Minimum Folding Unit of Dystrophin Rod Domain

Edith Kahana and W. B. Gratzer*

Medical Research Council Muscle and Cell Motility Unit, King's College, 26-29 Drury Lane, London WC2B 5RL, U.K.

Received February 22, 1995; Revised Manuscript Received April 17, 1995[⊗]

ABSTRACT: Fragments of the rod domain of dystrophin, which consists of spectrin-like repeating sequences, have been prepared by expression in *Escherichia coli*. The phasing established earlier for the dystrophin rod, as well as for *Drosophila* spectrin and smooth muscle α-actinin, suggested a length of less than 113 residues for the dystrophin repeat that we have chosen. Fragments with a common N-terminus and lengths between 113 and 119 residues were prepared. The formation of the stable native tertiary fold could be recognized by resistance to proteolysis, the circular dichroism spectrum in the regions of both peptide and aromatic absorption bands and the resolution of the long-wavelength component in the tryptophan absorption spectrum. It was found that the critical length for folding was 117 residues: shortening the chain by 1 further residue resulted in loss of the capacity to form a defined tertiary structure. Residue 117 is a glutamine; replacement of this by a methionine residue did not impair the ability of the chain to enter the folded conformation, implying that it is the length of the C-terminal α-helix, rather than any specific side-chain interaction, that is critical in determining the stability of the native structure. The fragment of 119 residues forms a significantly more stable structure than that of 117. It appears that the minimum unit capable of forming the native fold extends some residues into the adjoining sequence repeat.

The dystrophin molecule is made up of 4 domains: an actin-binding element at the N-terminal end and a cysteine-rich domain with an adjoining dystroglycan-binding domain at the C-terminal end are separated by a rod, comprising 25 spectrin-like repeating units (Koenig et al., 1988). The average length of the repeating units is greater by a few amino acids than the 106 residues that characterize most repeats of erythroid spectrin (Speicher & Marchesi, 1984), and they also display a less stringent homology among themselves. However, in both these proteins, as well as in α-actinin, a triple-helical coiled-coil structure is predicted by the sequence (Speicher & Marchesi, 1984; Davidson & Critchley, 1988; Cross et al., 1990; Parry & Cohen, 1991). We have previously found (Kahana et al., 1994) that the phasing of the repeating sequence closely resembles that determined for *Drosophila* spectrin (Winograd et al., 1991) and more recently for smooth muscle α-actinin (Gilmore et al., 1994). The rod domains of all three proteins are highly α-helical, as expected, and the phasing in all cases implies that a single continuous α-helix extends through pairs of adjoining repeating units, the N- and C-terminal halves forming parts of two separate three-helix bundles. There is strong evidence for the validity of this model from the high-resolution crystallographic structure of a *Drosophila* spectrin repeat (Yan et al., 1993). However, if, as seems likely (Cross et al., 1990; Parry & Cohen, 1991), the common helix extending across successive repeats leads to a nested structure, it remains unclear how the minimum folding unit is defined. Both the spectrin and dystrophin repeats were identified by enzymic digestion of longer stretches of sequence to a protease-resistant core. The positions of the N-termini of the fragments were determined by sequencing, but that of the C-terminus could only be inferred from the apparent molecular weight (Kahana et al., 1994). We have now established the minimum length required for folding of the same dystrophin repeating unit by expressing and characterizing a series of fragments differing only in the position of the C-terminus.

MATERIALS AND METHODS

Fragments of the human dystrophin sequence were prepared by expression in *Escherichia coli*. Dystrophin DNA was kindly provided by Professor K. E. Davies. The N-terminus of the second sequence repeat of the dystrophin rod was taken to be residue 439 of the protein (Kahana et al., 1994). The necessary DNA sequences were prepared by PCR. The forward primer, common to all required fragments, corresponded to nt 1523−1546, and reverse primers were nt 1844−1861, 1845−1864, 1845−1867, 1845−1870, 1845−1873, and 1858−1879. Cloning and expression of all fragments followed the procedure employed previously to prepare longer sequences from the same part of the dystrophin chain (Kahana et al., 1994). A mutant of the 117-residue fragment, in which the C-terminal glutamine was replaced by methionine, was prepared. The mutation was defined by the anti-sense oligonucleotide, 1856−1873, in which the terminal (1871−1873) triplet, CAA, was replaced by ATG. The required DNA was obtained by PCR, as described before. All DNA used for expression was sequenced.

The expressed protein was found in all cases to be largely in inclusion bodies in the *E. coli* (expression strain BL21/DE3), from which it was recovered by dispersion in 6 M guanidinium chloride, 5 mM dithiothreitol. It was then purified by chromatography in 6 M guanidinium chloride on Sephacryl S-300 (2.5 × 100 cm column). Fractions were screened by $NaDodSO_4$−gel electrophoresis. The purified protein was kept in the guanidinium chloride solution and dialyzed as needed against the required buffer, usually 0.15

[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1995.

0006-2960/95/0434-8110$09.00/0

M sodium chloride, 0.05 M sodium phosphate, pH 7.6. We found that storage in the guanidinium chloride solution led after a few weeks to modification of the protein, apparent from its failure to enter the folded state after removal of the denaturant. All conformational studies were therefore performed within about 2 weeks from the time of recovery from the bacteria.

Protein concentrations were determined spectrophotometrically, using a molar absorptivity at 280 nm, calculated from the composition (Perkins, 1986), of 23 350 cm^{-1} M^{-1}. For proteolysis with chymotrypsin, the protein at about 0.15 mg mL^{-1} in the isotonic sodium phosphate−chloride buffer was incubated with 25 μg mL^{-1} enzyme for varying lengths of time at room temperature. The reaction was stopped by addition of excess phenylmethanesulfonyl fluoride in ethanol, and the product was examined by electrophoresis in a 10% polyacrylamide gel in $NaDodSO_4$ (Laemmli, 1970). Chemical cross-linking with dimethyl suberimidate was carried out as described previously (Kahana & Gratzer, 1991).

Circular dichroism was measured in a Jobin-Yvon CD6 instrument, in a path length of 0.2 or 0.5 mm (or longer in the region of aromatic absorption). For thermal denaturation profiles, the temperature was in general varied between 5 and 70 °C in steps of 5 °C. Urea denaturation profiles by circular dichroism or absorbance were performed by adding aliquots of the protein stock solution to solid urea. Time for equilibrium to be attained was less than 30 min. The urea concentrations and dilution factors were calculated from the standard density data of Kawahara and Tanford (1966). To determine equilibrium constants for folding from the thermal and urea denaturation profiles, correction was made for the gradual (linear) change in ellipticity on either side of the sigmoid transition (Pace & Laurents, 1989). Then if the corrected molar residue ellipticity at any temperature or urea concentration is $[\theta]$, while those of the folded and unfolded forms are $[\theta]_f$ and $[\theta]_u$, respectively,

$$K = ([\theta]_f - [\theta])/([\theta]_f - [\theta]_u) \quad (1)$$

Fluorescence spectra were measured in a Spex Fluoromax spectrofluorometer at 20 °C, with excitation at 290 nm.

RESULTS

The second repeat of the dystrophin rod domain, the boundaries of which we had previously defined to within a few residues (Kahana et al., 1994), was selected for study. All expressed fragments were judged by $NaDodSO_4$−gel electrophoresis to be at least 95% pure.

As has been noted before for repeating units of an α-spectrin (Winograd et al., 1991), dystrophin (Kahana et al., 1994), and α-actinin (Gilmore et al., 1994), the native fold is characterized by resistance to proteolysis and high α-helicity. A fragment that does not conform to the correct phasing has significantly lower but still considerable α-helicity. In the case of dystrophin, we found that an additional criterion for folding was the appearance of a highly resolved tryptophan fine-structure band in the ultraviolet absorption spectrum (Kahana et al., 1994). The near-ultraviolet circular dichroism, which arises from asymmetric environments of the aromatic, especially indole chromophores (Woody, 1994), also showed increased amplitudes of the indole Cotton effects in the folded state. All four criteria for folding gave concordant conclusions for the fragments examined here.

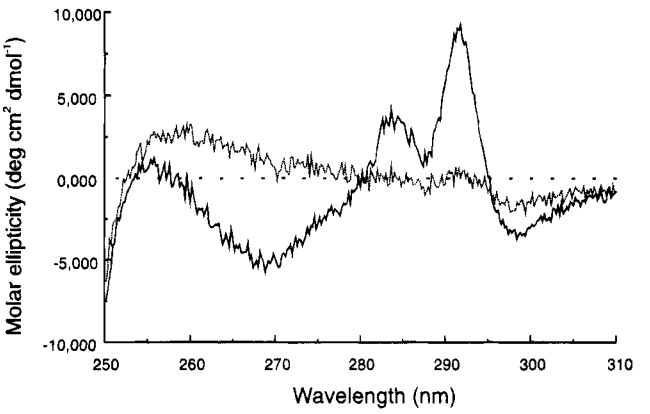

FIGURE 1: Circular dichroism spectra in the aromatic absorption region of dystrophin fragments F115 (broken curve) and F119, showing the effect of formation of the tertiary fold. Molar ellipticity refers to molar concentration of polypeptide chains.

From the positions of probable chymotrypsin-sensitive sites, the C-terminus of the stable, folded dystrophin fragment was inferred to correspond most probably to a chain length of about 109 amino acids (Kahana et al., 1994). Moreover, alignment on the N-terminal sequence, which can be precisely defined in terms of the heptad hydrophobic repeating motif (Cross et al., 1990), places the start of the adjacent repeat at position 110, counting as residue 1 the N-terminal methionine of the stable fragment (residue 439 in the dystrophin sequence). A fragment of 113 residues with the same N-terminus (F113, below) proved, however, not to be folded, as judged by the criteria enumerated above. A series of fragments, extended at the C-terminal end, were expressed and purified. They had the following sequences:

F113: MEKQ...EDRW

F114: MEKQ...EDRWV

F115: MEKQ...EDRWVL

F116: MEKQ...EDRWVLL

F117: MEKQ...EDRWVLLQ

F119: MEKQ...EDRWVLLQDI

Only the last two showed high resistance to chymotryptic hydrolysis. By our spectroscopic criteria, moreover, the first four of these fragments do not form the stable fold (Figures 1 and 2 and Table 1). The change in the appearance of the absorption spectrum (Figure 2) is unusual and clear-cut. Fluorescence spectra showed only a small difference (emission maximum shifted by about 2 nm) as between folded and unfolded chains. In aqueous solutions of all products, there was some gradual, concentration-dependent formation of aggregated material, which could be removed by sedimentation. As judged by chemical cross-linking with dimethyl suberimidate, the freshly prepared polypeptides were at least largely monomeric, when compared, for example, to the isolated α-actinin rod domain (Kahana & Gratzer, 1991).

We next sought to determine whether the failure of fragments shorter than 117 residues to assume the native fold reflected simply a decrease in stability (numerically smaller

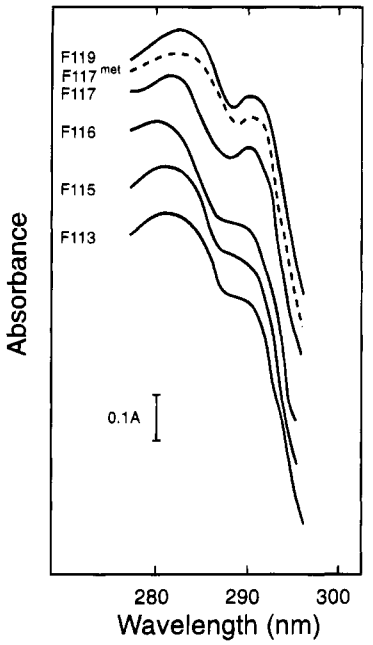

FIGURE 2: Tryptophan fine-structure absorption bands of expressed dystrophin fragments. Successive spectra are displaced on the absorbance axis. Note resolution of the feature near 290 nm (^1L_b transition) in folded polypeptides only.

Table 1: Molar Residue Ellipticities of Expressed Dystrophin Fragments at 222 nm[a]

fragment	molar residue ellipticity at 222 nm[b]	% α-helicity[c]
F113	−16 300	46
F114	−19 200	53
F115	−17 500	49
F116	−17 200	48
F117	−25 600	71
F117 mutant[d]	−24 900	69
F119	−26 500	74

[a] Measured at 20 °C. [b] deg cm^2 dmol^{-1}. [c] Approximate value, taking molar residue ellipticity for the α-helix as −36 000 deg cm^2 dmol^{-1} (Greenfield & Fasman, 1969). [d] Gln-117 → met.

free energy of folding). Thermal denaturation profiles are shown in Figure 3. Fragments F117 and F119 show well-defined sigmoidal transitions, although, as commonly observed (e.g., Pace & Laurents, 1989), there is a more gradual, monotonic change in ellipticity both below and above the transition region; in some preparations, the slope above the transition was considerably less, suggesting the presence of a variable, though small proportion of incompletely folded or aggregated material. The shorter fragments show only a progressive fall in ellipticity with increasing temperature, not preceded by a plateau. Thus, an abrupt loss in the ability of the polypeptide chain to enter the characteristic native fold occurs on loss of the C-terminal residue of fragment F117.

The profiles for F117 and F119 yielded linear van't Hoff plots within experimental error, leading to values of 41 and 62 kcal mol^{-1}, respectively, for the enthalpies, $\Delta H°$, of denaturation. The change in heat capacity on denaturation, $\Delta C_p = (\partial \Delta H/\partial T)_P$, is smaller for both fragments than can be detected by this means.

Urea denaturation profiles were also determined for all fragments at 20 °C with the same result (Figure 4): again, fragments F117 and F119 show a sigmoidal transition, consistent with a two-state equilibrium. The shorter fragments showed as before only a monotonic reduction in α-helicity with increasing urea concentration. It can be seen that F119 is perceptibly more stable than F117 with respect to both thermal and urea denaturation. With the assumption of a linear dependence of the free energy of folding on urea concentration (Schellman, 1978; Dill, 1985), i.e.

$$\Delta G°(u) = \Delta G°(w) - m[\text{urea}]$$

where $\Delta G°(u)$ and $\Delta G°(w)$ are the free energies of folding in the presence and absence of urea and m is a constant, $\Delta G°(w)$ at 20 °C for F117 and F119 emerges as about −1.5 and −3.4 kcal mol^{-1}, respectively. These may be compared with the apparent free energies at the same temperature, derived from van't Hoff plots, of −1.5 and −3.5 kcal mol^{-1}.

The disappearance of the resolved feature in the tryptophan absorption band of F119 was observed as a function of urea concentration. The change in absorbance difference between 291.5 nm (maximum) and 289.4 nm (minimum) followed that of the circular dichroism, confirming that the resolved absorption feature is indeed a characteristic of the folded state.

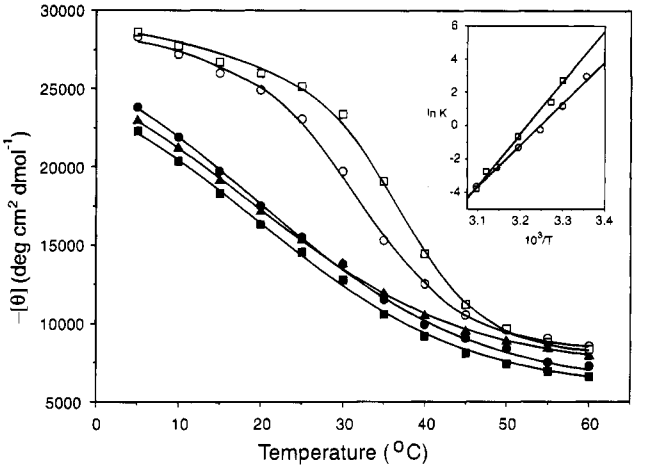

FIGURE 3: Thermal melting profiles of expressed dystrophin fragments, determined by circular dichroism (molar residue ellipticity at 222 nm): F119 (□), F117 (○), F116 (▲), F115 (●), F113 (■). Inset: van't Hoff plots for folding of F119 and F117 (symbols as above).

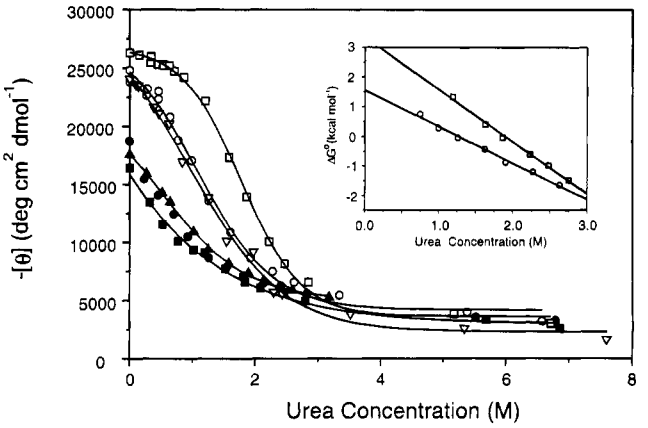

FIGURE 4: Urea unfolding profiles of expressed dystrophin fragments, determined by cicular dichroism (molar residue ellipticity at 222 nm): fragments F119 (□), F117 (○), mutant F117 (▽), F116 (▲), F115 (●), and F113 (■). Inset: Dependence of the apparent free energy of folding on urea concentration for F119 and F117.

The C-terminal residue of F117, which is essential for stability of the tertiary fold, is a glutamine. To determine whether this engages in an energetically important interaction in the tertiary structure, we prepared a mutant of F117 in which the glutamine was replaced by the approximately isomorphous amino acid methionine, for which interactions with an aromatic ring would be excluded, as would energetically significant hydrogen bonding. Circular dichroism (Table 1), as well as resistance to proteolysis and the appearance of the tryptophan fine-structure absorption band (Figure 2), reveals that the mutated fragment forms the native fold, and in fact the conformational stability is barely changed by the substitution (Figure 4).

We can make an estimate of the maximum effect of the substitution on conformational stability from the denaturation data. From the values of $\Delta G°$(w), derived from urea denaturation of the wild-type and mutant F117, the reduction in the free energy of folding is in the range 0.1−0.2 kcal mol^{-1}. A similarly small difference can be inferred from thermal unfolding: at any temperature, T, the ratio of the equilibrium constants, K and K^*, of the wild-type and mutant polypeptides is related to the perturbation in the enthalpy of folding, $\delta(\Delta H°)$, resulting from the mutation, by the equation $\ln(K/K^*) = \delta(\Delta H°)/RT$. The unfolding data gave an estimate of some 2 kcal mol^{-1} for $\delta(\Delta H°)$, with a free energy difference at 20 °C of no more than about 0.2 kcal mol^{-1}.

DISCUSSION

Despite the significant but gradual change in α-helicity with increasing temperature above and below the transition [the former attributable probably to fraying of the helix-bundle ends (Holtzer & Holtzer, 1992), as well as no doubt a small proportion of incompletely folded molecules], it is clear that a cooperative unfolding transition occurs in dystrophin fragments F119 and F117. Reduction in the length of the polypeptide chain from 119 to 117 residues is accompanied by a decrease of some 2 kcal mol^{-1} in the negative free energy of folding at 20 °C. The stable fragment generated by proteolysis of a longer polypeptide was inferred from sequence alignment to be shorter than F113 (Kahana et al., 1994), but as is now clear, it must in fact be considerably longer. Indeed, the greater stability with respect to urea denaturation suggests that it is even longer than F119. Elimination of one further residue from the C-terminus of F117 causes a qualitative change in the properties of the polypeptide, which then fails to enter the tertiary fold. This conclusion is based on conformational stability, reflected not only by cooperative melting profiles but also by resistance to proteolysis, reduced α-helicity, and perturbation of the absorption, circular dichroism, and fluorescence spectra of the tryptophan residues.

The spectroscopic perturbations arise probably from inhibition of free rotation of indole side chains, leading to asymmetric interactions with other side chains [see, e.g., Woody (1994)], and from transfer to a less polar environment, resulting in enhanced resolution of the two electronic transitions in the 280−310 nm region and vibronic structure (Strickland et al., 1969).

The entropy of folding, which should be dominated by hydrophobic side chain interactions on the one hand and configurational entropy on the other, should be minimally affected by a change in length of the polypeptide chain by one residue. Thus, the slope at the midpoint of the unfolding profile should also be unchanged, since $d(\Delta G)/dT = -\Delta S$, and the curves would be displaced toward lower temperature, while retaining approximately the same shape [*cf.* also the sets of such melting profiles generated by a series of α-helical polypeptides, differing only in chain length [see, e.g., Ingwall et al. (1968)]. In actuality, no sigmoidal transition can be discerned for any of the fragments shorter than 117 residues.

The C-terminal amino acid is evidently indispensable for a stable tertiary fold. Specific interactions that a glutamine residue could undergo include side chain hydrogen bonding with an appropriate partner and weak association with an aromatic ring (Brocchieri & Karlin, 1994). To test whether any such sterically specific interactions are required to stabilize the fold, a mutation was chosen that would preclude them. Since the mutant F117 fragment with methionine in place of glutamine at the C-terminus has the capacity to fold and the mutation causes minimal change in the free energy of folding, other explanations for the importance of the terminal residue than a specific side chain interaction must be sought.

Considering especially the high α-helicity of the polypeptide and the rapidity with which it forms the native fold when diluted out of guanidinium chloride solution (Kahana et al., 1994), it seems unlikely that loss of the C-terminal residue from F117 could merely disrupt the folding pathway. A global destabilization of the tertiary structure is thus implied. Such an extreme effect from the truncation of a relatively long α-helix by only one residue is unexpected. A contribution from a relative stabilization of the unfolded state in water with consequent reduction in the magnitude of the free energy of folding (Flanagan et al., 1993) is not excluded. Examination of the crystal structure of the *Drosophila* spectrin repeat (Yan et al., 1993) reveals that, if the dystrophin and spectrin structures are superimposable, Gln-117 of the dystrophin repeat should be located on the outside of the triple-helix bundle and could thus engage in no intramolecular interactions. At the same time, fitting the F117 sequence to the inferred structure of the monomer of the *Drosophila* spectrin repeat (W. R. Taylor, personal communication) shows that Leu-116 (equivalent to Leu-104 in the spectrin structure) packs into the loop between helices A and B. This leucine may be important for stabilizing the structure which may require a bulky side chain in position 117 to shield it from the solvent.

It has been shown that for a series of oligopeptides, containing the heptad hydrophobic motif, formation of two-stranded coiled-coils in aqueous solution requires not less than three complete heptads (Su et al., 1994). The critical residue 117 of our dystrophin fragments falls in position c of the heptad; i.e., the last complete heptad terminates in residue 114. It is possible, however, that the requirements for stabilization of a three-stranded coiled-coil are more stringent and that constraints are also imposed by the bend in the C-terminal helical strand (helix C) of the bundle, as well as the unusual packing of the three helices, implied by the *Drosophila* spectrin structure (Yan et al., 1993).

Model building (Cross et al., 1990; Parry & Cohen, 1991) suggests that the helix bundles which make up the dystrophin rod may be nested and there would then be no structural discontinuity between successive sequence repeats. In this case, the stable fragment of 117 residues described here, which extends into the adjoining phased repeating sequence

at the C-terminal end, may represent the minimum chain segment that is able to fold.

ACKNOWLEDGMENT

We thank the Muscular Dystrophy Group for support which made this work possible. We are grateful to P. J. Marsh and A. J. Henry for much helpful advice on cloning and expression, W. R. Taylor for illuminating discussions on the conformation of the fragments, and H. Patel for growing bacteria and purifying DNA.

REFERENCES

Brocchieri, L., & Karlin, S. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9297−9301.

Cross, R. A., Stewart, M., & Kendrick-Jones, J. (1990) *FEBS Lett. 262*, 87−92.

Davison, M. D., & Critchley, D. R. (1988) *Cell 52*, 159−160.

Dill, K. A. (1985) *Biochemistry 24*, 1501−1509.

Flanagan, J. M., Kataoka, M., Fujisawa, T., & Engelman, D. M. (1993) *Biochemistry 32*, 10359−10370.

Gilmore, A. P., Parr, T., Patel, B., Gratzer, W. B., & Critchley, D. R. (1994) *Eur. J. Biochem. 225*, 235−242.

Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108−4116.

Holtzer, M. E., & Holtzer, A. (1992) *Biopolymers 32*, 7589−7591.

Ingwall, R. T., Scheraga, H. A., Lotan, N., Berger, A., & Katchalski, E. (1968) *Biopolymers 6*, 331−368.

Kahana, E., & Gratzer, W. B. (1991) *Cell Motil. Cytoskel. 20*, 242−248.

Kahana, E., Marsh, P. J., Henry, A. J., Way, M., & Gratzer, W. B. (1994) *J. Mol. Biol. 235*, 1271−1277.

Kawahara, K., & Tanford, C. (1966) *J. Biol. Chem. 241*, 3228−3232.

Koenig, M., Monaco, A. P., & Kunkel, L. M. (1988) *Cell 53*, 219−228.

Laemmli, U. K. (1970) *Nature (London) 227*, 680−685.

Pace, C. N., & Laurents, D. V. (1989) *Biochemistry 28*, 2520−2525.

Parry, D. A. D., & Cohen, C. (1991) *AIP Conf. Proc. 226*, 367−377.

Perkins, S. J. (1986) *Eur. J. Biochem. 157*, 169−180.

Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.

Speicher, D. W., & Marchesi, V. T. (1984) *Nature (London) 311*, 177−180.

Strickland, E. H., Horwitz, J., & Billups, C. (1969) *Biochemistry 8*, 3205−3213.

Su, J. Y., Hodges, R. S., & Kay, C. M. (1994) *Biochemistry 33*, 15501−15510.

Winograd, E., Hume, D., & Branton, D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10788−10791.

Woody, R. W. (1994) *Eur. Biophys. J. 23*, 253−262.

Yan, Y., Winograd, E., Viel, A., Harrison, S. C., & Branton, D. (1993) *Science 262*, 2027−2030.

BI950393S