et al., 1990; Fersht, 1993). The Hammond−anti-Hammond behavior must be caused by the cooperative nature of the formation of the helix being coupled to the formation to the formation of the protein as a whole and the whole process thus being sensitive to mutation.

The movement of the transition state on mutation could possibly obscure the interpretation of ϕ values since it is seen that the mutation Ala-17→Gly gives an artifactually low value. This clearly does not happen in general since we see many values of ϕ that are very close to 1, both in $helix_1$ and in other regions of the protein (Matouschek & Fersht, 1993), showing that mutations that are sensibly designed not to perturb the energies too much do not lower ϕ values in regions of the protein whose formation are not highly cooperative. Examination of values of $m^{\ddagger}_{app}$ provides the diagnostic clue for movement for movement of the transition state. As noted in the last section, most mutations we have studied cause very small changes in $m^{\ddagger}_{app}$. In general, the ϕ-value analysis does appear to work satisfactorily, but care must be taken when analyzing radical mutations.

Implication for Protein Folding. As $helix_1$ is destabilized, its structure begins to be formed less and less in the major transition state. Conversely, the overall structure of the transition state, and hence the structure of the rest of the protein, becomes formed more completely. The structure of the transition state thus moves around the potential energy surface. The average degree of unfolding in the transition

state, given by $RTm^{\ddagger}_{app}/m_{U\text{-}F}$, changes from 0.45 to 0.38 in water, and 0.32 to 0.24 in 7.25 M urea, as the helix is destabilized. These changes show a significant change in structure. The plots in Figures 2 and 5 show a progressive Hammond response to changes in activation energy of unfolding. The progressiveness implies that there is not a leap from one pathway to another, but rather a gradual movement around the multidimensional energy surface as it is perturbed. The energy levels must be very close together and respond to a mutation that has a coordinated effect on structure.

The data also show that $helix_1$, despite forming early in wild-type protein and having characteristics that could be consistent with an initiation site, can be radically altered until it forms late in the folding process. The value of $\phi_{I\text{-}U}$ at position 12 in the Tyr-17→Gly mutant is 0 (Table 7), showing that the helix is not formed at the central position in the folding intermediate of the destabilized mutant. Yet that mutant still folds at a fast rate, showing that $helix_1$ is not an essential initiation site. This is consistent with observations on the folding and complementation of two fragments of barnase; residues 1−22 that contain the α-helix and residues 23−110 that contain the remaining α-helices and the β-sheet (Kippen et al., 1994). The large fragment is able to fold partially to a catalytically competent state, and the full native-like structure is attained on addition of fragment 1−22.

Sali et al. (1994) have recently presented calculations on the pathway of protein folding suggesting that the main criteria for folding pathways to exist is that there is a well-defined energy minimum for the folded state and a lack of kinetic barriers for its approach. The finding that the folding pathway for barnase has a certain plasticity is not inconsistent with those calculations.

ACKNOWLEDGMENT

J.M.M. is a recipient of a Packer Foundation scholarship.

REFERENCES

Bycroft, M., Matouschek, A., Kellis, J. T., Jr, Serrano, L., & Fersht, A. R. (1990) *Nature 346*, 488−490.
Clarke, J., & Fersht, A. R. (1993) *Biochemistry 32*, 4322−4329.
Fersht, A. R. (1988) *Biochemistry 27*, 1577−1580.
Fersht, A. R. (1993) *FEBS Lett. 325*, 5−16.
Fersht, A. R. (1995) *Curr. Opin. Biol. 5*, 79−84.
Fersht, A. R., Matouschek, A., & Serrano, L. (1992) *J. Mol. Biol. 224*, 771−782.
Hammond, G. S. (1955) *J. Am. Chem. Soc. 77*, 334−338.
Horovitz, A., & Fersht, A. R. (1992) *J. Mol. Biol. 224*, 733−740.
Horovitz, A., Serrano, L., & Fersht, A. R. (1991) *J. Mol. Biol. 219*, 5−9.
Horovitz, A., Matthews, J. M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 560−568.
Horovitz, A., Serrano, L., Avron, B., Bycroft, M., & Fersht, A. R. (1990) *J. Mol. Biol. 216*, 1031−1044.
Jencks, W. P. (1985) *Chem. Rev. 85*, 511−527.
Johnson, C. J., & Fersht, A. R. (1995) *Biochemistry* (in press).
Kellis, J. T., Jr., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914−4922.
Kippen, A. D., Sancho, J., & Fersht, A. R. (1994) *Biochemistry 33*, 3778−3786.
Kippen, A. D., Arcus, V. L., & Fersht, A. R. (1994) *Biochemistry 33*, 10013−10021.
Matouschek, A., & Fersht, A. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7814−7818.
Matouschek, A., Kellis, J. T., Jr., Serrano, L., & Fersht, A. R. (1989) *Nature 342*, 122−126.
Matouschek, A., Kellis, J. T., Jr., Serrano, L., Bycroft, M., & Fersht, A. R. (1990) *Nature 346*, 440−445.
Matouschek, A., Serrano, L., & Fersht, A. R. (1992a) *J. Mol. Biol. 224*, 819−835.
Matouschek, A., Serrano, L., Meiering, E. M., Bycroft, M., & Fersht, A. R. (1992b) *J. Mol. Biol. 224*, 837−845.
Matouschek, A., Matthews, J. M., Johnson, C. M., & Fersht, A. R. (1994) *Protein Eng. 7*, 1089−1095.
Meiering, E. M., Serrano, L., & Fersht, A. R. (1992) *J. Mol. Biol. 225*, 585−589.
Sali, A., Shakhnovich, E., & Karplus, M. (1994) *J. Mol. Biol. 235*, 1614−1636.
Serrano, L., & Fersht, A. R. (1989) *Nature 342*, 296−299.
Serrano, L., Horovitz, A., Avron, B., Bycroft, M., & Fersht, A. R. (1990) *Biochemistry 29*, 9343−9352.
Serrano, L., Bycroft, M., & Fersht, A. R. (1991) *J. Mol. Biol. 218*, 465−475.
Serrano, L., Matouschek, A., & Fersht, A. R. (1992a) *J. Mol. Biol. 224*, 805−818.
Serrano, L., Matouschek, A., & Fersht, A. R. (1992b) *J. Mol. Biol. 224*, 847−859.
Serrano, L., Neira, J. L., Sancho, J., & Fersht, A. R. (1992c) *Nature 356*, 453−455.
Serrano, L., Sancho, J., Hirshberg, M., & Fersht, A. R. (1992d) *J. Mol. Biol. 227*, 544−559.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1−95.
Vuilleumier, S., Sancho, J., Loewenthal, R., & Fersht, A. R. (1993) *Biochemistry 32*, 10303−10313.

BI9500436