Identification of Proline Residues Responsible for the Slow Folding Kinetics in Pectate Lyase C by Mutagenesis[†,‡]

Douglas E. Kamen[§] and Robert W. Woody*

Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523

Received July 19, 2001; Revised Manuscript Received December 4, 2001

ABSTRACT: The folding mechanism of pectate lyase C (pelC) involves two slow phases that have been attributed to proline isomerization. To have a more detailed and complete understanding of the folding mechanism, experiments have been carried out to identify the prolyl−peptide bonds responsible for the slow kinetics. Site-directed mutagenesis has been used to mutate each of the prolines in pelC to alanine or valine. It has been determined that isomerization of the Leu219−Pro220 peptide bond is responsible for the slowest folding phase observed. The mutant P220A shows kinetic behavior that is identical to the wild-type protein except that the 46-s phase is eliminated. The Leu219−Pro220 peptide bond is *cis* in the native enzyme. An analysis of the free energy of unfolding of this mutant indicates that the mutation destabilizes the protein by about 4 kcal/mol. However, it appears that the major refolding pathways are unaltered. Further mutations were carried out in order to assign the peptide bond responsible for the 21-s folding phase in pelC. Mutation of the remaining 11 prolines, which are *trans* in the native enzyme, resulted in no significant changes in the kinetic folding behavior. The conclusion from these experiments is that the 21-s phase involves isomerization of more than one prolyl−peptide bond with similar activation energies.

It is well-established that slow phases in protein folding exist. One of the most important developments in understanding slow kinetics of protein folding came in 1975 (*1*). Brandts et al. (*1*) proposed that the slow step in folding is due to the *cis*−*trans* isomerization of prolyl−peptide bonds in the unfolded state. X-Pro peptide bonds (X is any amino acid) can populate the *cis* conformation much more readily than can any other peptide bond. The highly unfavorable *cis:trans* ratio in non-prolyl peptide bonds is due in part to the *trans* isomer having greater conformational entropy relative to the *cis* isomer and, in the *cis* isomer, steric clashing of the side chain with the preceding amino acid. The bonding of C_δ to the backbone nitrogen in proline reduces the advantages of *trans* over *cis*. In an unordered polypeptide the *trans:cis* ratio of X-Pro peptide bonds is about 70:30 (*1*). Many proteins contain prolyl−peptide bonds that are *cis* in the native conformation. Conformational constraints in the folded protein can overcome the less favorable interactions and stabilize the *cis* bond. This is often important in forming the correct architecture of the enzyme active site for binding substrate or properly positioning a catalytically active group.

In the folding kinetics of pelC,[1] a total of four phases are observed (see the previous paper in this issue, ref *2*). Two of these events occur on the second or faster time scale and are attributed to the formation of secondary structure and ordering of side chains. Two phases show kinetics occurring on the order of tens of seconds. Proline isomerization in model peptides occurs in the time range of 10−100 s near room temperature (*1*). Evidence was presented previously (*2*) that indicates that both of these slow phases are due to proline isomerization. However, although these experiments are conclusive, they give no clue as to which specific prolines are responsible for the observed slow kinetics. The Leu219−Pro220 peptide bond is in the *cis* conformation as indicated by the X-ray structure of pelC (*3*). This proline is conserved among several pectate lyases (*4*). It serves a critical role in that it properly positions Arg218, an important group in catalysis. In addition to Pro220, pelC contains 11 more prolines, at positions 51, 79, 124, 139, 159, 263, 286, 302, 316, 324, and 335. All of these prolines are involved in *trans* peptide bonds. Figure 1 shows a model of pelC (*5*) highlighting all of the proline residues.

To have a more complete model for the folding mechanism of pelC, one needs to characterize the slow kinetics and to determine which prolines are responsible. Three methods have been used to identify "essential" prolines. Lin and Brandts developed the technique of isomer-specific proteolysis (*6*, *7*). This utilizes the fact that certain proteases only recognize *trans* peptide bonds and therefore cannot hydrolyze prolyl−peptide bonds in the *cis* configuration.

[†] Supported by USPHS Grant GM-22994.

[‡] From a dissertation submitted to the Academic Faculty of Colorado State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

* To whom correspondence should be addressed. E-mail: rww@lamar.colostate.edu. Fax: (970) 491-0494.

[§] Current address: Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY 10461.

[1] Abbreviations: CD, circular dichroism; CM, carboxymethyl; $D_{1/2}$, denaturant concentration at the transition midpoint; DEAE, diethylaminoethyl; gdn-HCl, guanidine-hydrochloride; LB, Luria-Bertani; MES, 2-[*N*-morpholino]ethanesulfonic acid; MOPS, *N*-morpholino propanesulfonic acid; pelC, pectate lyase C; PPI, peptidyl-prolyl isomerase; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; WT, wild-type. Standard abbreviations are used for the amino acids and nucleic acid bases.

10.1021/bi0115131 CCC: $22.00
Published on Web 03/15/2002

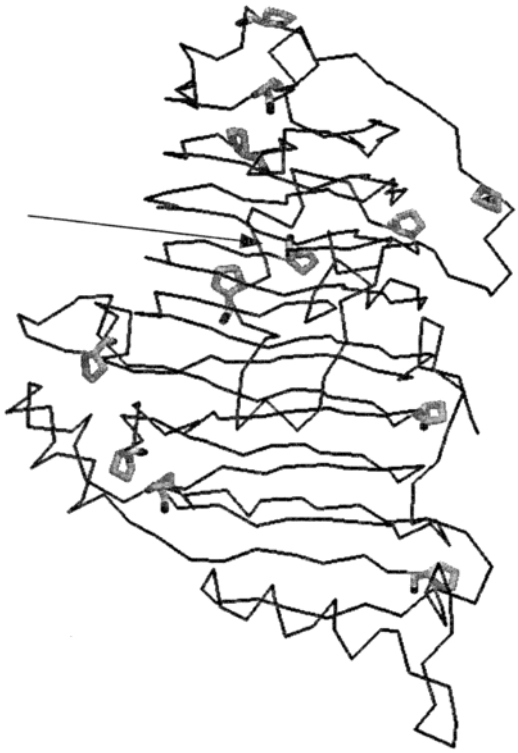

FIGURE 1: A structural model of pelC (*5*) showing all of the proline residues. Pro220, which is *cis* in the native enzyme, is indicated by an arrow. The figure was generated using RasMol (*30*).

Peptide fragments are analyzed by peptide sequencing and slowly isomerizing bonds can be identified. This method has yielded controversial results (*8*, *9*) and would be very complicated to use with pelC, considering that there are 12 prolines. NMR has also been used to measure the *cis:trans* ratio in heat-unfolded RNase A (*8*). Prolines that had a significant population of the incorrect native isomer were identified. The percentage measured by NMR was in good agreement with the percentage of incorrect isomer measured by kinetic experiments. The two prolines identified were proposed to be the ones responsible for the observed slow kinetics. This method was effective, but would be very difficult to apply to a protein the size of pelC because it requires assignment of resonances in the NMR spectrum. Advances in the field of protein engineering in recent years have led to numerous insights into protein structure and function. Several groups have used site-directed mutagenesis to study slow kinetics in folding. Schultz et al. (*9*) has demonstrated conclusively that the slow kinetics in RNase A are due to the two *cis* prolines, 93 and 114. Kelley and Richards (*10*) showed that replacement of Pro76 with Ala in thioredoxin eliminates the slow phase. In both of these cases, the suspect prolines were *cis* in the native structure. Schindler et al. (*11*) have investigated the role of *trans* prolines in the folding of RNase T1. They concluded that *trans* prolines do not noticeably affect folding. More recently, Eyles and Gierasch (*12*) investigated the *trans* proline contribution to the folding of CRABP 1. They found that mutation of Pro85 to Ala resulted in elimination of the slowest folding phase.

The mutagenesis approach seemed to be the best way to characterize the slow folding kinetics in pelC. In all, 10 Pro→Ala (including P220A) and two Pro→Val mutants were constructed, and the folding of these mutants was characterized. P79A and P124A were constructed but several attempts to express these mutants failed. Instead, Pro79 and Pro124 were mutated to valine. Pro220 isomerization is responsible for the slowest kinetic phase of pelC folding. The intermediate slow phase may result from multiple proline isomerizations with similar activation energies. Finally, the intermediate slow phase isomerization is strongly coupled to the formation of an intermediate on the folding pathway.

MATERIALS AND METHODS

Site-Directed Mutagenesis. The Quickchange Site-Directed Mutagenesis kit (Stratagene Inc, La Jolla, CA) was used to construct all mutants. Plasmid pPEL410 containing the gene for wild-type pelC was used as the template in all cases except for P139A (this utilized the P220A plasmid and therefore was a double mutation). The mutagenic primers used are shown in Table 1. All primers were purified by the supplier (Integrated DNA Technologies, Coralville, IA) by polyacrylamide gel electrophoresis (PAGE). In most cases, the prolines were mutated to alanine, and where possible, the optimum codon preferences were used. In two cases, Pro79 and 124, the prolines were mutated to valine.

Mutated plasmids were transformed into *Escherichia coli* XL1-Blue supercompetent cells (Strategene Inc, La Jolla, CA) and grown for about 16 h at 37 °C on LB agar plates supplemented with 50 μg/mL ampicillin. Plasmid DNA was isolated and purified using a Qiagen plasmid mini-prep kit (Valencia, CA). Mutations were verified by DNA sequencing. Mutant plasmids were then transformed into *E. coli* strain HB101 for protein expression. Transformation-competent cells were prepared by the $CaCl_2$ method and transformations were carried out by methods described by Maniatis (*13*). Transformants were grown at 37 °C on LB agar plates supplemented with 50 μg/mL ampicillin for 13 h to limit the number of background colonies.

Purification of Mutant pelC Proteins. The purification of all mutant proteins was similar to the wild-type pelC purification (*14*), with some modification. Growth of cells and formation of spheroplasts were identical. It was found that protein from the periplasmic fraction of some mutants tended to aggregate on the CM Biogel column used for wild-type pelC and caused the column to clog. Because pelC is the only protein in the *E. coli* periplasm that appears to bind to the negatively charged carboxymethyl resin of the CM Biogel, it was determined whether it is the only protein to pass through the corresponding DEAE Biogel. In fact, this procedure showed results comparable to the wild-type preparation. The main modification to purify the mutant proteins was to pass the dialyzed periplasmic fraction over a DEAE Biogel-A column equilibrated to pH 8 in 5 mM Tris buffer. The column flow-through that showed absorbance at 280 nm was retained. Proteins were assayed for purity by SDS−PAGE. Proteins that migrated on the gel at the same rate as wild-type pelC were assumed to be the mutants. In general, this procedure provided protein that was of comparable purity to the wild-type protein, i.e., >98% purity, as demonstrated by a single band on SDS−PAGE. In very rare cases, the flow-through was passed through a CM column and eluted with the procedure for the wild-type protein. With the majority of contaminating proteins re-

Table 1: Mutagenesis Primer Sequences. Bold Sequences Indicate the Altered Bases

mutation	coding strand primer
P51A	5′-GGCGGCGCCTAT**G**CGCTGGTAATCACC-3′
P79V	5′-GTGGAGCAAAGAC**GTT**CGTGGCGTAG-3′
P124V	5′-CGTATCGGCTACCTG**GTT**GGCGCTAAAGATGGC-3′
P220A/P139A[a]	5′-CCGCGTCGACGATTCG**G**CGAATGTCTGGGTTGACC-3′
P159A	5′-GTGCGACGGCACA**G**CGGACAACGACACC-3′
P220A	5′-CGTTAACGCCCGTCTG**G**CGTTGCAACGTGGTGG-3′
P263A	5′-GGTTCGAGAAGGCGATAAAC**G**CGGTAACGTCCCGTTATGACGGC-3′
P286A	5′-GCAACAACATCACCAAA**G**CGGCCGACTTCTCTACC-3′
P302A	5′-CGGCCGACACCAAG**G**CTTATGTGAATGCCG-3′
P316A	5′-CGGCCACCTTC**G**CGACCGTGGCCTAC-3′
P324A	5′-GCCTACAACTACA**G**CGCGGTCAGCGCACAATGC-3′
P335A	5′-GAAGGACAAACTG**G**CTGGCTATGCCGGC-3′

[a] The sequence shown for this double mutant is that which leads to the change of Pro→Ala at position 139. The sequence in the vicinity coding for residue 220 is the same as that for P220A.

moved, mutant pelC proteins did not appear to aggregate. Mutant protein yields ranged from about 4−50 mg/L of growth media.

Characterization of Proline Mutants. All experiments on pelC mutants were carried out identically to those on the wild-type pelC (*2*, *15*), with the following exceptions. Some refolding experiments were carried out by pH-jump. P220A rapidly unfolds at pH 2 in 25 mM glycine-HCl buffer. Refolding was initiated by dilution (generally 10-fold) into 100 mM MES/50 mM NaCl at pH 6.1. The resulting solution had a pH of 6.0−6.1. The activation enthalpy of P220A was measured in a manner similar to wild-type pelC, except the pH-jump method was used to initiate refolding. This avoided performing the experiment at several gdn-HCl concentrations and extrapolating to 0 M denaturant. For equilibrium denaturation experiments, 25 mM *N*-morpholino propanesulfonic acid (MOPS) buffer was used at pH 7.0.

RESULTS

Expression and Purification of pelC Mutants. All 12 mutant constructs were successfully expressed and purified. In all cases, a two-liter culture of bacteria yielded between 8 and 100 mg of pure protein. For expression tests, 100 mL of LB cultures were grown and lysed according to the standard protocol (*14*). For P220A, P139A/P220A, and P263A this method yielded good expression and sufficient quantities for some initial experiments. For the rest of the mutants, expression in 100-ml cultures was very low and sometimes hardly detectable. A very interesting result occurred when the culture volume was increased to 2 L. The level of mutant protein expression compared to other *E. coli* proteins was dramatically increased. This resulted in sufficient yield of pelC mutant proteins of purity comparable to that of the wild-type preparations.

Structure and Activity of P220A. P220A is the most drastic of the Pro→Ala mutations. A proline involved in a *cis* peptide bond is changed to alanine that favors the *trans* configuration. One might expect significant perturbations in the structure and activity of this mutant. The catalytic activity of P220A was measured in a manner identical to that for wild-type pelC (*15*). The spectrophotometric assay that monitors the cleavage of polypectate over time indicated that the specific activity of wild-type pelC is 1.9×10^6 units/g. The same assay under identical conditions indicated that the specific activity of P220A was only 3.0×10^4 units/g. The mutant is about 65-fold less active than the wild-type enzyme. Such a decrease in activity could suggest a large alteration in the overall structure of pelC due to the mutation. CD and fluorescence experiments were carried out to determine the extent of structural perturbations.

The fluorescence emission spectrum of P220A at pH 7 indicates an intense band with an emission maximum at 331 nm (Figure 2a). The spectrum of wild-type pelC is also shown. The intensities of the two spectra are similar, but the mutant maximum is shifted 2 nm to the blue. The intensity is arbitrary, but both are typical of folded proteins. However, the shift in the maximum suggests that there may be a slight decrease in the solvent exposure of tryptophan side chains in the mutant.

Circular dichroism experiments were performed to assess the degree of change in the secondary structure and ordering of aromatic side chains. The far-UV CD spectrum indicates a slight loss of intensity in both the long- and short-wavelength regions (Figure 2b). This is consistent with a minor perturbation in some elements of secondary structure. An analysis of the secondary structure using CDPro is shown in Table 2. The secondary structure contents are the same as for wild-type pelC, within the accuracy of the analysis.

A comparison of the near-UV CD of wild-type pelC and P220A mutant shows no change (Figure 2c), in contrast to the small but detectable changes in the far-UV CD and fluorescence spectra. This argues against any overall changes in the geometries of the aromatic side chains or in the coupling of the side-chain transitions.

Equilibrium Denaturation of P220A. Elimination of a proline involved in a *cis* peptide bond can result in large changes to the overall stability of the protein. The free energy of unfolding for P220A in water ($\Delta G^{\circ}_{H_20}$) was measured using gdn-HCl titration in a manner identical to that for the wild-type protein (*15*). Figure 3a shows the gdn-HCl denaturation transition monitored by CD at 218 nm for wild-type pelC and P220A. A comparison of the denaturation of the mutant with that of wild-type pelC indicates that the midpoint for the mutant is significantly lower than that of the wild-type. There also appears to be a slight difference in the magnitudes of the folded and unfolded baseline. In the folded baseline region, the magnitude of the wild-type protein is greater by ∼500 deg cm^2/dmol than that of the mutant. In addition, the magnitude of the wild-type unfolded baseline is less by ∼400 deg cm^2/dmol than that of the mutant. This indicates that the overall change in CD intensity for the folding process is less for the mutant P220A. However, the overall shape of

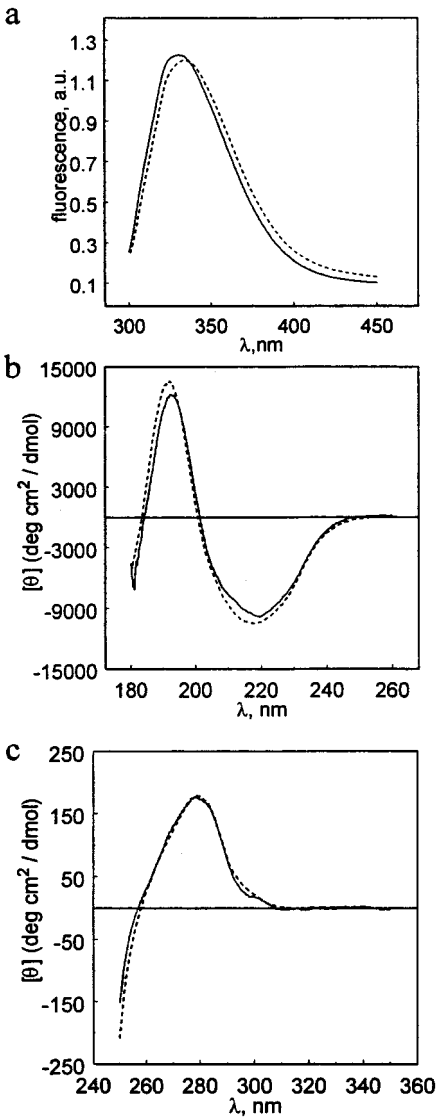

FIGURE 2: Spectroscopic analysis of P220A (−) compared to wild-type pelC (- - -) measured by fluorescence emission (a); far-UV CD (b); and near-UV CD (c). Conditions were 25 °C in water for both proteins.

Table 2: Fractions of Secondary Structure in pelC and P220A Determined by X-ray Diffraction and Circular Dichroism

	α-helix	β-sheet	turns	other
X-ray[a] WT pelC	11	27	4	59
DSSP[b] WT pelC	9	32	8	51
CDPro[c] WT pelC	24	21	22	33
CDPro[c] P220A	20	25	25	30

[a] Estimates from the X-ray diffraction structure (*3*). [b] Objective analysis of the structure from X-ray diffraction using the method of Kabsch and Sander (*26*). [c] Estimation of secondary structure from circular dichroism (*27*−*29*).

the transition curve appears to be unchanged. Table 3 shows the results of the thermodynamic analysis. $\Delta\Delta G^{\circ}_{H_2O}$ is about −4 kcal/mol but the *m* value is unchanged within experimental error. This indicates that, although there is a large loss of stability, the change in accessible surface area is not altered, and hence, the overall denaturation mechanism is probably unchanged.

To demonstrate that the two-state assumption is valid, denaturation was monitored by fluorescence emission (Figure 3b). The results are in excellent agreement with the CD

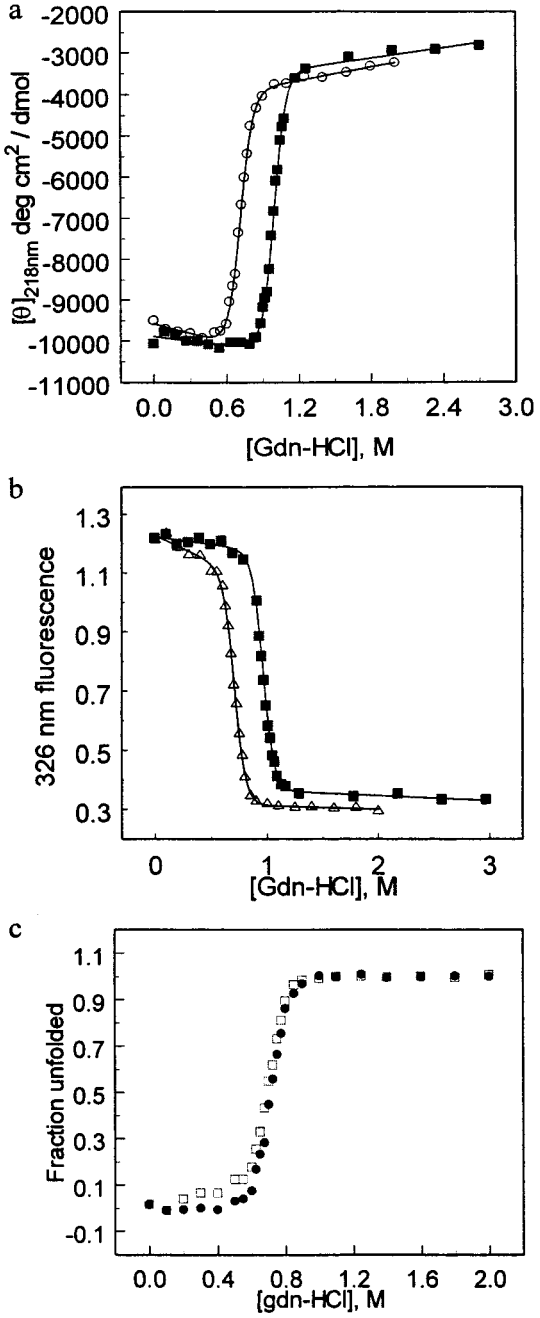

FIGURE 3: Gdn-HCl titration of P220A. (a) CD values at 218 nm for P220A (○) and WT pelC (■) fit to a two-state model. (b) Fluorescence emission values at 326 nm (excitation at 290 nm) for P220A (△) and WT pelC (■) fit to a two-state model. (c) The fraction of P220A unfolded vs gdn-HCl concentration. Filled circles (●) are from far-UV CD and open squares (□) are from fluorescence. Conditions were 25 °C, pH 6, MES 25 mM, NaCl 50 mM.

results on the mutant (Table 3). The spectroscopic data were converted to fraction of protein unfolded (Figure 3c). The curves do not overlay perfectly, however. This may be because of the significant slope of the folded baseline for the fluorescence titration, which could result from solvation effects in the unordered region around the mutation. Alternatively, perhaps deviation from two-state behavior is greater in the mutant than in the wild-type (*15*). Because the free-energy and *m* values measured by the two probes are identical within experimental error, it is assumed that the denaturation mechanism is unaffected by the mutation.

Table 3: Thermodynamic Results from gdn-HCl Denaturation at pH 7 for pelC and P220A

	218 nm CD WT pelC	218 nm CD P220A	fluorescence WT pelC	fluorescence P220A	$\Delta\Delta G$ CD	$\Delta\Delta G$ Fl
$\Delta G°$ (kcal/mol)	12.06 ± 0.59	8.28 ± 0.36	13.03 ± 0.67	8.35 ± 0.41	−3.78	−4.68
m (kcal/mol M)	12.19 ± 0.59	11.59 ± 0.49	13.04 ± 0.67	11.80 ± 0.56		
$D_{1/2}$ (M)	0.99	0.71	1.00	0.71		

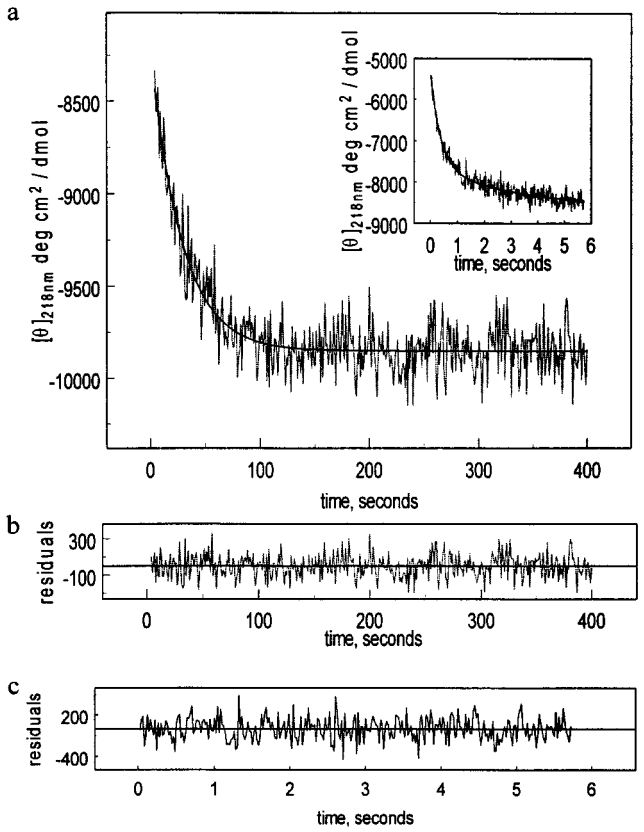

FIGURE 4: (a) Refolding kinetics of P220A in 0.3 M gdn-HCl monitored by far-UV CD at 218 nm. The solid line through the data is a single-exponential fit. The inset to panel a shows the first six seconds of the folding process measured by stopped-flow CD. Panels b and c show the residuals to the fits for the slow and fast phases, respectively. Conditions were 25 °C, pH 6, MES 25 mM, NaCl 50 mM.

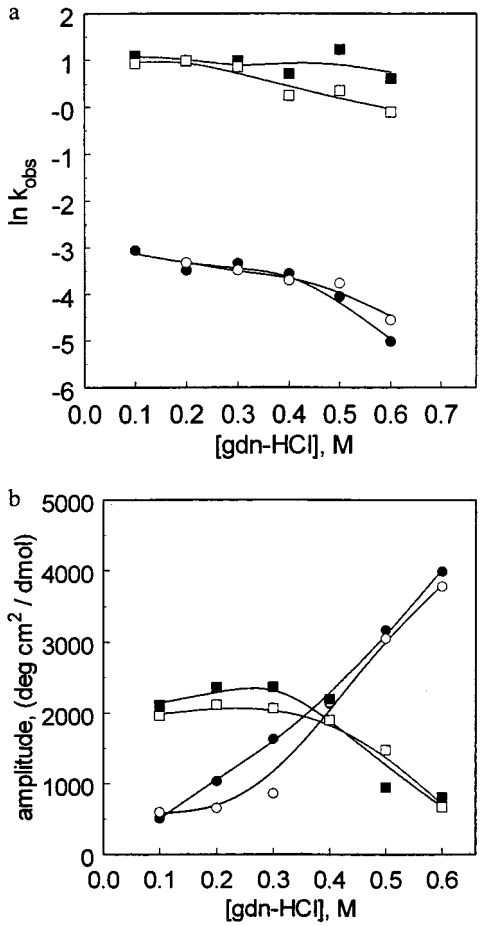

FIGURE 5: (a) Semilog plot of the observed rate constants vs gdn-HCl concentration for P220A (filled symbols) and WT pelC (open symbols) as measured by CD at 218 nm; for the fast phase (squares) and 21-s phase (circles). (b) Folding amplitude for each phase in P220A and WT pelC folding. Solid lines are a smooth line drawn through the data points. Conditions were 25 °C, pH 6, MES 25 mM, NaCl 50 mM.

Refolding Kinetics of P220A. Many examples have been reported indicating the disappearance of slow kinetic folding phases upon mutation of prolines involved in *cis* peptide bonds (*9*, *10*, *16*). The folding kinetics of P220A was studied by far-UV CD and fluorescence. In all cases examined, the manual-mixing refolding monitored by CD showed a single slow phase with a relaxation time in denaturant-free buffer of $\tau = 21$ s. The wild-type protein shows two slow phases with relaxation times $\tau_1 = 21$ s and $\tau_2 = 46$ s. Figure 4a shows a representative refolding experiment at 0.3 M gdn-HCl. Stopped-flow CD experiments were also performed (inset to Figure 4a). These data indicate the presence of a fast kinetic phase very similar to that observed in wild-type pelC. Closer examination of the kinetic data reveals that the refolding rates for both kinetic events are essentially the same in wild-type pelC and P220A. Figure 5 shows the dependence of the folding rate constants and amplitudes on gdn-HCl concentration. This indicates that the folding mechanism is identical in both proteins, except the slow isomerization of the Leu219−Pro220 peptide bond is eliminated. The amplitudes of each respective phase show excellent agreement between wild-type pelC and P220A. There is an increase in the amplitude of the fast phase in P220A. One might expect this increase to be larger because the slowest phase is eliminated in the mutant. The molecules that would have populated that phase should now be able to populate the fast phase. The increase is not large because the equilibrium amplitude is lower in the mutant, resulting in a smaller total change. It also appears that for the fast phase, the dead-time amplitude is slightly greater for the mutant than for the wild-type [see the previous paper in this issue (*2*)]. This is consistent with the difference in the equilibrium magnitude between the unfolded mutant and wild-type protein (Figure 3a). In addition, there is good agreement between the overall change in amplitude as measured by kinetic and equilibrium experiments for P220A.

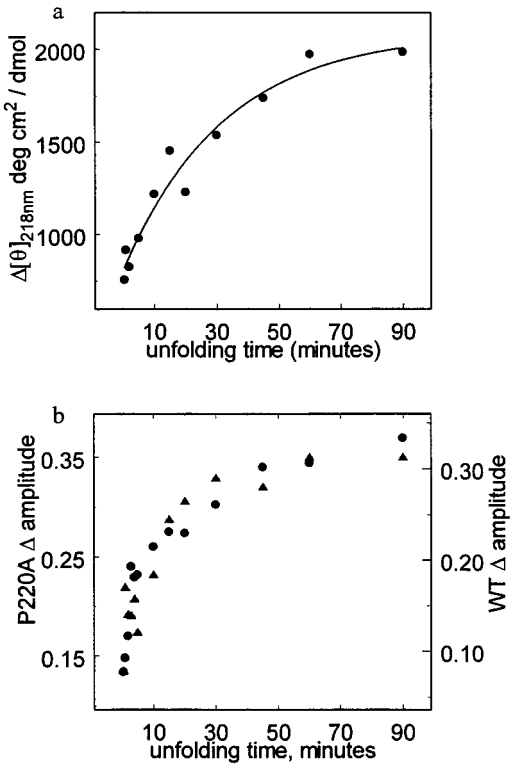

FIGURE 6: Double-jump refolding experiments for P220A monitored by CD at 218 nm (a), and by fluorescence emission (b). In panel b, P220A is represented by (●) and wild-type pelC is represented by (▲). Each is scaled to their respective axes. Conditions were 20 °C, pH 6, MES 25 mM, NaCl 50 mM.

It was observed that a pH jump from 6 to 2 resulted in extremely rapid and reversible unfolding of P220A (data not shown). This change occurs in the dead-time of manual mixing. pH jumps from 2 to 6 were carried out to determine the folding rate of the slow phase in denaturant-free buffer. This experiment is a good control to verify that extrapolating to 0 M denaturant is in fact accurate. CD was not used because the folding amplitude was less than the noise level of the instrument under the conditions employed. Fluorescence emission measurements gave good signals and low noise levels. A pH-jump folding experiment from 2 to 6 at 25 °C indicated a slow phase with a relaxation time of $\tau = 21 \pm 1$ s. This is in quantitative agreement with the rate extrapolated from the CD results in gdn-HCl for both wild-type pelC and the P220A mutant.

Analysis of the 21-s Phase in P220A. In wild-type pelC, it was demonstrated by double-jump assays and other criteria that the 21-s phase is due to proline isomerization in the unfolded protein. Double-jump assays were carried out on P220A to further assess if this phase is the same in P220A as in wild-type pelC. Both CD and fluorescence were used for this experiment. In both cases, an increase in the folding amplitude is observed with increasing unfolding delay time (Figure 6). The results obtained for the mutant by fluorescence are in excellent agreement with those for wild-type pelC (Figure 6b). The relaxation time for isomerization for wild-type pelC at 0 °C and pH 9.5 is 16.5 min and for P220A it is 14.7 min. These data, combined with the refolding kinetics, suggest that the same proline causes the slow kinetics in wild-type pelC and P220A.

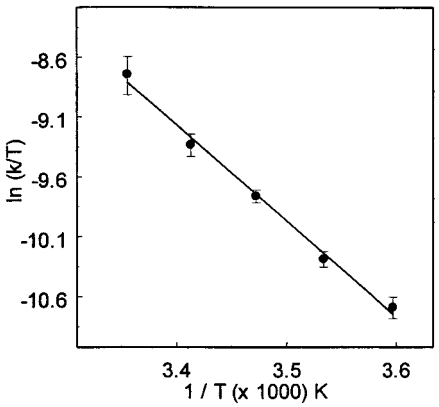

FIGURE 7: (a) Eyring plot for P220A. The error bars represent the standard deviation for the average of three runs, assuming that the temperature is known exactly, for each temperature, and the line is a straight-line fit to the Eyring equation. Conditions were pH 6, MES 25 mM, NaCl 50 mM.

Activation Enthalpy of the 21-s Phase in P220A. In wild-type pelC, it was difficult to obtain an accurate measurement of the activation enthalpy ($\Delta H^{\ddagger}$) for the 21-s phase. This was because the amplitude was low and an extrapolation to 0 M denaturant concentration was necessary. An accurate measurement of $\Delta H^{\ddagger}$ was obtained for P220A. Because the same event is going on in wild-type pelC as in the mutant, this can serve as an accurate measurement for the wild-type. Refolding was initiated by a pH jump from 2 to 6. The refolding kinetics was monitored by fluorescence emission at temperatures from 5 to 25 °C. Figure 7 shows an Eyring plot to determine $\Delta H^{\ddagger}$. This method yielded excellent data with very little scatter in the individual points. The results were $\Delta H^{\ddagger} = 16.0 \pm 0.7$ kcal/mol, and $\Delta S^{\ddagger} = -11.2 \pm 2.6$ cal/mol K. This value for $\Delta H^{\ddagger}$ is in agreement with results from model peptides and other known proteins (*1*), providing further evidence that this phase is due to proline isomerization.

Structure of Trans Proline to Alanine Mutants. To identify the proline responsible for the 21-s folding phase, nine of the remaining 11 prolines were mutated to alanine. The mutagenesis and purification procedure used for P220A worked well for all of these other mutants. DNA sequencing demonstrated that the mutation was present in all cases.

A possible result of any mutation is an observable change in the overall secondary or tertiary structure of the protein. CD experiments were carried out on all Pro→Ala mutants of pelC. Figure 8a shows the far-UV CD spectra of some representative mutants compared to that of the wild-type pelC. In most cases, the mutant CD spectrum is almost identical to the wild-type spectrum. P139A/P220A appears to be slightly less intense in the long-wavelength region and even more so in the short-wavelength region. The loss of intensity in the 218 nm band can be attributed to the presence of the P220A mutation in this double mutant. However, the decrease in the 190 nm band is at least twice as large as that in P220A. Pro139 is located within a tight turn. This is coincidentally next to Ser138 of the serine stack, a stack of polar residues in the core of pelC (*3*). It is possible that this turn is disrupted, thus resulting in the loss of some interactions. The CD spectrum of P263A appears to be somewhat different from that of wild-type pelC. Pro263 is also located in another tight turn at the end of a β-strand. Replacement

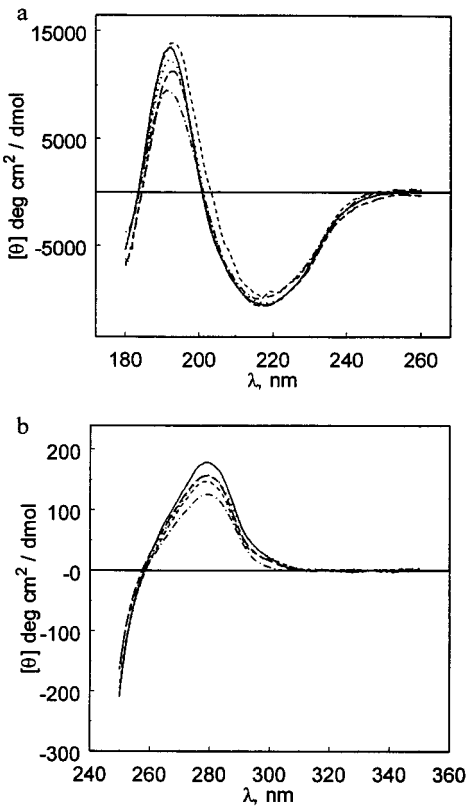

FIGURE 8: (a) Representative far-UV CD spectra for some proline mutants. Shown are wild-type pelC (−), P139A/P220A (-•-•-), P263A (- - -), P51A (····) and P79V (- − - −). (b) Representative near-UV CD spectra for some proline mutants. Shown are wild-type pelC (−), P139A/P220A (-•-•-), P159A (- - -), P316A (····) and P79V (- − - −). CD spectra for mutants not shown are essentially indistinguishable from those for wild-type pelC. Conditions were 25 °C in water for both proteins.

by alanine could introduce some flexibility normally restricted by the proline, thus causing some loss of interactions. The far-UV CD spectra of all of the other mutants (Pro 51, 159, 286, 302, 316, 324, and 335) show no significant changes. (The far- and near-UV CD spectra for all mutants are given in the Supporting Information.) Most of these mutants are located in regions of the protein that have been assigned as unordered. Only Pro335 is involved in some secondary structure, the middle residue of a 3_{10}-helix. Pro286 is located at the N-terminus of a short α-helix, and Pro51 is located at the N-terminus of a β-strand. Mutation of these prolines to alanine appears to have no noticeable affect on the secondary structure.

The P220A mutation had no effect on the near-UV CD spectrum. However, several of the *trans* proline mutants showed small effects (Figure 8b). P139A/P220A showed the largest of these differences. This must be attributed to the mutation P139A, because the mutation P220A results in no change to the near-UV CD. Trp142 is close in sequence to Pro139. Distortion of the local sequence could change the geometry of Trp142, resulting in a loss of CD intensity or a change in sign. Pro159 also shows a lower near-UV CD intensity. Trp75 is only about 4 Å away. Pro159 caps a short segment of α-helix, and mutation to alanine might result in a slightly longer helix. This might result in altered excited-state coupling between the tryptophan indole ring and the peptide groups in the helical geometry. The CD spectrum of P316A also shows some loss of intensity. Phe315 might have its geometry distorted. This residue is within 4 Å of several tyrosines involved in an aromatic residue cluster between the N- and C-termini. Distortion of this geometry might affect coupling within this cluster.

Although many of the observed changes in the CD spectra are small, these changes are reproducible. In addition, the purity of the mutant proteins is comparable to the wild-type protein. These facts suggest that the observed changes are real and represent localized changes in secondary or tertiary structure. None of the mutations, however, lead to substantial disruption of the native structure.

Refolding of Trans Mutants, Pro→Ala. Mutation of Pro220 to Ala resulted in the elimination of the slowest phase of folding. The intermediate slow phase is also due to proline isomerization. The Pro→Ala mutants that were made were subjected to kinetic experiments in order to identify the essential proline responsible for this phase. Kinetic refolding experiments were performed on these mutants in a manner identical to those for the wild-type protein (*2*). In all cases the data show two kinetic phases with rates and amplitudes comparable to those for wild-type pelC. Kinetic traces for these and the two Pro→Val mutants are provided in the Supporting Information (Figure S.3). Comparison of the residuals for two-exponential fits (Figure S.4) with those for one-exponential fits (Figure S.5) shows that there are two slow phases for all mutants, except those in which Pro 220 has been mutated. P139A/P220A shows a single phase in manual-mixing experiments with kinetic parameters comparable to those for the corresponding intermediate phase in wild-type pelC.

Table 4 lists the kinetic parameters from these experiments. There is some variation in the kinetic parameters, probably a result of slight changes in the slow folding process due to the mutations (see Discussion). The average of the rate constants for the Pro→Ala mutants agrees well with those for the wild-type (Table 4). However, the amplitude for the 21-s phase in the mutants generally exceeds that for the wild-type, whereas that for the 46-s phase is smaller than that for the wild-type. The total amplitude for the slow phases is smaller than for the wild-type. We do not have an explanation for these amplitude differences. It appears that none of these Pro residues are solely responsible for the observed 21-s phase. We assumed that the fast kinetic events were not likely to be affected by these mutations. This assumption is supported by the close similarity of the kinetic parameters for the fastest phase in P220A to those for the wild-type. Therefore, stopped-flow experiments were not performed on the *trans* mutants.

CD Spectra of Pro→Val Mutants. In two cases (Pro79 and Pro124), the Pro→Ala mutant could not be expressed at significant levels. It is interesting that these residues, although 45 residues apart in the sequence, are within about 6 Å of each other in the tertiary structure. Both of these residues have ϕ and ψ angles that correspond to the poly-proline II structure. Alanine does not preferentially populate this conformation. However, valine and other β-branched amino acids do have some preference for this structure (*17*). These two prolines were mutated to valine, and experiments similar to those for the alanine mutants were carried out. Figure 8 shows the CD spectra of P79V compared to the spectra of wild-type pelC. The spectra of P124V are essentially identical to those of P79V and are therefore not shown (see Supporting

Table 4: Rates and Amplitudes for WT and Pro→Ala or Val pelC Mutants[a]

mutant	k_2 (s^{-1})	k_3 (s^{-1})	a_2 (deg cm^2/dmol)	a_3 (deg cm^2/dmol)	$a_2 + a_3$ (deg cm^2/dmol)
P51A	0.037 ± 0.006	0.0065 ± 0.0003	1600 ± 100	1400 ± 60	3000
P139A/ P220A	0.019 ± 0.002		1000 ± 80		
P159A	0.037 ± 0.004	0.0079 ± 0.0006	1300 ± 160	1300 ± 180	2600
P263A	0.037 ± 0.01	0.0068 ± 0.0006	1400 ± 220	1600 ± 170	3000
P286A	0.025 ± 0.004	0.0057 ± 0.0005	1400 ± 100	1200 ± 140	2600
P302A	0.022 ± 0.002	0.0030 ± 0.0003	1300 ± 70	1000 ± 60	2300
P316A	0.026 ± 0.003	0.0051 ± 0.0003	1400 ± 100	1600 ± 100	3000
P324A	0.029 ± 0.003	0.0083 ± 0.0005	1500 ± 130	1200 ± 140	2700
P335A	0.034 ± 0.007	0.0070 ± 0.0004	1100 ± 120	1600 ± 130	2700
P79V	0.028 ± 0.003	0.0072 ± 0.0006	2000 ± 150	1400 ± 170	3400
P124V	0.023 ± 0.003	0.0080 ± 0.001	2200 ± 360	1500 ± 380	3700
P220A	0.036 ± 0.002		1600 ± 60		
WT	0.031 ± 0.004	0.0068 ± 0.0002	900 ± 60	2500 ± 70	3400
average (all mutants)	0.030 ± 0.002	0.0066 ± 0.0005	1500 ± 100	1400 ± 60	2900
average Pro→Ala mutants	0.030 ± 0.002	0.0063 ± 0.0006	1300 ± 60	1400 ± 80	2700

[a] Conditions for the kinetic experiments were 0.3 M residual gdn-HCl, 25 °C, pH 6, MES 5 mM, NaCl 50 mM. Error estimates are the mean errors from the nonlinear least-squares analysis.

Information). The far-UV CD spectra of both mutants are comparable to those for the wild-type enzyme. The only notable difference is the intensity of the 190 nm band. In both mutants, this is decreased slightly. Both spectra differ slightly from that for wild-type pelC in the intensity of the near-UV band (Figure 8b). There is probably some distortion in the structure of nearby aromatic side chains, but it is not substantial. Overall the mutations do not significantly perturb the structure of pelC to the point that kinetic experiments are likely to show a substantial change in the folding mechanism.

Folding Kinetics of P79V and P124V. Refolding kinetic experiments on the Pro→Val mutants were carried out identically to those for the Pro→Ala mutants. In both cases, as with the previous Pro→Ala mutations, the data fit to the sum of two exponential terms with rate constants and amplitudes similar to those for the wild-type protein (Table 4). In the case of the Pro→Val mutants, the amplitude for the 21-s phase is about twice as large as for the wild-type and about 50% larger than for the Pro→Ala mutants. It is this phase that should be eliminated upon mutation of an essential proline. It appears that neither of these prolines is responsible for the intermediate slow phase. For the 46-s phase, the amplitude of the Pro→Val mutants is comparable to that for the Pro→Ala mutants, and smaller than that for the wild-type. The resultant total amplitude for the Pro→Val is similar to that for the wild-type.

DISCUSSION

The Slowest Step in Folding Is Due to the Isomerization of Leu219−Pro220 Peptide Bond. The hypothesis that slow folding results from proline isomerization first was proposed in 1975 (*1*). It was suggested that a general intermediate in folding was the result of incorrect *cis*−*trans* prolyl peptide bonds. In many small- to medium-sized proteins lacking *cis* peptide bonds, no kinetic phases slower than 1−5 s are observed, but in almost all proteins containing *cis* prolyl peptide bonds, slow phases are observed ranging from tens of seconds to hours. In several cases, proteins containing prolyl peptide bonds in the *trans* conformation also show slow kinetics (*12*, *16*, *18*−*21*). The mutagenesis results presented conclusively prove that the slowest folding phase for pelC is due to isomerization of the Leu219−Pro220 peptide bond. Mutation to alanine eliminates that slow phase completely. Aside from that, this mutation has little effect on the overall secondary and tertiary structure and no detectable effect on the refolding mechanism.

The catalytic activity of P220A is reduced to levels uncharacteristic of the native enzyme. This is not surprising. Heffron et al. (*4*) carried out a structure-based sequence alignment of 14 pectate lyases from five organisms. Pro220 is strictly conserved in all of these proteins. In fact it is the only proline so strictly conserved out of the 12 in *E. chrysanthemi* pelC. This amino acid is obviously functionally important. It maintains the *cis* form in order to align properly another invariant residue, Arg218, for important catalytic function (*4*).

The free energy of unfolding, $\Delta G^{\circ}_{H_2O}$, is also significantly reduced in P220A, $\Delta\Delta G^{\circ}_{H_2O} \approx -4$ kcal /mol. This most likely results from the loss of some intramolecular interactions in the local region of the protein around the mutation. Pro220 is the first residue of a β strand. Forcing the *trans* conformation could significantly distort the backbone. The *m* value is only slightly reduced, by 0.6 kcal/mol M. This would suggest that the overall denaturation mechanism remains unchanged. The only differences lie in the region of the mutation, where some buried surface area might be exposed to the solvent. It is also noted that the denaturation midpoint decreases by ~0.3 to 0.71 M. This correlates with the lower free energy of unfolding at essentially constant *m*.

The Slow Step in P220A Folding Is Coupled to an Intermediate. In simpler proteins that exhibit slow steps in folding, such as RNase A, the rate and amplitude of the slow steps are independent of the final folding conditions, such as denaturant concentration (*22*). This is because the population distribution in the unfolded protein is solely dependent on the *cis:trans* ratio in the unfolded protein. The percentage of molecules with the incorrect isomeric configuration is established in the unfolded protein. For wild-type pelC and P220A, the rate and amplitude for the 21-s phase is strongly dependent on final denaturant concentration. The probable explanation for this is that the isomerization is coupled with an intermediate (*20*, *23*). At low gdn-HCl concentrations (0.1 M) the folding amplitude of the 21-s phase in WT pelC is about 11% of the total. One might expect a prolyl peptide bond to be 10−30% *cis* in the equilibrium-unfolded form.

However, the amplitude increases dramatically at 0.6 M gdn-HCl to over 50%. In P220A the magnitude of the amplitude is about the same as that for the wild-type protein. However, because the total amplitude is lower, and there is one less phase, the fractional amplitudes actually increase to 20% and 90% (for 0.1 and 0.6 M gdn-HCl**,** respectively). At low denaturant (0.1 M), the intermediate is stable and most of the molecules populate this form. As with WT pelC, this intermediate is in rapid equilibrium with unfolded protein [see the previous paper in this issue (*2*)].

The 21-s Phase Is Due to Multiple Proline Isomerizations. Mutation of Pro220 results in elimination of the slowest kinetic phase. All evidence points to the notion that the intermediate slow phase is also due to proline isomerization. This evidence includes the slowly interconverting population of unfolded molecules, the 16 kcal/mol activation enthalpy, the relaxation time of 21 s, and the effect of peptidyl-prolyl isomerase [see the previous paper in this issue (*2*)]. It is interesting that no single proline mutation results in elimination of this phase. In some cases the kinetic parameters are slightly different (Table 4). Differences in amplitudes have been discussed previously. Three mutants have rates that are noticeably lower than the wild-type rate. These are P139A/P220A, P302A, and P124V. They are only reduced by 30 or 40%, however. It is puzzling, however, that none of these mutations show the pronounced effect on the kinetics that would be expected if the 21-s phase were due to a single Pro residue. Several groups have reported the elimination of slow kinetic phases by mutation of prolines (*10*, *12*, *16*) and results from the P220A mutation agree with these. Other groups have reported that mutation of prolines does not eliminate slow phases (*11*) or that certain slow phases are not the result of proline isomerization (*24*). A very interesting paper has recently reported that slow-folding phases can result from isomerization of nonprolyl peptide bonds (*25*). It was suggested that a very small population of peptide bonds could exist as *cis* when the protein is unfolded. Therefore, as the size of the protein increases, slow kinetics resulting from nonprolyl peptide bond isomerization will also be more prevalent. This phenomenon might occur in pelC, but this is uncertain. Indeed many proteins that lack prolines show no slow phases at all. Another possible scenario, similar to the above, is that the 21-s phase results from isomerization of more than one proline with similar activation energies. If several prolines are undergoing isomerization with similar transition state energies, and hence similar rates, their individual folding processes are unlikely to be resolved with the techniques employed. Considering the evidence in favor of an isomerization in the unfolded state, these two scenarios seem equally likely. One piece of evidence that tips the scales in favor of the multiple proline isomerization argument is the fact that the 21-s phase is catalyzed by peptidyl-prolyl isomerase [see the previous paper in this issue (*2*)]. Considering this evidence, it is most likely that the 21-s phase results from an ensemble of prolines isomerizing from the nonnative *cis* isomer to the native *trans* isomer.

ACKNOWLEDGMENT

We would like to express our thanks to Professor Noel Keen for providing starter cultures of cloned pelC.

SUPPORTING INFORMATION AVAILABLE

Far- and near-UV CD spectra, refolding kinetic data for pelC proline mutants, and analysis of nonlinear curve fitting. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Brandts, J. F., Halvorson, H. R., and Brennan, M. (1975) *Biochemistry 14*, 4953−4963.
2. Kamen, D. E., and Woody, R. W. (2002) *Biochemistry 41*, 4173−4723.
3. Yoder, M. D., Lietzke, S. E., and Jurnak, F. (1993) *Structure 1*, 241−251.
4. Heffron, S., Henrissat, B., Yoder, M. D., Lietzke, S., and Jurnak, F. (1995) *Mol. Plant-Microbe Interact. 8*, 331−334.
5. Yoder, M. D., Keen, N. T., and Jurnak, F. (1993) *Science 260*, 1503−1507.
6. Lin, L. N., and Brandts, J. F. (1983) *Biochemistry 22*, 559−563.
7. Lin, L. N., and Brandts, J. F. (1984) *Biochemistry 23*, 5713−5723.
8. Adler, M., and Scheraga, H. A. (1990) *Biochemistry 29*, 8211−8216.
9. Schultz, D. A., Schmid, F. X., and Baldwin, R. L. (1992) *Protein Sci. 1*, 917−924.
10. Kelley, R. F., and Richards, F. M. (1987) *Biochemistry 26*, 6765−6774.
11. Schindler, T., Mayr, L. M., Landt, O., Hahn, U., and Schmid, F. X. (1996) *Eur. J. Biochem. 241*, 516−524.
12. Eyles, S. J., and Gierasch, L. M. (2000) *J. Mol. Biol. 301*, 737−747.
13. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory, Plainview, NY.
14. Kita, N., Boyd, C. M., Garrett, M. R., Jurnak, F., and Keen, N. T. (1996) *J. Biol. Chem. 271*, 26529−26535.
15. Kamen, D. E., Griko, Y., and Woody, R. W. (2000) *Biochemistry 39*, 15932−15943.
16. Wood, L. C., White, T. B., Ramdas, L., and Nall, B. T. (1988) *Biochemistry 27*, 8562−8568.
17. O'Connell, T. M., Wang, L., Tropsha, A., and Hermans, J. (1999) *Proteins 36*, 407−418.
18. Jullien, M., and Baldwin, R. L. (1981) *J. Mol. Biol. 145*, 265−280.
19. Veeraraghavan, S., and Nall, B. T. (1994) *Biochemistry 33*, 687−692.
20. Jackson, S. E., and Fersht, A. R. (1991) *Biochemistry 30*, 10428−10435.
21. Ogasahara, K., and Yutani, K. (1997) *Biochemistry 36*, 932−940.
22. Schmid, F. X., and Baldwin, R. L. (1979) *J. Mol. Biol. 133*, 285−287.
23. Schmid, F. X. (1986) *FEBS Lett. 198*, 217−220.
24. Herning, T., Yutani, K., Taniyama, Y., and Kikuchi, M. (1991) *Biochemistry 30*, 9882−9891.
25. Pappenberger, G., Aygun, H., Engels, J. W., Reimer, U., Fischer, G., and Kiefhaber, T. (2001) *Nat. Struct. Biol. 8*, 452−458.
26. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−2637.
27. Johnson, W. C. (1999) *Proteins 35*, 307−312.
28. Provencher, S. W., and Glöckner, J. (1981) *Biochemistry 20*, 33−37.
29. Sreerama, N., and Woody, R. W. (1993) *Anal. Biochem. 209*, 32−44.
30. Sayle, R. A., and Milner-White, E. J. (1995) *Trends Biochem. Sci. 20*, 374−376.

BI0115131