U.S.A. 82, 3182.
Povirk, L. F., Han, Y.-H., & Steighner, R. J. (1989) Biochemistry 28, 8508.
Singer, B. (1975) Prog. Nucleic Acids Res. Mol. Biol. 15, 219.
Strauss, B., Rabkin, S., Sagher, D., & Moore, P. (1982) Biochimie 64, 829.
Wang, P., Bennett, R. A. O., & Povirk, L. F. (1990) Cancer Res. 50, 7527.
Zimmer, C., & Wahnert, U. (1986) Prog. Biophys. Mol. Biol. 47, 31.

Structural and Thermodynamic Consequences of Burying a Charged Residue within the Hydrophobic Core of T4 Lysozyme†,‡

S. Dao-pin,§ D. E. Anderson, W. A. Baase, F. W. Dahlquist, and B. W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute, and Departments of Physics and Chemistry, University of Oregon, Eugene, Oregon 97403

Received July 22, 1991; Revised Manuscript Received September 16, 1991

ABSTRACT: To determine the energetic and structural consequences of placing a charged group within the core of a protein, two "buried charge" mutants, Met 102 → Lys (M102K) and Leu 133 → Asp (L133D) were constructed in phage T4 lysozyme. Both proteins fold at neutral pH, although they are substantially less stable than wild type. The activity of M102K is about 35% that of wild type, while that of L133D is about 4%. M102K could be crystallized, and its structure was determined at high resolution. The crystal structure (at pH 6.8) of the mutant is very similar to that of wild type except for the α-helix that includes residues 108–113. In wild-type lysozyme, one side of this helix is exposed to solvent and the other contacts Met 102. In the M102K structure this α-helix becomes much more mobile, possibly allowing partial access of Lys 102 to solvent. The stability of M102K, determined by monitoring the unfolding of the protein with CD, is pH-dependent, consistent with the charged form of the substituted amino acid being more destabilizing than the uncharged form. The pK_a of Lys 102 was estimated to be 6.5 both by differential titration and also by NMR analysis of isotopically labeled protein with ^{13}C incorporated at the C^{ϵ} position of all lysines. As the pH is lowered below pH 6.5, the overall three-dimensional structure of M102K at room temperature appears to be maintained to pH 3 or so, although there is evidence for some structural adjustment possibly allowing solvent accessibility to the protonated form of Lys 102.

Charged residues in protein molecules are usually located on the surface (Rashin & Honig, 1984; Wada et al., 1985). When such residues are mobile and solvent-exposed, they appear to contribute very little to protein stability (Dao-pin et al., 1991a). Charged residues that are buried or largely buried are often found to be catalytically or functionally important [e.g., Gelin and Karplus (1977), Kraut (1977), Perutz (1978), Baldwin and Chothia (1979), and Kossiakoff (1983)]. When a charged amino acid is buried, it is usually paired with a residue or ion of the opposite charge and, in addition, has most if not all of its hydrogen-bonding potential satisfied (Rashin & Honig, 1984). This suggests that the burial of an isolated charge in the nonpolar protein interior would be energetically costly. If so, can a protein tolerate such a replacement and still fold? The description of the dielectric environment inside a protein molecule also remains problematic. Different theoretical calculations (Gilson et al., 1985, 1987; Klapper et al., 1986; Dao-pin et al., 1989; Gilson & Honig, 1987; Sternberg et al., 1987) have used a dielectric constant ranging from 2 to 8 and have predicted that it would cost 10–40 kcal/mol to bury a charge in a protein interior (Gilson et al., 1985).

To experimentally probe the nature of the dielectric constant, to measure the energetics of burying a charge, and to observe the tolerance of a protein molecule toward a potentially catastrophic amino acid replacement, two "buried charge" mutations, Met 102 → Lys (M102K) and Leu 133 → Asp (L133D), have been constructed in phage T4 lysozyme. Both sites of substitution are buried in the native protein structure, as shown in Figure 1. The thermal stabilities of both M102K and L133D have been measured by CD-monitored thermal denaturation. Attempts were also made to determine the pK_a values of the substituted residues by differential titration and by NMR. The structure of one of the mutants, M102K, has been determined by X-ray crystallography.

Materials and Methods

(1) Choice and Construction of Mutants. The two mutations M102K and L133D were chosen as representative replacements for two principal reasons. (a) The mutation sites are located near the center of the C-terminal lobe of T4 lysozyme. This is the largest hydrophobic region in the molecule. Residue 102 is located in a buried α-helix which extends from residue 93 to 106. Residue 133 is located within an α-helix that includes residues 127–136 and is exposed to solvent on one side and buried on the other. Both the main-chain and

† This work was supported in part by grants from the NIH (GM21967 to B.W.M. and 1S10RR04926 to F.W.D.), the NSF (DMB8905322 and DIR8815047 to F.W.D.), and the Lucille P. Markey Charitable Trust.

‡ The coordinates of the refined structure of M102K have been deposited in the Brookhaven Protein Data Bank (Ref. 1L54).

* To whom correspondence should be addressed.

§ Present address: Laboratory of Molecular Biology, NIDDKD, National Institutes of Health, Bethesda, MD 20892.

0006-2960/91/0430-11521$02.50/0

the side-chain atoms of residue 133 are buried, as is residue 102. The calculated percentage of solvent-accessible surface area (Richards, 1977) is 0% for residue 102 and 1% for residue 133 in wild-type lysozyme. (b) Both mutations replace a buried hydrophobic residue with a charged residue that is approximately the same overall size and shape. This was intended to minimize disruption of the structure either through the introduction of cavities or through unfavorable steric clashes.

Both the M102K and L133D mutants were created by site-directed mutagenesis following Kunkel's uracil template method (Kunkel et al., 1987). The cysteine-free variant of the wild-type T4 lysozyme (C54T/C97A) (Pjura et al., 1990), referred to as pseudo-wild-type lysozyme and abbreviated as WT*, was cloned into phage M13mp18 which was then transformed into the *dut*$^-$, *ung*$^-$ bacterium strain CJ236 for purification of single-stranded uracil-containing DNA template. Both M102K and L133D were constructed using the WT* template with their complete designations therefore being M102K/C54T/C97A and L133D/C54T/C97A, respectively. Since the two auxiliary WT* mutations (C54T/C97A) were introduced to facilitate thermodynamic measurements, both M012K/C54T/C97A and L133D/C54T/C97A (or M102K/WT* and L133D/WT*) will hereafter be abbreviated to M102K and L133D. The mutations were each confirmed by sequencing the respective T4 lysozyme genes. This also verified that there were no mutations other than those desired. Each mutant lysozyme gene was subcloned into expression vector pHN1403 (Poteete et al., 1991) and transformed into bacterium strain RR1. Subsequently, the expression of lysozyme was screened on the basis of halo-forming ability on IPTG plates (Poteete et al., 1991).

(*2*) *Protein Purification.* The mutant proteins were purified following the normal procedure (Alber & Matthews, 1987; Poteete et al., 1991). The cells bearing the mutant expression vectors were grown first in 100 mL of LBH broth as an overnight culture. This culture was then added to 3 L of LB broth and grown to an OD_{595} between 1.2 and 1.3. At this point, 700 mg of IPTG was added to the growth media to induce lysozyme expression. After about 90 min, the cells were then harvested and centrifuged. From this point, all the purification procedures were carried out at 4 °C. Both the supernatant and the pellet contained lysozyme protein. The supernatant was dialyzed against double-deionized water until its conductivity was less than 4 mmhos/cm. The supernatant was then loaded onto a 2.5 × 10 cm CM-Sepharose (Sigma CL-6B) ion-exchange column. The pellet was resuspended in 100 mL of lysis buffer (0.1 M sodium phosphate buffer, pH 6.5, 0.2 M NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$). A total of 1 mL of 0.5 M EDTA was added, and the suspension was stirred for 2 h to promote cell lysis. After the cells lysed, a few grains of DNase I together with 1 mL of 1 M $MgCl_2$ were added to the suspension to degrade chromosomal DNA. Then, the suspension was centrifuged, the pellet from the centrifugation was discarded, and the supernatant was diluted with double-deionized H_2O until its conductivity was less than 4 mmhos/cm. The sample was then loaded onto the same CM-Sepharose column. A salt gradient from 50 to 300 mM NaCl was used to elute the lysozyme protein. The lysozyme-containing fractions from the CM-Sepharose column were dialyzed against 50 mM sodium phosphate buffer, pH 5.8, and concentrated on a 1 × 5 cm SP Sephadex (Pharmacia C-50) column. The yield of M102K mutant protein per preparation was about 30 mg and that of L133D was about 20 mg. The mutant lysozyme proteins were examined by reverse-phase HPLC (Brownlee aquapore RP-300, 2.1 × 30 mm) and SDS–PAGE and appeared to be homogeneous.

(*3*) *Activity and Stability Measurements.* In vivo lysozyme activities were assayed using the halo-forming abilities of the mutants relative to the cysteine-free pseudo wild type at different temperatures on IPTG lysis assay plates (Poteete et al., 1991).

In vitro lysozyme activities were determined by measuring the initial rates of cell wall hydrolysis at 4 °C. A suspension of lyophilized *Escherichia coli* (obtained from Sigma) cell walls was made in 50 mM Tris-HCl, pH 7.0 (Tsugita et al., 1968). Into 1 mL of this suspension was added 100 ng of lysozyme, and the cell wall hydrolysis was followed for up to 1 min by monitoring the absorbance at 450 nm. The activity is defined as the initial rate of decrease of cell wall turbidity. The activity of WT* was measured in the same experiment under identical conditions.

The thermostabilities of the mutants were measured by CD-monitored reversible thermal denaturation at 223 nm using a J-600 Jasco spectrophotometer (Dao-pin et al., 1990). Thermal denaturation was carried out at pH values from 3.0 to 10.4. From pH 3.0 to 5.3 (Table II) the unfolding of the protein was fully reversible. At pH 6.5, only about 50–75% of the sample unfolded reversibly, and at pH 10, this fraction decreased to about 20%. This loss in reversibility occurred, notwithstanding reduction in the concentration of protein from 20 to 2 μg/mL. Under conditions of partial reversibility, the unfolding T_m can still be determined fairly reliably, but estimates of the free energy of unfolding are subject to error.

(*4*) *Differential Titrations.* The differential titration technique, described by Parsons and Raftery (1972), was used to estimate the pK_a values of the substituted amino acids. About 5–10 mg/mL protein samples were first dialyzed against 0.1 M KCl for 2 days to achieve protein self-buffering. The concentrations of the dialyzed protein samples were adjusted with 0.1 M KCl so that the difference in the sample concentrations between a mutant and WT* was less than 1%. An aliquot (2 mL) of a given protein solution was first titrated by adding 0.2 N KOH until the pH of the solution reached about 9–10. The same solution was then titrated with 0.2 N HCl until the pH was reduced to between 2.0 and 3.0. The protein was kept in a 7-mL test tube, with a flow of water-saturated cold nitrogen gas (0 °C) directed into the tube to prevent CO_2 in the air from dissolving in the sample. The pH meter was a PHM84 radiometer with a GK2421C electrode. The titrants were added through polyethylene no. 20 tubing connected to a 0.25-mL Hamilton syringe driven by a micrometer. The amount of titrant added at each step varied from 0.2 to 20 μL. Because the mutants were expected to be unstable, the experiment was carried out at 0 °C. Each mutant was titrated together with WT* on the same day under the same experimental conditions. The overall procedure was checked by repeating it at a different time with different protein. The amount of acid and base at each pH in the WT* titration was subtracted from the corresponding amount in the mutant titration. The differences, normalized to moles of protein in each sample, were plotted as a function of pH to give the differential titration curve. The pK_a's were estimated by fitting the differential titration curve with the Henderson–Hasselbalch equation (Stryer, 1975) using the nonlinear curve-fitting program NONL kindly provided by Michael Johnson (Johnson & Frasier, 1985).

(*5*) *NMR Titration.* To facilitate identification of the resonances of lysine protons, wild-type and M102K lysozymes were prepared with ^{13}C incorporated at the C^{ϵ} position (ad-

Table I: Activity Measurements

Peptidoglycan Plate Assay	halo size[a] (mm)	
lysozyme	23 °C	37 °C
WT*	9.2	12.7
M102K	4.5	0.0
L133D	0.0	0.0

lysozyme	Cell Wall Assay rate (−OD/min)	rel act. (%)
WT*	0.34 ± 0.05	100
M102K	0.12 ± 0.02	35
L133D	~0.01	4

[a] The estimated accuracy of the halo size assay is ~20%.

jacent to the amino group) in all lysines. Isotopic labeling was achieved by providing 60 mg/L L-[6-^{13}C]lysine (CIL, Lot F-3383) to lysine auxotrophs as described by Muchmore et al. (1989). These samples were transferred to D_2O buffers by repeated microconcentration using Centricon-10's (Amicon). Amides were outexchanged by heating samples to 45 °C for 15 min at pH 2 in 10 mM KCl/10 mM DCl. Final conditions for NMR experiments were 30 mM P_i and 100 mM KCl.

A two-dimensional ^{1}H–^{13}C HSMQC experiment (Zuiderweg, 1990) was performed for both samples at pH* 7.0 to identify the peak due to Lys 102. Thereafter, a one-dimensional version of this experiment (simply the first block of a two-dimensional experiment) was performed to selectively observe the ^{13}C bound protons without separation in the ^{13}C dimension. Nonselective ^{1}H NMR spectra were obtained along with selective spectra to provide correct phases and to monitor pH-dependent transitions. The same presaturation to suppress residual water was used in both experiments. Spectra were collected on a General Electric GN500 NMR spectrometer. The temperature was maintained at 8 °C. The pH was adjusted by addition of buffer and microconcentration.

(*6*) *Structure Determination.* Mutant M102K was crystallized in a batch experiment under conditions similar to those for wild-type (Weaver & Matthews, 1987), and the crystals were found to be isomorphous. Prior to X-ray photography, the crystals were equilibrated with a standard mother liquor consisting of 0.23 M NaCl, 1.26 M NaH_2PO_4, 1.05 M K_2HPO_4, and 1.4 mM BME, pH 6.8. X-ray crystallographic data were collected by oscillation photography (Weaver & Matthews, 1987) to 1.9-Å resolution with 71% completeness. The initial model of the mutant structure was built on the basis of the wild-type structure (Matsumura et al., 1989) and then refined with the TNT program (Tronrud et al., 1987). The typical procedure used for the refinement of mutant lysozymes has been described by Dao-pin et al. (1991a).

Crystals of L133D have not been obtained.

Results

When the M102K protein was expressed in the plasmid pHN1403 in bacterial RR1 cells in vivo, its halo formed on IPTG halo assay plates indicated that this mutation caused the protein to become temperature-sensitive. Table I compares the halo sizes of M102K mutant and of cysteine-free wild type after incubation for 24 h at 23 and at 37 °C.

L133D failed to form halos at either tempertaure, suggesting that the protein was of low stability.

(*1*) *Catalytic Activity.* The activities of the mutants were determined by cell wall hydrolysis at 4 °C in 50 mM Tris-HCl buffer, pH 7.0. About 100 ng of protein was used in each assay. Table I summarizes the results. Under these experimental conditions, M102K has about 35% of the activity of WT* while the apparent activity of L133D is only about 4% of that of WT*.

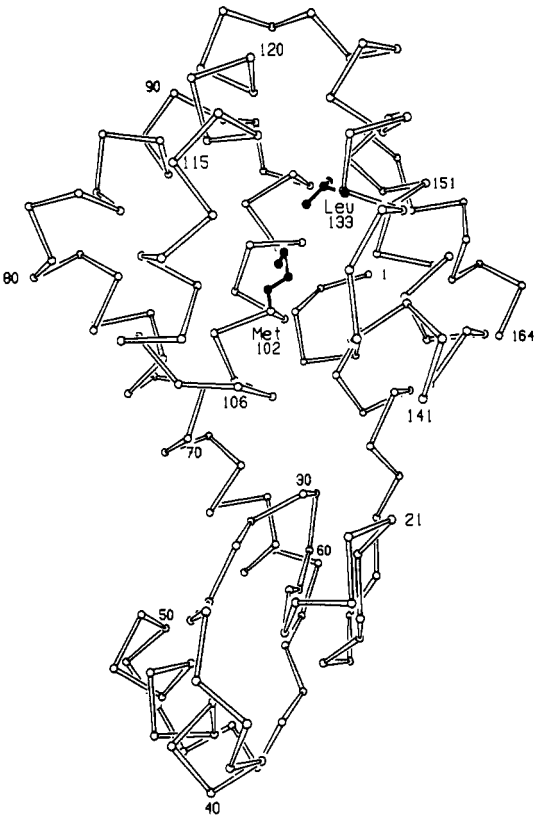

FIGURE 1: Drawing of the backbone structure of T4 lysozyme showing the locations of Met 102 and Leu 133.

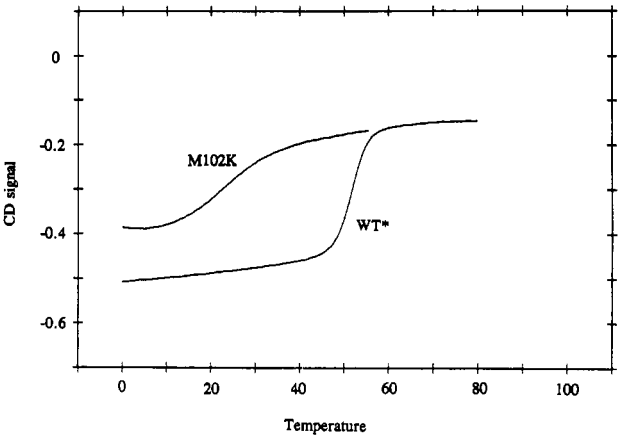

FIGURE 2: CD-monitored thermal melting curves of mutant M102K and WT* at pH 3.0. The curves are normalized so that the CD signals for the unfolded proteins are equal.

(*2*) *Thermal Stabilities.* Thermal denaturation studies were attempted over a wide range of pH values. The buffer was 25 mM KCl/20 mM KPO_4 in all cases except pH 10.4 where 0.1 mM EDTA was also present. The results are listed in Table II. Unfolding of M102K, L133D, and WT* was reversible at pH 3.0–5.3, but only about 20% of the sample was reversible at pH 10.4. All thermal denaturations were analyzed within the framework of a two-state van't Hoff model. At pH 3.0, M102K was only partially folded (equivalent to a 43% loss in α-helix) and was interpreted using the base line for folded WT* (Figure 2).

(*3*) *Differential Titrations.* Before attempting to apply the differential titration technique to M102K, a control was carried out with S90H/Q122D, a mutant for which the pK_a of His

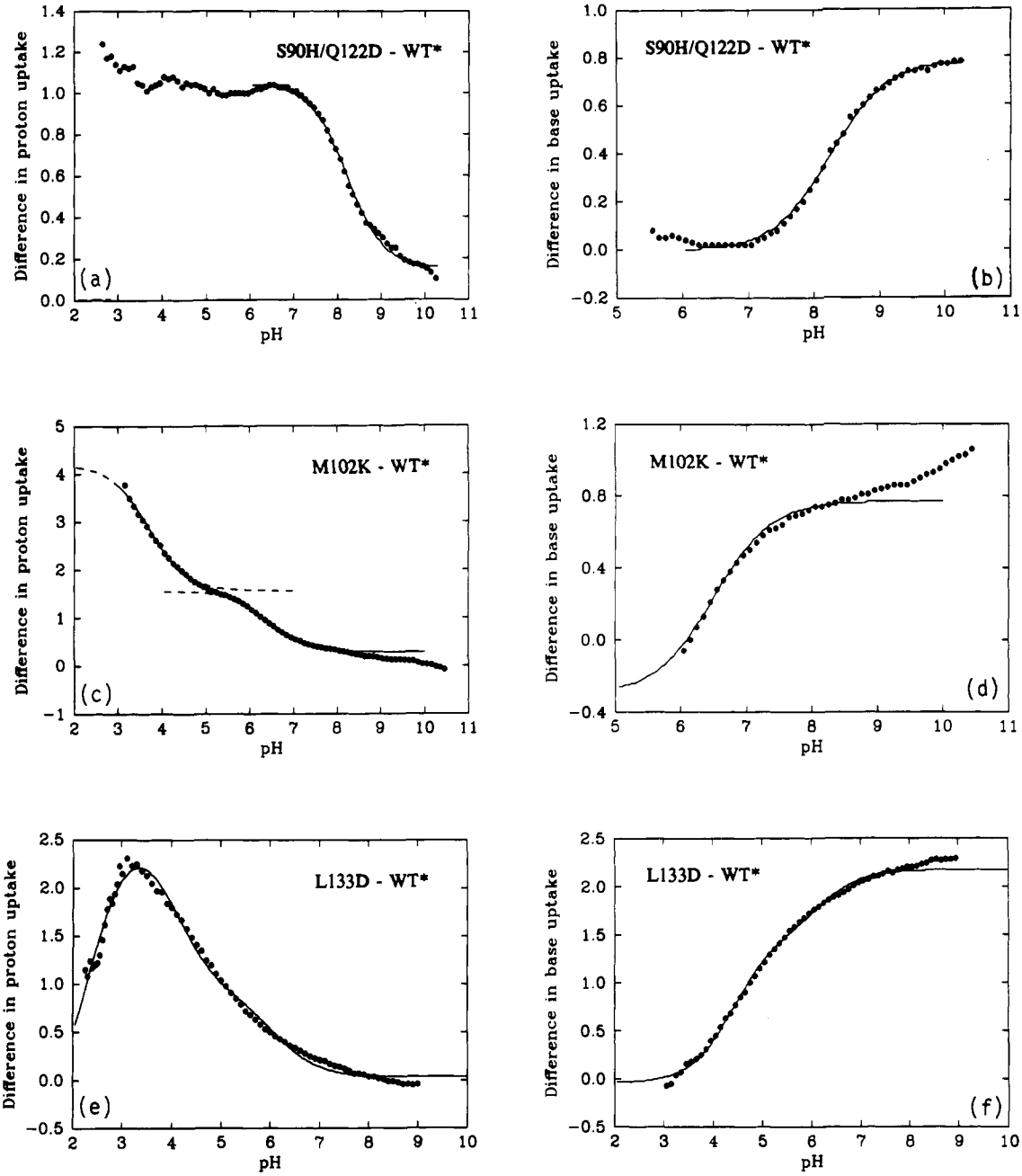

FIGURE 3: Differential titrations of mutant lysozymes relative to the pseudo-wild-type (cysteine-free) lysozyme, WT*. The base titrations are carried out by adding 0.2 N KOH, and the acid titrations are with 0.2 N HCl. In each case the points show the individual experimental values and the continuous curves were obtained by curve fitting using the nonlinear fitting procedure of Johnson (1985). (a) S90H/Q122D vs WT*, acid titration, pH 10.3–2.6. (b) S90H/Q122D vs WT*, base titration, pH 5.5–10.3. (c) M102K vs WT*, acid titration, pH 10.5–3.1. (d) M102K vs WT*, base titration, pH 6.0–10.5. (e) L133D vs WT*, acid titration, pH 9.0–2.2. (f) L133D vs WT*, base titration, pH 3.0–9.0.

90 had been determined previously by NMR to be 7.7 at room temperature (Dao-pin et al., 1991b; E. Anderson and F. W. Dahlquist, unpublished results). Acid and base titrations of this mutant at 0 °C yielded pK_a values of 8.21 and 8.25, respectively (Figure 3a,b). The apparent discrepancy of 0.5 pH unit between the NMR value and those determined by differential titration is presumably due to the difference in the experimental conditions. The NMR experiments were carried out at room temperature using 0.1 M KCl/10 mM D_3PO_4/K_3PO_4 buffer in D_2O. The differential titrations were performed at 0 °C in 0.1 M KCl, but were self-buffered.

The concentration of M102K was 8.64 mg/mL, and that of wild type was 8.67 mg/mL. The amount of 0.2 N KOH consumed in titrating from pH 6.0 to 10.5 was 0.029 mL for the M102K mutant and 0.023 mL for WT*. The amount of 0.2 N HCl consumed in titrating from pH 10.5 to 3.1 was 0.103 mL for M102K and 0.084 mL for WT*, respectively. The differential titration curves for both acid and base titrations are shown in Figure 3b. The base differential titration curve can be reasonably fit by a single group titrating with a pK_a of 6.5. The acid differential titration, however, suggests that more than one proton is taken up in the low pH range and can be decomposed into the sum of two titration curves with pK_a's of 6.4 and 3.7. The pK_a of 6.4 agrees with the base titration result, while the pK_a of 3.7 is outside the range of pH explored in the base titration.

The mutant L133D was titrated in a similar manner. The protein concentration of L133D was 5.74 mg/mL while that of WT* was 5.83 mg/mL. Base titration with 0.2 N KOH was carried out from pH 3.0 to 9.0, and the total amount of 0.2 N KOH consumed in the titration was 0.061 mL for L133D and 0.054 mL for WT*. Acid titration of the same

Table II: Stability of Mutant Lysozymes[a]

	pH	T_m (°C)	ΔT_m (°C)	ΔH (kcal/mol)	$\Delta\Delta G$ at T_m of the mutant (kcal/mol)
WT*	3.0	51.6		117	
M102K		16.6	−35.0	12.4	−8.9
L133D		36.7	−14.9	47.9	−4.9
WT*	4.4	64.6		138	
L133D		49.6	−15.0	59.1	−5.5
WT*	5.3	65.3		135	
M102K		45.0	−20.3	53	−6.9
WT*	6.5	61.9		121	
L133D		44.0	−17.9	50.8	−5.7
WT*	10.4	47.9		79.5	
M102K		37.8	−10.1	77.4	−2.2
L133D		28.9	−19.0	47.1	−3.7

[a] T_m is the melting temperature of the mutant lysozyme, and ΔT_m is the difference between mutant and wild type. ΔH is the enthalpy of unfolding at T_m, and $\Delta\Delta G$ is the difference between the free energy of unfolding of the mutant and that of WT*. $\Delta\Delta G$ values at each pH value were determined at the T_m of the mutant by means of a van't Hoff plot of the WT* data using a constant ΔC_p of 1800 cal/(mol·deg). This comparison is not isothermal; instead, free energy changes are based on the data from the best studied protein, WT*, at temperatures where the ΔG is zero for the mutants. The accuracy of $\Delta\Delta G$ values determined this way is limited by random error only at small values of ΔT_m. For the proteins studied here, ΔT_m's are large and accuracy is limited by the choice for ΔC_p as well as the use of a constant ΔC_p model. We estimate error in $\Delta\Delta G$ to be no greater than ±20% except in the case of pH 10.4, where low reversibility may have reduced accuracy to ±50% primarily by means of an effect on ΔH. However, at pH 10.4, the inversion in the order of instability is clear-cut.

sample with 0.2 N HCl was carried out from pH 9.0 to 2.2 immediately following the base titration. The total amount of 0.2 N HCl consumed in the acid titration was 0.17 mL for L133D and 0.163 mL for WT*. The base differential titration can be interpreted as arising from two groups with pK_a's of 6.2 and 4.4 (Figure 3c). The acid differential titration can be considered as the sum of two positive titration curves with pK_a's of 6.1 and 4.0 and a negative titration curve with a pK_a of 2.3 (Figure 3c). On one hand, it is encouraging to observe that both curves are consistent in displaying apparent titrations at about pH 4.2 and 6.1. On the other hand, the obvious complexity of the titration curves means that the inferred pK_a's must be regarded with considerable caution.

(4) NMR Titrations. Figure 4 compares part of the methylene region of the [6-^{13}C]lysine isotope edited spectra of WT* and M102K T4 lysozymes at pH* 7. Both samples show an intense peak at 2.81 ppm (not shown) and less intense peaks at approximately 2.5 and 2.6 ppm, which represent the 11 protonated lysine residues of the wild-type protein. M102K shows two additional peaks, not seen in wild type, of ^{1}H chemical shift 1.63 and 1.16 ppm (Figure 4). In a two-dimensional ^{1}H–^{13}C HSMQC spectrum (not shown), these two proton resonances give cross-peaks to a single carbon resonance of chemical shift 40.39 ppm. We assign these two resonances to the prochiral pair of protons on C^ϵ of Lys 102. At pH 7, these two peaks are shifted 1.3–1.7 ppm upfield of the main peak of the C^ϵ protons of the other lysine residues. This shift is consistent with an unprotonated lysine residue. As the pH is lowered below 6.5, these two peaks broaden and/or decrease in intensity. These resonances may become part of the intense unresolved peaks near 2.8 ppm corresponding to the other protonated lysine residues. The observations are consistent with the view that Lys 102 protonates near pH 6.5 in the mutant M102K.

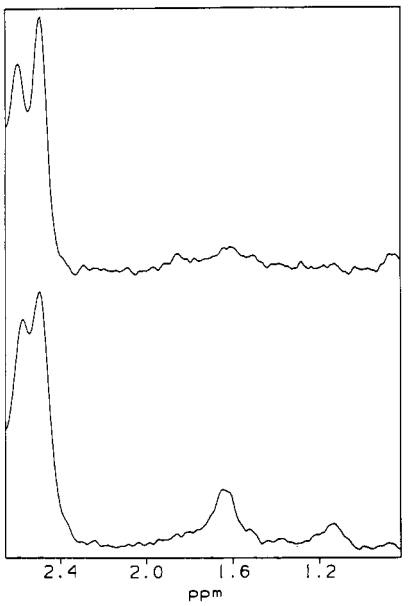

FIGURE 4: A detail of the methylene region of [^{13}C]lysine isotope edited spectra of WT* (top) and M102K (bottom) T4 lysozymes at pH* 7.0. Chemical shift is relative to internal DSS. Phases were obtained from unedited spectra with a 180° correction in the zeroth order phase correction. Residual intensity between 2 and 1 ppm is seen in the 2D experiment to originate from natural abundance resonances 10 ppm upfield in ^{13}C from the lysine resonances.

Table III: Data Processing and Refinement Statistics for M102K

data processing	
cell dimensions (Å)	$a = b = 61.0$ $c = 96.4$
no. of independent reflections	12166
R_{merge}	0.058
refinement[a]	
resolution (Å)	6.0–1.9
rms bond length (Å)	0.018
rms bond angle (°)	2.72
rms planar group (Å)	0.016
crystallographic residual (%)	15.4

[a] The values of the rms bond length, bond angle, and planar groups give the average discrepancies of the final refined model from the values expected for an "ideal" structure.

Direct observation of the nonexchangeable resonances downfield of water provides some insight into the nature of the mutant and wild-type proteins under acidic conditions. Between pH 4 and 6, M102K shows spectra characteristic of a folded protein. There is dispersion of the aromatic protons, and three resolved resonances (the protons on $C^{\epsilon 1}$ and $C^{\gamma 2}$ of His 31 and the C^α proton of Lys 15) have chemical shifts characteristic of the folded state of the wild-type protein. However, there are substantial changes in the pattern of aromatic residue chemical shifts when M102K is compared to the wild type. These changes are accentuated below pH 6 and suggest some reorganization of the aromatic residues of M102K as the pH is lowered from 8 to 6. This change is not seen in the wild-type protein. When the pH is lowered below 3.5, M102K begins to unfold and all side-chain chemical shift dispersion is lost by pH 3. The wild-type protein is still folded under these conditions.

(5) The Structure of M102K. The initial map showing the difference in electron density between M102K and wild-type lysozyme is shown in Figure 5a. The large negative feature centered on S^δ of wild-type Met 102 is due to the replacement of the electron-dense sulfur atom with a less dense carbon atom, as expected for replacement of a methionine with a lysine. There are also negative features centered on some of

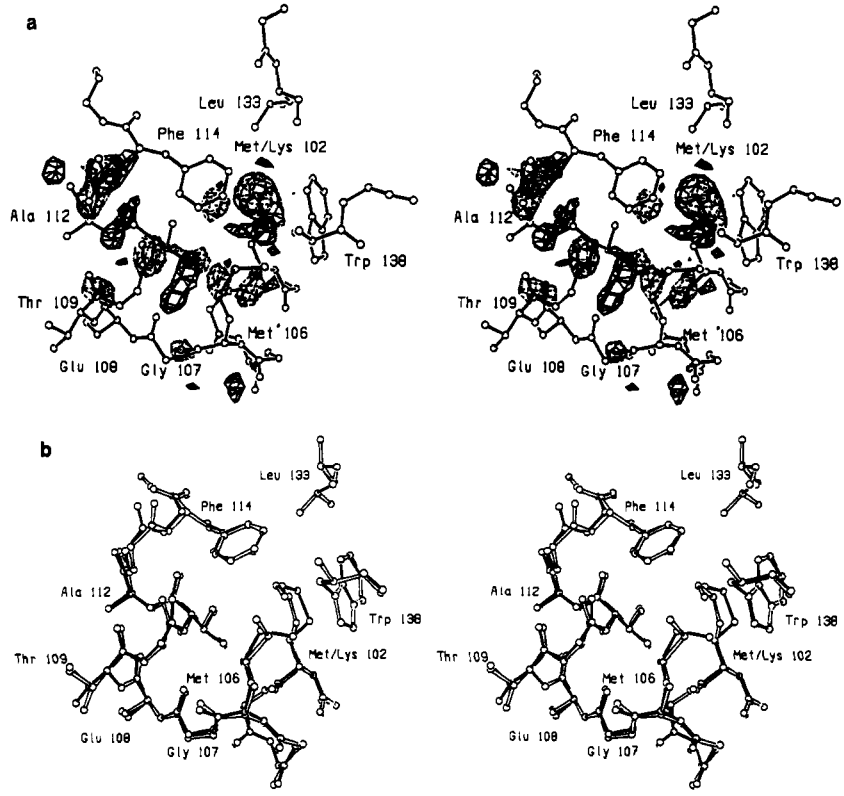

FIGURE 5: (a) Map showing the difference in density between mutant M102K and wild-type lysozyme. Amplitudes are ($F_{mut} - F_{WT}$) where F_{mut} and F_{WT} are the observed amplitudes for the mutant and WT crystals and the phases were from the refined structure of pseudo-wild-type lysozyme (Weaver & Matthews, 1987). Solid contour lines (positive density) represent features present in the mutant structure but absent in the mutant. Negative contours (broken lines) indicate electron density that is weaker or absent in the mutant structure relative to that of the wild type. Contours were drawn at $\pm 3.5\sigma$ where σ is the root mean square difference density throughout the unit cell. Note that in this calculation the amplitudes and phases for WT rather than WT* lysozyme were used since, at the time the calculation was performed, the refined model for WT lysozyme was superior to that for WT*. (b) Superposition of the refined structure of M102K lysozyme (open bonds) on that of wild-type (solid bonds) in the vicinity of the amino acid replacement.

the main-chain atoms between residues 106 and 114, suggesting that this part of the mutant structure has become more mobile in the mutant structure than in that of the wild type.

Table III summarizes the statistics for the X-ray data collection and structure refinement. The superposition of the refined M102K structure on that of wild type in the vicinity of the mutation (Figure 5b) shows that the coordinates of the two structures are very similar. In particular, the α-helix 106–114 remains in the same position as in wild type. If this helix were to have shifted (in contrast to becoming more mobile), one would expect to see pairs of positive and negative density features in Figure 5a. This is not observed. The root mean square difference between all atoms in the two structures is 0.16 Å. At the mutation site, the side-chain conformation of Lys 102 is very similar to that of the wild-type Met 102, except with an extra atom (N^{ζ}) added. The ϵ-methyl of Met 106 does, however, assume a different conformation in the mutant relative to in the wild type. The new conformation of Met 106 permits a 2.9-Å hydrogen bond between S^{δ} of Met 106 and N^{ζ} of Lys 102. Although the coordinates of the mutant and wild-type structures are very similar, there is a striking difference in the mobility of residues 105–114. This 10-residue segment, which forms a turn and a short α-helix, becomes much more mobile in the mutant structure than in wild-type lysozyme. The average crystallographic thermal factor (B value) for residues 105–114 is 24.9 Å^2 in wild-type lysozyme and increases to 47.4 Å^2 in the mutant structure. Figure 6 shows the difference between the main-chain B values in the two structures. Except for residues 105–114, the thermal factors throughout the remainder of the two structures are similar.

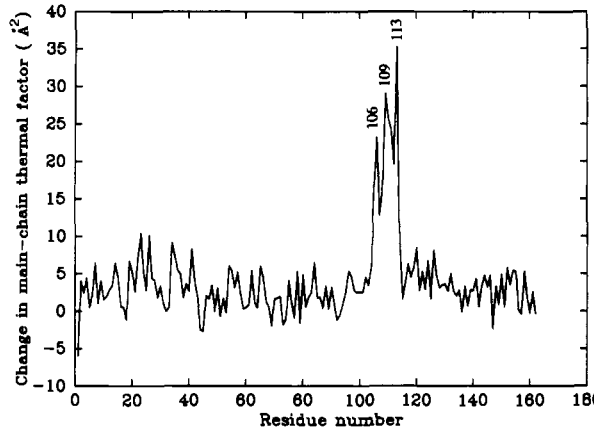

FIGURE 6: Plot showing the difference between the backbone mobility of mutant M102K and wild-type lysozyme. For each residue in the respective structures the individual crystallographic thermal factors for N, CA, C, and O were averaged to give a mean value. The plot shows the differences between these values for M102K and wild-type lysozymes.

Discussion

(1) Adaptability of the Protein to Mutations. Perhaps the most striking result of the present study is that the protein can accept the introduction of a lysine or an aspartic acid within its hydrophobic core and yet still fold. The ability of proteins to adapt to different mutations has been emphasized before (Matthews, 1987), but the present examples are more extreme. Not surprisingly, the introduction of the charged residue substantially decreases the stability of the protein. For L133D, the melting temperature is decreased by 15–20 °C in the pH

Table IV: Interatomic Contacts for Met 102 in WT Lysozyme and Lys 102 in M102K[a]

Met 102	contact		distance (Å)	Lys 102	contact		distance (Å)
CB	O	Ala 98	4.0	CB	O	Ala 98	3.8
CB	O	Leu 99	3.7	CB	O	Leu 99	3.7
CB	CE	Met 106	4.0	CB	CG1	Val 111	4.0
CB	CG2	Val 111	3.7				
CG	O	Ala 98	3.7	CG	OH	Sol 217	4.0
CG	CG1	Val 149	4.0				
CG	OH	Sol 208	3.9				
SD	CE1	Phe 153	3.7				
CE	CE	Met 106	3.5	CE	CD1	Phe 114	3.8
				CE	CE1	Phe 114	3.8
				NZ	SD	Met 106	2.9(H)
				NZ	CE	Met 106	3.9
				NZ	CE1	Phe 114	3.8
				NZ	CE2	Phe 114	3.9
				NZ	CZ	Phe 114	3.6
				NZ	CG	Trp 138	4.0
				NZ	CD2	Trp 138	3.6
				NZ	CE2	Trp 138	3.8
				NZ	CE3	Trp 138	3.7

[a] The table gives all contact distances less than 4.0 Å involving the side-chain atoms of Met 102 and Lys 102. The symbol (H) denotes a hydrogen bond.

range 3.0–10.4. Presumably, the aspartate is negatively charged within at least the upper part of this range (see below), but since we do not have the crystal structure of this variant, we do not know whether or not the aspartic acid has access to solvent.

In the case of the mutant with the buried lysine, M102K, the stability is very pH-dependent, with the decrease in melting temperature relative to that of wild type increasing from 10.1 °C at pH 10.4 to 35 °C at pH 3.0. The destabilization of the mutant structure relative to that of wild type increases from 2.2 kcal/mol at pH 10.4 to 8.9 kcal/mol at pH 3.0 (Table II). In general terms, this pH dependence is as expected on the basis of the reasonable assumption that it requires less energy to bury the lysine in its uncharged form (i.e., at high pH) than at low pH when the lysine is protonated.

The crystal structure of M102K confirms that the lysine is buried and inaccessible to solvent at pH 6.8, the pH of the crystals (but see the discussion below on the mobility of α-helix 108–113). One cannot tell, however, from the crystallographic data whether or not the buried lysine is positively charged. The environment of the lysine side chain, including the ε-amino group, is very hydrophobic (Figure 5b). Table IV summarizes the closest interatomic distances between the lysine side chain and the surrounding protein. The approach of 2.9 Å between the ε-amino group of the lysine and S^{δ} of methionine 106 indicates the presence of an N-H···S hydrogen bond. There is no suggestion of a second hydrogen bond from N^{ϵ} of the lysine.

As judged from the crystallographic thermal factors (Table V), the atoms within the side chain of Lys 102 are substantially more mobile than the atoms of the methionine that they replace. This suggests that the atoms within the lysine side chain lack the specific van der Waals contacts with the surrounding atoms that occur for the methionine and serve to restrict its motion to a value comparable with the rest of the hydrophobic core.

The increased mobility that is induced in the 108–113 α-helix by the Met 102 → Lys replacement is of special interest. It should be noted that the helix does not significantly change its average position (Figure 5b). It is not as if the helix moves "inward" to help fill a cavity created by the amino acid replacement. Neither does the helix move "outward" because the lysine is too big to fit into the space vacated by the methionine. Rather, the helix retains essentially the same equilibrium position but becomes more mobile. The exact nature of this increased mobility is not certain. One interpretation is that the helix remains essentially "in place" but undergoes displacements of larger amplitude about its rest position than is the case in the wild-type structure. An alternative explanation, however, is that the helix becomes completely disordered. According to this model, which is equally consistent with the X-ray data, the crystal structure is an equilibrium mixture of two conformational states. In one state the 108–113 helix remains in place with mobility essentially the same as in wild-type lysozyme. In the alternative state the 108–113 helix is disordered and not visible in the electron density map. In this state Lys 102 would presumably be accessible to solvent. Helix 108–113 is on the surface of T4 lysozyme but is not involved in any crystal contacts (Weaver & Matthews, 1987). Therefore, the helix could become disordered even within the constraints imposed by the crystal lattice. The disordering of α-helix 108–113 would be assumed to occur on protonation of Lys 102 and would therefore be favored by low pH. The disordering of this helix would help to account for the reduction in CD signal seen in the folded structure of M102K at pH 3.0 (Figure 2), although it would not explain the apparent reduction of the overall α-helix content by 43%.

Table V: Mobility of Residue 102

Met 102		Lys 102	
atom	thermal factor B (Å^2)	atom	thermal factor B (Å^2)
N	9.9	N	12.0
C	16.0	C	21.9
O	14.2	O	18.1
CA	17.8	CA	24.1
CB	12.9	CB	27.7
CG	14.8	CG	31.0
SD	18.6	CD	45.5
CE	16.7	CE	55.2
		NZ	62.3

[a] The mean square amplitude of atomic vibration ($\overline{r^2}$) is related to the thermal factor B by $B = 8\pi^2\overline{r^2}$.

(2) The pK_a Values of the Mutant Side Chains and the Energetic Cost of Burying a Charge. The difference titrations suggest that the mutant M102K differs from wild type in having two groups that titrate respectively at pH 6.5 and 3.7. Presumably, it is the introduced Lys 102 that has an apparent pK_a of 6.5, although the lower value must also be considered. If, indeed, the pK_a of Lys 102 is 6.5, then the pK_a of 3.7 must be caused by the titration of one or two acidic groups whose pK_a's are 3.7 in the mutant but less than 3 in WT*. There is only one histidine residue, His 31, in T4 lysozyme, and its pK_a in WT* has been determined by NMR to be about 9.1 (Anderson et al., 1990). Proton NMR was also used to verify that the pK_a of His 31 in M102K is greater than 8.0 (data not shown). This excludes the possibility that the observed pK_a of 6.5 is due to His 31. The difference titrations therefore indicate that the pK_a of Lys 102 in the mutant M102K is 6.5.

The NMR results also support the assignment of a pK_a value of 6.5 to Lys 102. In addition, the NMR data suggest that M102K begins to unfold between pH 3 and 4, providing an explanation for the other proton adsorptions (pH midpoint 3.7) seen by difference titration.

The apparent titration chemical shift of Lys 102 is about 0.7 ppm larger than that observed for model compounds. This suggests that factors other than the amino-group protonation may be involved in determining the chemical shift of Lys 102. Examination of the crystal structure (see Figure 5) of M102K

reveals that the ε-hydrogens of Lys 102 are positioned over the indole ring of Trp 138. Ring current effects from the tryptophan would lead to an upfield shift of these hydrogens.

Two transitions were observed in nonselective 1D NMR spectra of M102K in D_2O. One occurred coincident with the loss of the upfield lysine resonances. At pH values below 6 and above 4, features of the aromatic and upfield resonances differ from those of wild type. This could arise if the conformation of the protein changed. At pH 2 the 1D NMR spectrum is that of unfolded lysozyme. This spectroscopic result suggests that M102K unfolds between pH 4 and 2. The uptake of protons by M102K detected by the differential titration should then arise from the known uptake of protons that will occur upon unfolding due to changes in titration of carboxyls involved in salt bridges and other interactions (Anderson et al., 1990; Nicholson et al., 1991).

The differential titration of mutant L133D suggests three groups with pK_a's of 6.2, 4.2, and 2.3. Residue 133 in wild-type lysozyme is buried with no opposite charge nearby. It would be expected that an Asp at such a site would have an abnormally high pK_a, assuming the wild-type fold is retained. The pK_a of an aspartate in solution is 3.9 (Stryer, 1975). Thus, the pK_a of Asp 133 is unlikely to be 4.2 and is most probably 6.2. If the pK_a of Asp 133 were 4.2, then the pK_a of 6.2 would have to be assigned to another residue, the most logical alternative being His 31. It is known that His 31 in WT* has a pK_a of about 9.1 and forms a strong salt bridge with Asp 70, whose pK_a is measured to be less than 0.5 (Anderson et al., 1990). If the pK_a of His 31 in the mutant reverts to 6.2, it would indicate that the His 31 to Asp 70 salt bridge was disrupted in the mutant, and the pK_a of Asp 70 observed should thus be around 4. Following this scenario, a total of three protons should be observed in the differential titration of L133D between pH 3 and 7. However, only two protons are observed in both the acid and base titrations. Therefore, we tentatively interpret the pK_a of Asp 133 to be 6.2. Possibly, the two remaining pK_a's of 4.2 and 2.3 might be coupled, representing the pK_a of an acidic group shifted from 2.3 in WT* to 4.2 in L133D, but this is very speculative.

In general, an acidic residue buried in a hydrophobic environment is expected to have a higher pK_a value than one that is solvent-exposed. Similarly, a buried basic residue is expected to have a lower pK_a value than an exposed one. The results obtained here provide rather clearcut evidence that the pK_a of Lys 102 in the M102K mutant is 6.5. The evidence that the pK_a of Asp 133 in the L133D mutant is 6.2 is much weaker. Assuming that Lys 102 in M102K and Asp 133 in L133D have normal pK_a values in the unfolded state, i.e., about 10.8 and 3.9, respectively (Stryer, 1975), then the shifts in pK_a upon folding of M102K and L133D can be estimated as 4.3 and 2.3, respectively, which correspond to free energies of about 5.9 and 3.2 kcal/mol. In the case of M102K the agreement between the energy estimated from the shift in pK_a of Lys 102 (5.9 kcal/mol) agrees quite well with the decrease in stability of the mutant protein at pH 5.3 where the lysine would be expected to be largely protonated ($\Delta\Delta G$ = 6.9 kcal/mol; Table II). It should be remembered, however, that the experimental conditions are not identical. (In particular, the NMR determination is at 0 °C, the thermodynamic estimate is at the T_m of M102K, i.e., 45 °C.) It is in the case of M102K that the structure of the mutant is known and the shift in pK_a is reliable. In the case of L133D the agreement between the two energy estimates is somewhat poorer (3.2 kcal/mol from $\Delta pK_a'$ vs $\Delta\Delta G$ = 5.7 kcal/mol at pH 6.5), and it is also in this case that the pK_a of Asp 133 is not very secure. It is also to be remembered the energy term derived from the shift in pK_a gives an estimate of the cost of protonating (or deprotonating) the charged group in the folded relative to the unfolded protein. In contrast, $\Delta\Delta G$ gives the total destabilization of the mutant protein relative to wild type, which can include terms other than electrostatics.

ACKNOWLEDGMENTS

We thank Sheila Pepiot and Joan Wozniak for excellent technical assistance. Drs. Hale Nicholson, Larry Weaver, and Dale Tronrud also provided invaluable technical advice.

REFERENCES

Alber, T., & Matthews, B. W. (1987) *Methods Enzymol. 154*, 511–533.

Alber, T., Dao-pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., & Matthews, B. W. (1987) *Nature 300*, 41–46.

Anderson, D. E., Becktel, W. J., & Dahlquist, F. W. (1990) *Biochemistry 29*, 2403–2408.

Baldwin, J., & Chothia, C. (1979) *J. Mol. Biol. 129*, 175–220.

Dao-pin, S., Liao, D.-I., & Remington, S. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5361–5365.

Dao-pin, S., Baase, W. A., & Matthews, B. W. (1990) *Proteins: Struct., Funct., Genet. 7*, 198–204.

Dao-pin, S., Alber, T., Baase, W. A., Wozniak, J. A., & Matthews, B. W. (1991a) *J. Mol. Biol. 221*, 647–667.

Dao-pin, S., Sauer, U., Nicholson, H., & Matthews, B. W. (1991b) *Biochemistry 30*, 7142–7153.

Dao-pin, S., Soderlind, E., Baase, W. A., Wozniak, J. A., Sauer, U., & Matthews, B. W. (1991c) *J. Mol. Biol.* (in press).

Gelin, B. R., & Karplus, M. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 801–805.

Gilson, M. K., & Honig, B. H. (1987) *Nature 330*, 84–86.

Gilson, M. K., Rashin, A., Fine, R., & Honig, B. (1985) *J. Mol. Biol. 184*, 503–516.

Gilson, M. K., Sharp, K. A., & Honig, B. (1987) *J. Comput. Chem. 9*, 327–335.

Johnson, M. L., & Frasier, S. G. (1985) *Methods Enzymol. 117*, 301–342.

Klapper, I., Hagstrom, R., Fine, F., Sharp, K., & Honig, B. (1986) *Proteins: Struct., Funct., Genet. 1*, 47–59.

Kossiakoff, A. A. (1983) *Annu. Rev. Biophys. Bioeng. 12*, 159–182.

Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331–358.

Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367–382.

Matsumura, M., Wozniak, J. A., Dao-pin, S., & Matthews, B. W. (1989) *J. Biol. Chem. 264*, 16059–16066.

Matthews, B. W. (1987) *Biochemistry 26*, 6885–6887.

Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., & Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44–73.

Nicholson, H., Anderson, D. E., Dao-pin, S., & Matthews, B. W. (1991) *Biochemistry* (in press).

Parsons, S. M., & Raftery, M. A. (1972) *Biochemistry 11*, 1623–1629.

Perutz, M. F. (1978) *Science 201*, 1187–1191.

Pjura, P., Matsumura, M., Wozniak, J., & Matthews, B. W. (1990) *Biochemistry 29*, 2592–2598.

Poteete, A. R., Dao-pin, S., Nicholson, H., & Matthews, B. W. (1991) *Biochemistry 30*, 1425–1432.

Rashin, A. & Honig, B. (1984) *J. Mol. Biol. 173*, 515–521.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.

Sternberg, M. J. E., Hayes, R. F. F., Russell, A. J., Thomas,

P. G., & Fersht, A. R. (1987) *Nature 330*, 86–88.
Stryer, L. (1975) *Biochemistry*, W. H. Freeman and Co., New York.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489–503.
Tsugita, A., Inouye, M., Terzaghi, E., & Streisinger, G. (1968) *J. Biol. Chem. 243*, 391–397.
Wada, A., Nakamura, H., & Sakamoto, T. (1985) *J. Phys. Soc. Jpn. 54*, 4042–4046.
Weaver, L. H., & Matthews, B. W. (1987) *J. Mol. Biol. 193*, 189–199.
Zuiderweg, E. R. P. (1990) *J. Magn. Reson. 86*, 346–357.

Alboaggregin-B: A New Platelet Agonist That Binds to Platelet Membrane Glycoprotein Ib†

Manling Peng,‡ Weiqi Lu,§ and Edward P. Kirby*,‡

Department of Biochemistry, Thrombosis Research Center, and Department of Physiology, Temple University School of Medicine, Philadelphia, Pennsylvania 19140

Received November 19, 1990; Revised Manuscript Received August 19, 1991

ABSTRACT: A new protein, called alboaggregin-B (AL-B), has been isolated from *Trimeresurus albolabris* venom by ion-exchange chromatography. It agglutinated platelets without the need for Ca^{2+} or any other cofactor. The purified protein showed an apparent molecular mass on SDS–PAGE and gel filtration of about 23 kDa under nonreducing conditions. Ristocetin did not alter the binding of AL-B to platelets or affect AL-B-induced platelet agglutination. Agglutinating activity was not dependent on either proteolytic or lectin-like activity in AL-B. Binding analysis showed that AL-B bound to platelets with high affinity (K_d = 13.6 ± 9.3 nM) at approximately 30 800 ± 14 300 binding sites per platelet. AL-B inhibited the binding of labeled bovine von Willebrand factor (vWF) to platelets. Monoclonal antibodies against the 45-kDa N-terminal domain of platelet glycoprotein Ib inhibited the binding both of AL-B and of bovine vWF to platelets, and also inhibited platelet agglutination induced by AL-B and bovine vWF. Specific removal of the N-terminal domain of GPIb by treatment of the plaelets with elastase or *Serratia marcescens* protease reduced the binding of labeled AL-B and bovine vWF to platelets and blocked platelet agglutination caused by both agonists. Monoclonal antibodies to glycoprotein IIb/IIIa, to bovine vWF, and to bovine serum albumin did not show any effect on the binding of AL-B to platelets. Our results indicate that the binding domain for AL-B on platelet GPIb is close to or identical with the one for vWF. This new protein may be a very useful tool for studying the interaction between platelets and vWF.

When blood vessels are damaged, von Willebrand factor (vWF)[1] is needed for platelets to recognize the damaged vascular endothelium and to form aggregates upon it. In vitro, the function of human vWF is assayed by its capacity to agglutinate platelets in the presence of the nonphysiological agonist ristocetin (Howard & Firkin, 1971). Bovine vWF agglutinates human platelets directly without the need for ristocetin (Forbes & Prentice, 1973). Agglutination of human platelets by bovine vWF, and by human vWF in the presence of ristocetin, is inhibited by antibodies to platelet membrane glycoprotein Ib (GPIb) (Tobelem et al., 1976). Bovine vWF and human vWF bind to platelets in a reversible manner and compete for binding to platelets (Suzuki et al., 1980).

Glycoprotein Ib is one of the major platelet membrane glycoproteins. There are approximately 25 000 copies of GPIb per platelet (Coller et al., 1983). This glycoprotein contains two disulfide-linked subunits, GPIbα (M_r 145K) and GPIbβ (M_r 24K), that are complexed with glycoprotein IX (M_r 22K) (Clemetson, 1985; Berndt et al., 1983, 1985). Two regions on GPIbα have been identified that are sensitive to proteolytic cleavage (Cooper et al., 1981; Berndt et al., 1986). One region is susceptible to hydrolysis by the platelet Ca^{2+}-dependent protease (calpain), and by *Serratia marcescens* protease. Cleavage at this site generates an N-terminal 135-kDa fragment termed glycocalicin, and a 25-kDa fragment that is disulfide-linked to GPIbβ. Another protease-sensitive region on GPIbα is located nearer to the N-terminal. Hydrolysis by elastase or by trypsin generates a soluble 45-kDa amino-terminal fragment, and a 100-kDa fragment that remains disulfide-linked to GPIbβ in the platelet membrane (Wicki & Clemetson, 1985). The external 45-kDa fragment may contain the binding sites for thrombin and vWF (Wicki & Clemetson, 1985; Kao et al., 1979; Handa et al., 1986). Proteolytic degradation of GPIb on platelet membranes is paralleled by a loss of vWF binding ability (Wicki & Clemetson, 1985; Coller, 1983) and a loss of responsiveness to bovine vWF and ristocetin (Cooper et al., 1977). Treatment of platelets with trypsin causes the GPIb–IX complex to be cleaved into four parts—an N-terminal domain (M_r 45K), a macroglycopeptide (M_r 85K), a remnant of the α-chain disulfide-linked to the β-subunit (M_r 41K), and undigested GPIX (M_r 22K) (Berndt et al., 1988; Du et al., 1987). Monoclonal antibodies AP1 and AK2, which are directed against the 45-kDa N-terminal

† This investigation was supported by National Institutes of Health Grant HL 27993.

* Author to whom correspondence should be addressed.
‡ Department of Biochemistry, Thrombosis Research Center.
§ Department of Physiology.

[1] Abbreviations: AL-B, alboaggregin-B; BSA, bovine serum albumin; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; FWP, formalin-fixed washed platelet(s); GPIb, glycoprotein Ib; PRP, platelet-rich plasma; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; Tris, tris(hydroxymethyl)aminomethane; vWF, von Willebrand factor; WP, washed platelet(s).

0006-2960/91/0430-11529$02.50/0