Stabilizing Amino Acid Replacements at Position 52 in Yeast Iso-1-cytochrome *c*: *In Vivo* and *in Vitro* Effects

Lisa I. Linske-O'Connell,[†,§] Fred Sherman,[†] and George McLendon*[,†,‡]

Department of Biochemistry, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642, and Department of Chemistry, University of Rochester, Rochester, New York 14627

Received July 13, 1994; Revised Manuscript Received February 6, 1995[⊗]

ABSTRACT: Position 52 of iso-1-cytochrome *c* in the yeast *Saccharomyces cerevisiae* was systematically replaced with all possible amino acids to investigate the molecular basis of the "global suppressor" activity for the N52I mutation. Isogenic strains containing the variant proteins were made with a mixed oligonucleotide-directed mutagenesis technique *in vivo*. A relationship between thermodynamic protein stability and cellular protein levels was established by comparing direct spectroscopic measurements of cytochrome *c* levels *in vivo* with the thermodynamic parameters from guanidine hydrochloride denaturation and microcalorimetric measurements. Reversible denaturation data show that, within a large group of amino acid side chains, the thermodynamic stabilization at position 52 is related to the hydrophobicity of the side chain at that site.

. One of the goals of protein science is to create an algorithm that is capable of taking the primary amino acid sequence and folding it "in computo" into the proper three-dimensional structure. For each amino acid in a protein, there are av ariety of forces that could lead to stabilization or destabilization. Hydrophobic interactions are thought to be among the strongest driving forces in protein folding. Thus, quantification of this effect is important in understanding its role in producing a stable structure.

There is a relatively small difference of about 5−15 kcal/mol (Creighton, 1984) in the free energy, ΔG_U, between a folded and unfolded protein. Most amino acid changes at evolutionarily conserved sites are either neutral or lead to a decrease in stability (Pakula & Sauer, 1989; Hampsey *et al.*, 1986; Hampsey *et al.*, 1988). Stabilizing amino acid replacements are less common. Generally, they have a comparatively small net increase in stability, $\Delta\Delta G_U$. By contrast, the N52I replacement of iso-1-cytochrome *c* results in an unusually large increase in the protein's stability. This single-amino acid change doubles the stability over that of the wild type protein in chemical and thermodenaturation studies (Das *et al.*, 1989; Hickey *et al.*, 1991a). Yeast strains containing the N52I cytochrome *c* exhibit increased levels of protein *in vivo*. Pulse-chase experiments on the wild type and N52I cytochromes *c* show that the increased levels *in vivo* appear to result from the increase in the intracellular half-life of the protein (Pearce and Sherman, manuscript in preparation). Thus, this "two-point" correlation suggests a relationship between fundamental physiochemical stability as measured *in vitro* and steady state levels *in vivo*. A broad range of variants would be helpful to test this relationship.

The crystal structure of wild type and N52I variant of the oxidized form of iso-1-cytochrome *c* shows that position 52 is 100% buried as measured by the Access program (Lee & Richards, 1971). The measurements were done using a solvent molecule with a radius of 1.4 Å. A comparison of crystallographic coordinates shows that there is little change in the overall α-carbon backbone between N52 and N52I (Hickey *et al.*, 1991a; Louie & Brayer, 1990; Berghuis & Brayer, 1992). At the site of the replacement, Wat-166, an internal water molecule, is missing from the isoleucine structure (Figure 1). There is a change in the hydrogen-bonding network in N52I that results in a net *decrease* in the number of H bonds. This area of the protein has high average crystallographic thermal *B* factors. Wat-166 has also been shown to undergo a large 1.7 Å displacement when

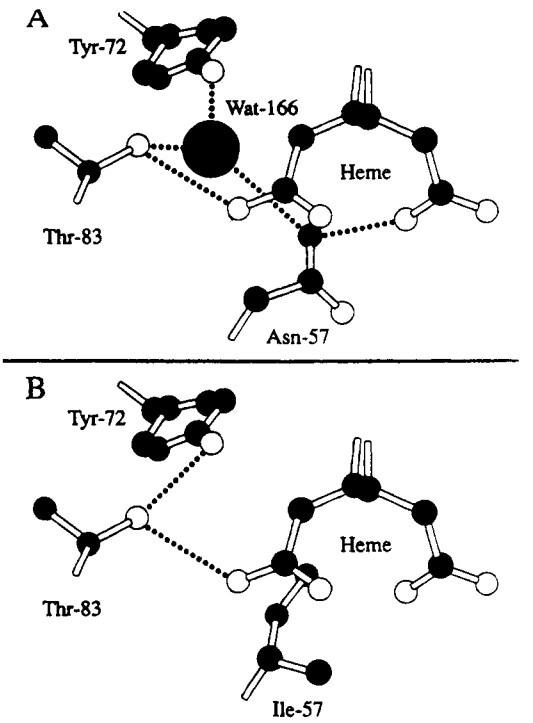

FIGURE 1: X-ray crystal structure of position 52. The ball and stick representation of the heme and amino acids near position 52 in iso-1-cytochrome *c*: (A) the wild type N52 showing the internal water molecule 166 and the hydrogen-bonding network and (B) the stabilized N52I with the loss of a water molecule and change in the hydrogen-bonding network.

[†] University of Rochester School of Medicine and Dentistry.
[‡] University of Rochester.
[§] Current address: ABL, NCI-FCRDC, P.O. Box B, Frederick, MD 21702-1201.
[⊗] Abstract published in *Advance ACS Abstracts,* May 15, 1995.

0006-2960/95/0434-7094$09.00/0

the redox state is changed (Berghuis & Brayer, 1992). This means there is a high degree of mobility and thus flexibility at the site of amino acid replacement. When the thermodynamic parameters from the wild type, N52I, and N52A proteins were compared, there appeared to be a linear relationship between the hydrophobicity of the amino acid side chain and the stability *in vitro* for this limited number of variants (Hickey *et al.*, 1991a).

To further understand the unusually large increase in stability at position 52, a series of isogenic strains was constructed using oligonucleotide-directed mutagenesis *in vivo* for all 20 amino acids at position 52. The levels of cellular cytochromes *c* and growth characteristics on nonfermentable carbon sources were measured. Of these strains, 11 had *in vivo* levels about equal to or higher than that of the wild type. These proteins were isolated and measured for thermodynamic parameters using guanidine hydrochloride denaturation and microcalorimetric analysis. A frequent assumption when studying protein stability *in vitro* is that the results are physiologically significant for the protein *in vivo*. In fact, there is little work concerning physiological protein folding, particularly in regard to protein turnover. (McLendon, 1977; McLendon & Radany, 1978; Drummond *et al.*, 1991; Leontiev *et al.*, 1993; Pakula *et al.*, 1986; Parcell & Sauer, 1989; Hecht *et al.*, 1984.) There are a variety of additional forces known to regulate a protein's stability *in vivo* (Jaenick, 1993). Although there are known exceptions for cytochrome *c* (Pearce and Sherman, manuscript in preparation), for the present set of replacements, the stability *in vivo* appears directly related to the thermodynamic stability *in vitro*.

MATERIALS AND METHODS

Genetic Nomenclature and Strain Construction. CYC1 codes for iso-1-cytochrome *c* in the yeast *Saccharomyces cerevisiae*. In the work discussed, all functional cytochromes *c* have the C102A replacement (as is seen in iso-2-cytochrome *c* at that position). This replacement prevents dimerization of the isolated protein that would complicate analysis. A cysteine to alanine substitution at that site should have a similar effect on each of the protein's stability. This change eliminates the need for chemically blocking C102, which may lead to a destabilization in the protein (Hickey *et al.*, 1991a). *CYC1-XXXX* denotes an altered functional cytochrome *c*. The nomenclature *cyc1-XXXX* stands for a nonfunctional cytochrome *c*. All haploid strains are isogenic with B-7628 (*MAT***a** *cyc1-31 cyc7-67 ura3-52 lys5-10*), except at the iso-1-cytochrome *c* locus (Moerschell *et al.*, 1988). The *cyc7-67* allele contains a major deletion in the iso-2-cytochrome *c* gene. Therefore, all effects seen in the strains produced are solely due to the *CYC1* gene product. The haploid series of position 52 replacements was crossed to B-7682 (*MAT*α *cyc1-31 cyc7-67 his1 trp2-1*), which contains no functional cytochromes *c*. These diploid strains were constructed to eliminate any recessive mutations that may have occurred during the transformation process that would affect growth.

Oligonucleotide-Directed Mutagenesis in Vivo. The parent strain, B-8162 (*MAT***a** *cyc1-1026 cyc7-67 ura3-52 lys5-10*), was constructed using classical site-directed mutagenesis techniques (Kunkel, 1987). The plasmid pAB569 (containing the *Bam*H1-*Hind*III fragment of *CYC1* with C102A) was modified using oligonucleotide OL90.069 that contained the frame shift and stop at position 52. The resulting nonfunctional gene (*cyc1-1026*) was substituted in strain B-7528 at the *CYC1* locus using methods previously described to produce B-8162 (Fetrow *et al.*, 1989). The integrity of the integrated cytochrome *c* was checked by using PCR amplification to obtain DNA for sequencing the chromosomal *CYC1* allele.

Site-directed oligonucleotide mutagenesis *in vivo* was carried out in yeast using the technique of Yamamoto *et al.* (1992). A series of four degenerate oligonucleotides was made that could correct the frame shift and stop at position 52 and insert all 20 amino acids (Figure 2). The transformations were done at 30 °C using 100 μg of oligonucleotide with ~1 × 10^6 cells. This method relies on the production of a strain with a functional cytochrome *c* that can outgrow the confluent growth of the parent strain on nonselective media. After 4 days, random transformants were picked and streaked for single colonies. The strains were put into groups depending on the levels of cytochrome *c* as determined by visual low-temperature spectra and growth characteristics. DNA from selected colonies was amplified by PCR and sequenced in the region of position 52. Amino acid residues not recovered in the first experiment were produced with a second round of oligonucleotide transformations. These transformants were incubated up to 8 days to recover slower growing colonies. Multiple isolates of the same amino acids were judged to have similar cytochrome *c* levels *in vivo* as scored by blind study of visual low-temperature spectra. All 20 amino acid substitutions were recovered by this method. The arginine 52 replacement showed reversion to a more active form of cytochrome *c*. It was dropped from further analysis. This method gives a single copy of the desired gene stably integrated at the *CYC1* locus, so that it is under the same control as the wild type gene. Any differences among these isogenic strains should, therefore, be attributable to the change at position 52.

Cytochrome c Levels in Vivo. To help differentiate the higher than wild type levels, diploids were made with a single copy of the mutated *CYC1* gene as the only source of functional cytochrome *c*. The isogenic haploid strains were crossed with B-7682 (*MAT*α *cyc1-31 cyc7-67 his1 trp2-1*). Strains were grown on YPE[1] media which contains ethanol as a carbon source at 30 °C for 3−4 days (Sherman *et al.*, 1987). A low-temperature (−196 °C) scanning spectrophotometer was used to measure the relative absorbance of cytochromes *c* and *b* in each of the strains (Hickey *et al.*, 1991b). Briefly, the cells were scraped from plates and blotted dry. The yeast paste was mounted in a sample holder with a 1 mm path length to give a quantitative number of cells. The samples were frozen in liquid nitrogen. The spectrophotometer was equipped with an unsilverred Dewar's flask containing liquid nitrogen. This way the sample is maintained in a frozen state while being scanned. Scans were done from 500 to 750 nm with a wavelength step size of 0.2 nm, a bandwidth of 0.4 nm, and an averaging time of 0.4 s. The cytochrome *c* peak was measured at 546 nm. Cytochrome *b* (558 nm) was used as an internal standard. A cytochrome *c*/*b* ratio was taken. A strain containing no iso-

[1] Abbreviations: YPD, yeast extract, peptone, and dextrose media; YPE, yeast extract, peptone, and ethanol media; SL+13, lactate and 13 essential amino acid media; GuHCl, guanidine hydrochloride.

A

```
    45                  50                  55                  60      102
Glu-Gly-Tyr-Ser-Tyr-Thr-Asp-Ala-Asn-Ile-Lys-Lys-Asn-Val-Leu-Trp-Asp    -Cys-   CYC1
GAA GGG TAT TCG TAC ACA GAT GCC ATT ATC AAG AAA AAC GTG TTG TGG GAC     TGT

                                                                       -Ala-   cyc1-1026
GAA GGG TAT TCG TAC ACA GAT GCC TAΔ AAG AAA AAC GTG TTG TTG TGG GAC     GCT

      G TAT TCG TAC ACA GAT GCC NNN AAG AAA AAC GTG TTG T                      Oligonucleotide
                                ⌣
                            Codon 52
```

B

Oligonucleotide	Codon 52	Possible Amino Acids	Recovered Amino Acids
Experiment 1			
OL90.1298	BNN	F,L,S,Y,C,W,L,P,H,Q,R,V,A,D,G,E	L,V,A,G,E,D,R
OL90.2299	ACN	T,N,K,S,R	T,S,N,K
OL90.3154	ATG	M	M
OL90.1067	ATT	I	I
Experiment 2			
OL92.7	TGG	W	W
OL92.8	ATT	I	I
OL92.9	TTC	F	F
OL92.10	TAC	Y	Y
OL92.11	TGT	C	C
OL92.17	CAA	Q	Q
OL92.18	CAC	H	H
OL92.19	CCA	P	P

FIGURE 2: Mixed oligonucleotide-directed site specific mutagenesis. (A) The protein and DNA sequence of the region flanking positions 52 and 102 in iso-1-cytochrome *c*, the corresponding sequence of the frame shift and stop at position 52 and C102A of the parental strain with *cyc1-1026*, and the general sequence of the oligonucleotides used for transformation. (B) The DNA sequence of each oligonucleotide for codon 52 (N = A, T, G, C; B = C, G, T) and the amino acids each oligonucleotide encoded and the amino acids recovered from each experiment.

1-cytochrome *c* is normally subtracted from all strains. Since ethanol media does not support the growth of strains without a functional electron transport chain, this could not be done. Within 20% of 100% cytochrome *c* levels, which includes most of the replacements, there would be less than a 5% effect on the result. As the protein level approaches 0%, the cytochrome *c* level would be increasingly overestimated. A mutant with 60% cytochrome *c* would be overestimated by about 10%. All the scans taken had a standard deviation of 3.2% of the average cytochrome *b* level. This deviation probably reflects slight differences in the thickness of the packed cells rather than any difference in the oxidation state of the strains, which would affect the results. Supporting this idea, repeated samples with slightly different cytochrome *b* levels gave very similar *c/b* ratios. The wide type control was repeated six times for a standard deviation of 3.3% in the *c/b* ratio. The cytochrome *c* levels were normalized to the diploid strain with N52 C102A (B-8450) as 100%. Cytochrome *c* levels were originally screened by visual examination by a Hartree low-dispersion spectroscope (Sherman & Slonimski, 1964; Sherman *et al.*, 1974). The strains with greater than normal levels were easily found in double-blind visual examination. The relative levels of the visually estimated cytochromes *c* for all the strains assessed matched those measured with the scanning spectrophotometer (unpublished results).

Growth. The haploid and diploid strains containing the position 52 replacements and the controls were spotted on nonselective media containing glucose (YPD_1), as well as selective media containing the nonfermentable carbon sources glycerol, lactate (SL + 13), and ethanol (YPE) as outlined in Sherman *et al.* (1987). These plates were incubated for 3 days at the standard temperature of 30 °C as well as elevated temperatures of 35 and 37 °C.

Protein Preparation. Yeast strains were grown in fermentors and permeabilized using the ethyl acetate/NaCl method of Sherman *et al.* (1968). The procedure was followed through the first CG-50 Amberlite column and dialysis against carbonate buffer. The proteins were then purified with an S Sepharose column in a Pharmacia FPLC system. Buffer A was 0.1 M potassium phosphate buffer, pH 7.2. A 12−30% gradient of buffer B (0.1 M potassium phosphate buffer, pH 7.2, with 0.8 M KCl) was used to elute the protein. If the cytochrome *c* was isolated as a mixture of oxidized and reduced cytochrome *c*, the first minor peak of, presumably, deaminated protein was discarded and the second (reduced) and third (oxidized) peaks were pooled. Most preparations were first oxidized with potassium ferricyanide that was removed by passage through a DEAE cellulose column before the S Sepharose separation. For these proteins, the first peak was discarded and the second oxidized peak was collected. The samples were concentrated

in an Amicon stirred ultrafiltration device with a YM30 membrane. The buffer was exchanged at least three times with 0.1 M potassium phosphate buffer, pH 7.0. All samples were shown to be pure with a single band on an SDS−PAGE gel and an A_{410}/A_{280} ratio of about 4.57 for the oxidized protein. The samples were flash frozen in liquid nitrogen and kept at −70 °C until used.

Guanidine Denaturation. A stock solution of 6.0 M guanidine hydrochloride in 0.1 M potassium phosphate, pH 7.0, was made. The concentration was verified using the refractive index of the solution at the sodium D line with a Bausch and Lomb Abbe refractometer (Pace, 1986). The proteins were quickly thawed in warm water. Reduced cytochromes *c*, as judged by the presence of absorbance peaks at 550 nm, were oxidized as described above. Each cytochrome *c* sample was diluted to 30 μM in 0.1 M potassium phosphate buffer, pH 7.0. A series of guanidine hydrochloride concentrations was made from the stock solution. The protein was added to a final concentration of 2 μM cytochrome *c* in a 300 μL volume. The samples were allowed to reach equilibrium at room temperature for about 15 min. The amount of unfolded protein was monitored on a Perkin-Elmer MPF-2A fluorometer (Perkin-Elmer Cetus Instruments) (excitation at 292 nm, emission at 350 nm, slit widths at 10 nm). The proteins showed reversible denaturation except for the N52C replacement. This probably formed cystine dimers in the unfolded state. It was dropped from further investigation.

A two-state denaturation was assumed for analysis. The percentage of unfolding versus concentration on denaturant was plotted.

$$\% \text{ unfolded} = 100[(f - f_o)/(f_u - f_o)] \quad (1)$$

f is the fluorescence at each denaturant concentration, and f_o and f_u are the fluorescence at zero concentration and the highest concentration of guanidine hydrochloride, respectively. The data were analyzed using a nonlinear least squares fitting method (Santoro & Bolen, 1988). The equation was modified to give the resulting eq 2:

$$y = [(m_N{*}D) + [100 + m_U(D - M_D)] \exp[-(\Delta G_U + m_G{*}D/RT)]]/1 + \exp[-(\Delta G_U + m_G{*}D/RT)] \quad (2)$$

The program used was written by Karlene Gunter (Department of Biophysics, University of Rochester). The percent unfolding (y) is related to the concentration of denaturant (D) through the fitting parameters m_N, m_U, ΔG_U, and m_G. One hundred percent unfolding was assumed at the highest denaturant concentration used (M_D). The terms m_N and m_U represent the slopes of the native and unfolded proteins, respectively. The Δe°_N and Δe°_U terms from the original paper, representing the intercept of the native and unfolded base lines, respectively, were replaced in the equation by 0% and the difference from 100% unfolding. ΔG_U and m_G are the change in free energy and the slope of unfolding as represented by a linear extrapolation of the unfolding free energies versus denaturant concentration. Duplicate experiments were done on the normal N52 and stabilized N52I proteins. Both proteins gave three standard deviations of 0.03 M for the C_m values. Three standard deviations of as a percentage of the ΔG_U were 5% and 3% for N52 and N52I, respectively. Each protein can be considered to have an estimated error of ±0.03 M for the C_m and ±5% for the ΔG_U.

Microcalorimetry. A Microcal DSC2 scanning microcalorimeter was used to measure thermodynamic parameters of the isolated cytochromes *c*. The oxidized proteins in 50 mM potassium acetate buffer at pH values between 3.5 and 5.0 were scanned at concentrations between 100 and 160 μM (micromolar extinction coefficient A_{410} of 0.1061). A pH of 4.6 was used to make measurements comparable to other published values for cytochromes *c* (Privalov, 1979; Betz & Pielak, 1992). Scans were run from 15 to 80 °C at a scan rate of 60 °C/h. Scans were successively repeated two or three times. The protein concentration was measured after the last scan. A decrease in cytochrome *c* as measured by A_{410} was between 5% and 10% for each run. The decrease was assumed to be linear. The values were normalized for this decrease. Calculations were done using the Origin program (version 2.9) supplied with the instrument. The estimated error for T_m of unfolding was well within 0.5 °C of the average for each of the proteins. The calculated ΔH values were all within 5% of the average. The experiments run at pH 4.6 were used for calculating ΔG_U. Since $\Delta G = \Delta H - T\Delta S$ and at equilibrium $\Delta G = 0$, the entropy was calculated by eq 3:

$$\Delta S = \Delta H_{cal}/T_m \quad (3)$$

The ΔG_U was calculated from eq 4 (Shortle, 1989):

$$\Delta G_U(T) = \Delta H_{cal} + \Delta C_p(T - T_m) - T\Delta S - T\Delta C_p \ln(T/T_m) \quad (4)$$

T is 300 K. ΔG_U is the change in free energy at temperature T. T_m is the midpoint temperature of the transition. ΔH°_{cal} and ΔS° are the enthalpy and entropy change at T_m as measured by microcalorimetry. ΔC_p is the change in heat capacity between native and unfolded states.

RESULTS

Degenerate Oligonucleotide-Directed Mutagenesis in Vivo. To further investigate the striking thermostabilization of the N52I amino acid replacement in iso-1-cytochrome *c*, a complete series of strains with position 52 substitutions was made (Figure 2, Table 1). Experiments have shown that using oligonucleotide to directly transform chromosomal DNA is a powerful method for the production of functional variants of yeast iso-1-cytochrome *c* (Moerschell *et al.*, 1988, 1990; Yamomoto *et al.*, 1992). This work at position 52 and other work at position 6 (accompanying manuscript) show that using oligonucleotides with mixed bases at the site of interest is a convenient and cost efficient variation of that method. Theoretically, it would be possible to recover all functional changes in a single experiment with only a few oligonucleotides. (These transformations would have recovered more amino acid replacements if the original transformation plates had been incubated for a longer period to recover slower growing colonies.)

Low-Temperature Cytochrome c Levels and Growth. The relative amounts of holocytochrome *c* in the intermitochondrial space as measured by a spectrophotometer are shown in parentheses in Figure 3. The *x* axis shows the average volume of a buried residue (Richards, 1977). The *y* axis

Table 1: Amino Acid Replacements at Position 52

allele	strain no. haploid	strain no. diploid	amino acid-52	codon 52
CYC1	B-8114	B-8540	N	AAT
CYC1-1163	B-8435	B-8519	W	TGG
CYC1-1126	B-8436	B-8520	I	ATT
CYC1-1164	B-8437	B-8521	F	TTC
CYC1-1165	B-8438	B-8522	L	TTG
CYC1-1166	B-8439	B-8523	M	ATG
CYC1-1167	B-8440	B-8524	V	GTG
CYC1-1168	B-8841	B-8525	C	TGT
CYC1-1169	B-8442	B-8526	Y	TAC
CYC1-1170	B-8443	B-8527	P	CCA
CYC1-1171	B-8444	B-8528	A	GCA
CYC1-1172	B-8445	B-8529	T	ACT
CYC1-1173	B-8446	B-8530	H	CAC
CYC1-1174	B-8447	B-8531	G	GGT
CYC1-1175	B-8448	B-8532	S	AGT
CYC1-1176	B-8449	B-8533	Q	CAA
CYC1-1112	B-8450	B-8534	N	AAT
CYC1-1177	B-8451	B-8535	E	GAG
CYC1-1178	B-8452	B-8536	D	GAC
CYC1-1179	B-8453	B-8537	K	AAG
CYC1-1180	B-8454	B-8538	R	CGA
cycl-1026	B-8162	B-8539	fs/s[a]	TAΔA

[a] fs/s—frame shift and stop.

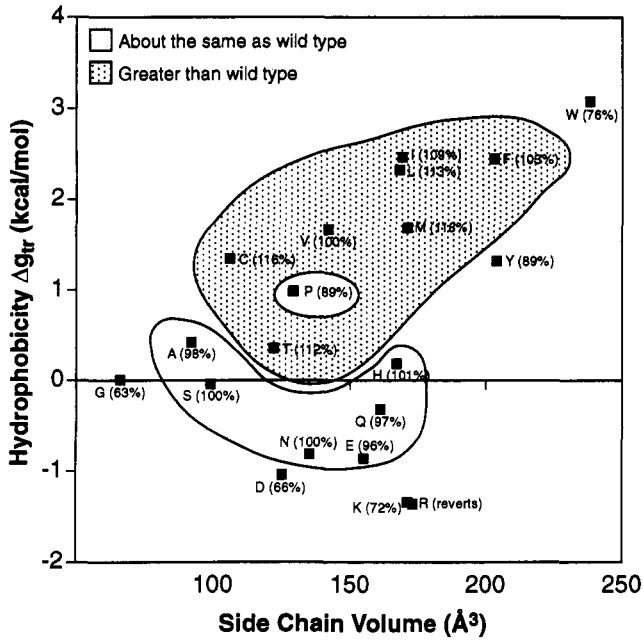

FIGURE 3: Relative levels of cytochrome *c in vivo* versus hydrophobicity and amino acid volume. The levels of cytochrome *c in vivo* relative to the native protein are shown in parenthesis (see Materials and Methods). The side chain size is measured by volume of the residue (Richards, 1977; Zamyatnin, 1972). The hydrophobicity is measured by Δg_{tr} (octanol/water). The value is *p*, reported by Fauchere and Pliska (1983), multiplied by 2.30*RT* = 1.36 kcal/mol. *p* represents the log *D*(acetyl amino acid amide) − log *D*(acetyl glycine amide). *D* is the coefficient of distribution between octanol and water.

shows the hydrophobicity as measured by Δg_{tr} (octanol/water) for each of the amino acids (Fauchere & Pliska, 1983). As will be shown by the thermodynamic experiments, *vide infra*, the levels *in vivo* are related to the stability of the protein. The strains can be divided into three categories based on their cytochrome *c* levels. Strains with lower protein levels have proline, glycine, and amino acids that are large or charged at position 52. Strains containing

Table 2: Guanidine Hydrochloride Denaturation Thermodynamic Parameters

cytochrome *c* (C102A)	ΔG_U (kcal/mol)	m_G (kcal L/mol²)	C_m (M)	$\Delta\Delta C_m$ (M)	$\Delta\Delta G_U$ (kcal/mol)
N52	5.86	−4.36	1.34	0.00	0.00
N52Q	5.02	−3.61	1.39	0.05	−0.84
N52S	5.49	−3.95	1.39	0.05	−0.37
N52H	4.16	−2.95	1.41	0.07	−1.70
N52T	7.76	−4.91	1.58	0.24	1.90
N52A	8.83	−5.50	1.52	0.18	2.97
N52V	7.88	−4.44	1.77	0.43	2.02
N52M	9.90	−4.62	2.14	0.80	4.04
N52L	10.59	−4.87	2.17	0.83	4.73
N52F	5.80	−3.42	1.70	0.36	−0.06
N52I	9.96	−4.77	2.09	0.75	4.10

proteins with about the same protein levels as wild type (C102A) include N52A, N52S, N52Q, and N52H. The remaining strains had cytochrome *c* levels higher than N52. Of the last two categories, there is a general increase in cytochrome *c* levels as the hydrophobicity of the amino acid at position 52 increases.

Recessive changes at other sites that occurred during the transformation process can be ruled out. The growth of the haploid cells was very similar to that of the diploid cells except in strains that had the lowest levels of cytochrome *c*. This low-level difference is probably attributable to a gene dosage effect rather than any recessive mutation. Most strains grew to the same colony size as the wild type at different temperatures on different nonfermentable carbon sources. Work with destabilizing changes at position 6 (following paper) shows that cytochrome *c* levels fall below 75% before the colony size on nonfermentable carbon sources was affected. At 30 °C only the positively charged N52K and N52R grew less well than the N52 strain. At 35 and 37 °C incubation temperatures, there was also a decrease in the growth of the negatively charged amino acids on lactate media (unpublished results). Work with the destabilized N52G protein (Hickey *et al.*, 1991a) suggests that isolated cytochromes *c* with 80% levels *in vivo* can be unstable at room temperature. Therefore, only the 12 proteins that had stabilities *in vivo* equal to or greater than that of the wild type N52 protein were isolated and subjected to thermodynamic measurements.

Guanidine Hydrochloride Denaturation. The C_m values (concentration at the midpoint of unfolding) for N52, N52A, and N52I correlate with those calculated by Hickey *et al.* (1991a) within experimental error (Table 2). The calculation of ΔG_U can depend on the method used (Pace, 1975). The slopes and ΔG_U of unfolding for these three proteins agree with those of Hickey *et al.* (1991a) when analyzed by the same method (Schellman, 1978). The nonlinear least squares method used in these calculations gave larger slopes and m_G and ΔG_U values.

Unfolding experiments with the N52C variant gave an unstable unfolded base line. This is possibly caused by a cysteine dimerization. Because this invalidates the free energy calculations, this protein ws dropped from further analysis. Three typical denaturation curves with a line representing the best fit with the nonlinear least squares calculation are shown in Figure 4. Results of this analysis are shown in Table 2. ΔG_U values for unfolding are plotted against the hydrophobicity of the side chain at position 52 (Figure 5). The correlation coefficient for guanidine hydro-

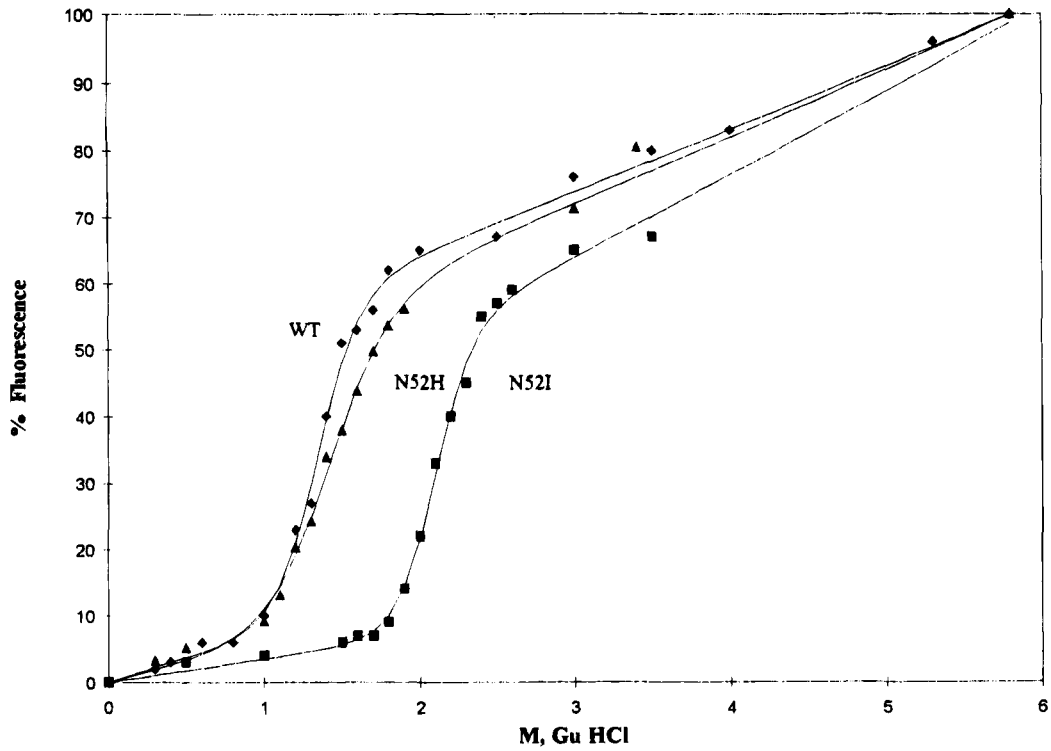

FIGURE 4: Guanidine hydrochloride denaturation of position 52 replacements. The relative fluorescence measurements of unfolding for the least stable protein N52H (▲), the wild type N52 (◆), and the stable protein N52I (■) at pH 7.0, 22 °C, with the best fit calculated by eq 2.

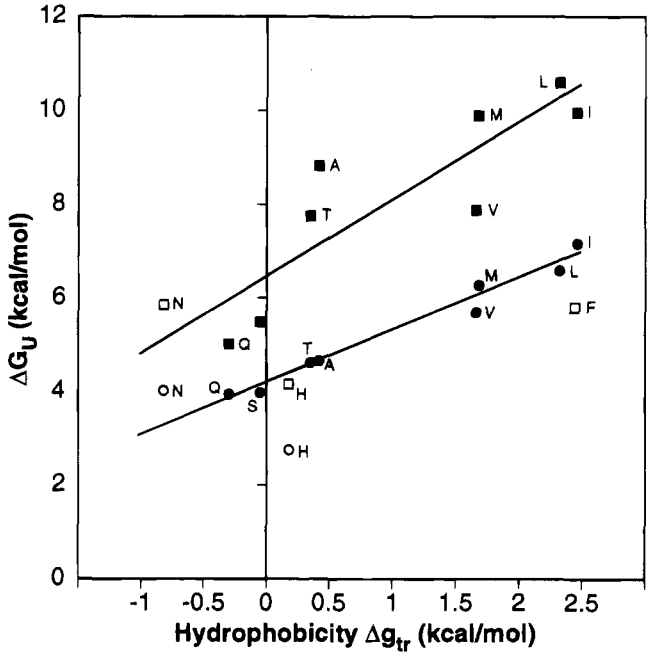

FIGURE 5: Free energy of unfolding versus hydrophobicity. The free energy of unfolding (ΔG_U) as measured by guanidine hydrochloride denaturation at pH 7.0 (■) and microcalorimetry at pH 4.6 (●) for position 52 replacements of iso-1-cytochrome *c* are plotted against the hydrophobicity as measured by Δg_{tr} (octanol/water). Linear regressions of the filled points are shown. Correlation coefficients are $R^2 = 0.86$ for guanidine hydrochloride denaturation and $R^2 = 0.99$ for microcalorimetry.

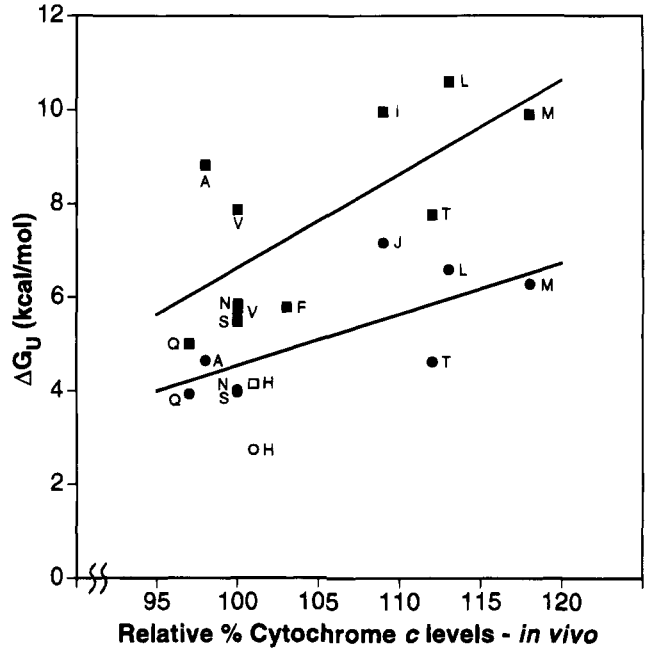

FIGURE 6: Free energy of unfolding versus relative % cytochrome *c*. The free energy of unfolding (ΔG_U) as measured by guanidine hydrochloride denaturation at pH 7.0 (■) and microcalorimetry at pH 4.6 (●) for position 52 replacements of iso-1-cytochrome *c* are plotted against the relative levels of cytochrome *c* with N52 as 100%. Linear regressions of the filled points are shown. Correlation coefficients are $R^2 = 0.71$ for guanidine hydrochloride denaturation and $R^2 = 0.68$ for microcalorimetry.

chloride denaturation excluding N52, N52H, and N52F is $R^2 = 0.86$ for ΔG_U. A higher correlation is found with C_m: $R^2 = 0.95$. The lower correlation with the ΔG_U calculations may reflect the inaccuracy inherent in determining the change in free energy as well as the possible differences in the slopes of the denaturation curves or the differences in the folded or unfolded states of the proteins.

Figure 6 shows the relationship between the free energy of protein unfolding measured *in vitro* and the protein levels *in vivo*. There is a general increase in cellular protein levels

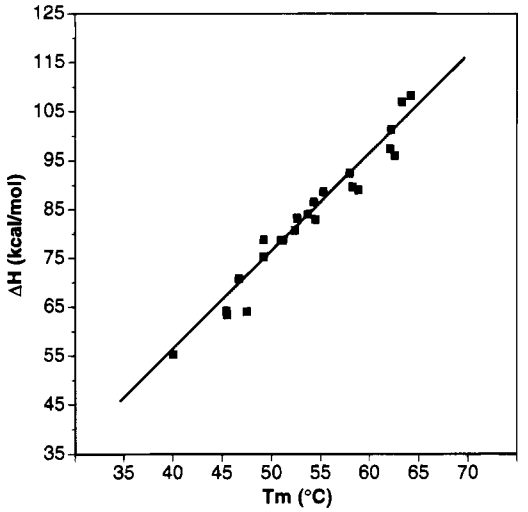

FIGURE 7: Thermal dependence of enthalpy of unfolding. The enthalpy of unfolding (ΔH) for position 52 replacements at different pH values are plotted against the T_m of unfolding. The slope of the line by linear regression equals ΔC_p. The points are the same as in Table 3.

Table 3: Microcalorimetry Thermodynamic Parameters

cytochrome *c* (C102A)	pH	T_m (°C)	ΔH (kcal/mol)	ΔS (cal/mol K)	ΔG (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
N52	5.0	63.3	107.0	318	5.62	
	4.6	49.2	78.8	244	4.03	
	4.0	45.4	64.2	202	2.52	
	3.5	40.1	55.3	177	1.64	
N52Q	4.6	51.2	78.7	243	3.95	−0.08
N52S	5.0	52.6	83.2	256	4.33	−1.29
	4.6	51.0	78.7	243	3.98	−0.05
	4.0	46.7	70.8	221	3.26	0.74
N52H	4.6	47.5	64.1	180	2.76	−1.27
N52T	5.0	55.3	88.6	270	5.09	−0.53
	4.6	54.5	82.9	253	4.62	0.59
	4.0	49.2	75.3	234	3.53	1.01
	3.5	45.4	63.4	199	2.62	0.98
N52A	4.6	53.7	84.0	257	4.66	0.63
N52V	4.6	58.0	92.4	279	5.70	1.67
	4.0	54.3	86.5	264	4.96	2.44
N52M	4.6	62.6	96.0	286	6.28	2.25
	4.0	58.9	89.0	268	5.43	2.91
N52L	4.6	62.1	97.4	290	6.59	2.56
N52I	5.0	64.2	108.2	321	7.64	2.02
	4.6	62.2	101.3	301	7.17	3.14
	4.0	58.3	89.6	270	5.55	3.03
	3.5	52.4	80.7	248	4.26	2.62

with an increase in thermodynamic stability as measured by ΔG_U. The correlation coefficient for this relationship as measured by guanidine hydrochloride denaturation is 0.71 excluding the N52H protein.

Microcalorimetry. It is necessary to have an accurate ΔC_p, change in heat capacity between the folded and unfolded forms of a protein, to measure the change in free energy, ΔG_U. The value of ΔC_p is related to the number of hydrophobic residues (Privalov, 1979). It has been shown that ΔC_p can change drastically for mutated proteins (Shortle, 1989). Because this series of cytochrome *c* amino acid replacements has vastly different hydrophobicities at position 52, it is important in assessing thermal stability to determine if a change in ΔC_p accompanies the replacement. Cytochromes *c* that covered the range of available hydrophobic residues at position 52 were tested (Figure 7, Table 3). Calculated ΔC_p's from a single experiment are inherently inaccurate because of the difficulty in obtaining a proper base line. Since ΔC_p is constant with pH, the change in heat capacity can be more accurately measured using Kirchhoff's relationship, $\Delta_d C_p = (\delta\Delta_d H)/(\delta T)$, with microcalorimetric data obtained at different pH values. Ionization heat effects are excluded by the use of acetate buffer, which has an enthalpy of ionization close to that of a protein's ionizable groups in the pH range studied. Figure 7 shows the strong correlation between $\delta\Delta H$ versus δT for all proteins despite the different levels of hydrophobicity at position 52. The slope of the plot gives a ΔC_p value of 2.0 kcal/K mol. The same value was previously calculated by Hickey *et al.* (1991a) by plotting the $\delta\Delta H$ versus δT_m of the normal C102A iso-1-cytochrome *c* with small amounts of added guanidine hydrochloride in thermodenaturation experiments. Since the variations in ΔC_p are within experimental error, this average value of ΔC_p = 2.0 kcal/K mol was used to calculate the ΔG_U values. For each of the proteins there is the expected nonlinear decrease in ΔH with decreasing pH. A decreasing ΔS compensation was seen with decreasing ΔH. A plot of ΔG_U versus hydrophobicity (Figure 5) of the amino acid at position 52 shows a linear dependence. The correlations are 0.99 for ΔG_U and 0.96 for T_m, excluding N52 and N52H. These amino acids have other possible effects beside hydrophobicity influencing stability as elaborated in the Discussion section. The stability of the proteins measured by microcalorimetry is related to the protein levels *in vivo* as seen by Figure 6. The correlation for this relationship is 0.68, excluding N52H. The histidine replacement has a lower thermodynamic stability than what would be predicted by the cellular protein levels. Other explanations such as a difference in the ionization state of the histidine side chain in buffer compared to *in vivo* or an increased resistance to protease in the N52H protein may be involved.

DISCUSSION

By using the *in vivo* mutagenesis technique described, a single stable copy of the desired cytochrome *c* gene was integrated at the *CYC1* locus in isogenic strains. This allows direct comparisons of amino acid replacements to explore the relationship between stability *in vivo* and *in vitro*. Low-temperature spectrophotometric recordings of intact cells give direct measurements of the holoprotein levels *in vivo*. Earlier pulse-chase experiments (Pearce and Sherman, manuscript in preparation) suggest that the increase in cellular protein levels, associated with increased stability, is the result of a longer half-life for the protein turnover of the intact holoprotein rather than other factors in the translational or posttranslational processing of cytochrome *c*. This work with changes at position 52 and other work with destabilizing changes at position 6 (accompanying manuscript) clearly show a general relationship between thermodynamic measurements of proteins *in vitro* and steady state protein levels *in vivo* (Figure 6).

Previous work (Das *et al.*, 1989; Hickey *et al.*, 1991a) shows that the N52I and N52A replacements of yeast iso-1-cytochrome *c* are more stable during unfolding *in vitro* than the wild type protein. The stability of the proteins with replacements at position 52 is correlated with the hydrophobicity of the side chain. Both chemical and heat denaturation show this direct relationship for the amino acid replacements tested (Figure 5). Furthermore, several amino acid replacements in this series have additional factors besides hydro-

phobicity that must be considered. There is an obvious optimal amino acid size defined by the dense packing in the protein interior (Richards, 1977). For example, work with gene V proteins has shown that for some proteins, the destabilization seen with amino acid substitutions is related to packing effects (Sandberg & Terwilliger, 1989). The cavity created by removing N52 and the internal water molecule can be estimated by the size of cavity left by the water molecule and the volume of an average asparagine residue as measured by Richards (1977). Water 166 has a cavity of ~10 Å^3 in the reduced protein and ~25 Å^3 in the oxidized protein (Berghuis & Brayer, 1992). An average asparagine has a volume of 135.2 Å^3. This volume (~145–160 Å^3) plus some degree of flexibility would define the largest possible amino acid substitution. Leucine is the largest amino acid replacement (average volume of 168.8 Å^3) that shows no other destabilizing effects. The somewhat larger N52F (203.4 Å^3) shows mixed effects. The variant N52F is more resistant to chemical denaturant, as measured by C_m, and has higher levels *in vivo* than the normal protein, but the $\Delta\Delta G_U$ is incongruently smaller than what would be predicted from hydrophobicity effects alone. Furthermore, the slope of unfolding, m_G, is less steep than that of the normal N52 protein. It is possible the size of the side chain is destabilizing. Larger side chains, W (average 237.6 Å^3) and Y (average 203.6 Å^3), are clearly destabilizing. Another possibility is that the replacement is stabilizing the unfolded form of the protein (Shortle, 1989). It is interesting to note that the addition of a single hydroxyl group from phenylalanine to tyrosine has such a large effect on the cytochrome *c* levels *in vivo* and, presumably, the protein stability.

The destabilizing effect *in vivo* of D, E, K, and presumably R is expected because the charged residue must be buried at a high thermodynamic cost. A partial positive charge on the histidine residues could also result in destabilization. The N52H replacement shows a large decrease in the $\Delta\Delta G_U$ in chemical and microcalorimetric measurements. This could signal a protonation step coupled to the unfolding. The N52A substitution has a large slope of unfolding in the guanidine hydrochloride experiments. The increased cooperativity could indicate a destabilization of a partially folded form. Glycine was shown to be less stable (Hickey *et al.*, 1991a), possibly due to an increase in flexibility and/or the creation of a cavity inside the protein. The crystal structure of the wild type, N52, shows an internal water molecule. The stability is higher than what would be predicted by hydrophobicity alone. This suggests a possible effect of an internal solvent molecule on the stability of the protein.

There is current debate about how to quantitate the effects of amino acid replacements in the interior of a protein as is seen at position 52 in iso-1-cytochrome *c*. The importance of side chain hydrophobicity in protein stability has been recognized for a long time (Tanford, 1962). Recent work has shown that internal packing of nonpolar residues is an important driving force in the *de novo* design of a globular protein (Kamtekar *et al.*, 1993). The strength of the free energy change due to hydrophobicity is conventionally estimated by partitioning amino acid derivatives between nonaqueous solutions and water. As seen in this work, the $\Delta\Delta G_U$ of unfolding values are larger than can be accounted for by the hydrophobicity as measured by the Δg_{tr} between octanol and water (Fauchere & Pliska, 1983). This method underestimates the hydrophobic effect in a folded protein, and/or other factors are involved. A problem often cited is that a nonaqueous solution does not mimic the crystalline nature of a protein's interior. A change in van der Waal's interactions from the alteration of an internal cavity is another possible source of thermodynamic free energy not accounted for by the partitioning experiments. Other factors such as hydrogen bonds and internal solvent can also be important in an altered protein. Eriksson *et al.* (1992) have shown that the $\Delta\Delta G_U$ seen in a change to a smaller amino acid can be accounted for by a constant change in free energy for each similar amino acid substitution with the addition of a factor related to the size of the cavity created in the replacement. Sharp *et al.* (1991) have applied solution thermodynamics to this problem. The Δg_{tr} values obtained by Fauchere and Pliska are recalculated with the entropy of mixing taken into account. These values are also corrected for the change in volume between solute and solvent. Pace (1992) has suggested that with these corrections and a correction for the percentage of the side chain that is buried, the change in free energy can be accounted for by the hydrophobic effect quantitatively based on the number of CH, -CH_2, and CH_3 groups with no consideration for cavity formation. Lee (1993) has estimated that the ΔG_U values from Sharp *et al.* (1991), uncorrected for volume change, give a good minimum value for the range of free energies for each type of amino acid change in a flexible area of the protein. In this treatment, as the site of mutation becomes more rigid, the size of the cavity delineates the upper range of free energy change. Therefore, the free energy of cavity formation becomes important in a rigid structure.

Without crystallographic information, it is impossible to quantitate the cavity formation and/or the presence of water molecules in the interior of these mutated proteins to analyze their importance in this work. The general treatments of Lee (1993) and Pace (1992) can be applied. The calculated $\Delta\Delta G_U$ values for the N52 amino acid replacement agree with the minimum values presented by Lee (1993). (The $\Delta\Delta G_U$ value for the V → A replacement was lower than predicted by the model for the guanidine hydrochloride experiments.) This agrees with what is predicted for a flexible area of the protein such as position 52. When the data were treated as in the Pace paper (1992), the values of $\Delta\Delta G_U/CH_x$ groups were somewhat lower than those calculated.

Defining and quantitating all the stability effects in a set of single-amino acid replacements is very complicated. At this single site, several factors are involved in the overall protein stability. Side chain size and hydrophobicity, cavity formation, and an internal solvent molecule have to be considered. Several of the replacements also show changed folding profiles by chemical denaturant that could predict a stabilization in the unfolded or partially folded form. A deeper understanding of these effects on position 52 replacements will be further explored through X-ray crystallography and folding studies.

REFERENCES

Berghuis, A. M., & Brayer, G. D. (1992) *J. Mol. Biol. 223*, 959–976.

Betz, S. F., & Pielak, G. J. (1992) *Biochemistry 31*, 12337–12344.

Creighton, T. (1984) *Proteins*, Freeman, New York.

Das, G., Hickey, D. R., McLendon, D., McLendon, G., & Sherman, F. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 496−499.
Drummond, D. R., Hennessey, E. S., & Sparrow, J. C. (1991) *Biochem. J. 274*, 301−303.
Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992) *Science 255*, 178−183.
Fauchere, J.-L., & Pliska, V. (1983) *Eur. J. Med. Chem. 18*, 369−375.
Fetrow, J. S., Cardillo, T. S., & Sherman, F. S. (1989) *Proteins 6*, 372−381.
Hecht, M., Sturtevant, J. M., & Sauer, R. T. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5685−5689.
hickey, D. R., Berghuis, A. M., Lafond, G., Jaeger, J. A., Cardillo, T. S., McLendon, D., Das, G., Sherman, F., Brayer, G. B., & McLendon, G. (1991a) *J. Biol. Chem. 266*, 11686−11694.
Hickey, D. R., Jayaraman, K., Goodhue, C. T., Shah, J., Clements, J. M., Tsunasawa, S., & Sherman, F. (1991b) *Gene 105*, 73−81.
Jaenick, R. (1993) *Philos. Trans. R. Soc. London B 339*, 287−295.
Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M., & Hecht, M. (1993) *Science 262*, 1680−1685.
Kim, A., & Szoka, F. C. (1992) *Pharm. Res. 9*, 504−514.
Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
Lee, B. (1993) *Protein Sci.* 2, 733−738.
Lee, B., & Richards, R. M. (1971) *J. Mol. Biol. 55*, 379−400.
Leotiev, V. V., Uversky, V. N., & Gudkov, A. T. (1993) *Protein Eng. 6*, 81−84.
Louie, G. V., & Brayer, G. D. (1990) *J. Mol. Biol. 214*, 527−555.
Moerschell, R. P., Tsunasawa, S., & Sherman, F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 524−528.
Moerschell, R. P., Hosokawa, Y., Tsunasawa, S., & Sherman, F. (1990) *J. Biol. Chem. 265*, 19638−19643.
Nozaki, Y., & Tanford, C. (1971) *J. Biol. Chem. 246*, 2211−2217.
Pace, C. N. (1975) *Crit. Rev. Biochem. 263*, 1−43.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
Pace, C. N. (1992) *J. Mol. Biol. 226*, 29−35.
Pakula, A. A., Young, V. B., & Sauer, R. T. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8829−8833.
Parsell, D. A., & Sauer, R. T. (1989) *J. Biol. Chem. 264*, 7590−7595.
Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167−241.
Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151−176.
Sandberg, W. S., & Terwilliger, T. C. (1989) *Science 245*, 54−57.
Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.
Sharp, K. A., Nicholls, A., Friedman, R., & Honig, B. (1991) *Biochemistry 30*, 9686−9697.
Sherman, F., & Slonimski, P. P. (1964) *Biochim. Biophys. Acta 90*, 1−15.
Sherman, F., Stewart, J. W., Jackson, M., Gilmore, R. A., & Parker, J. H. (1974) *Genetics 77*, 255−284.
Sherman, F., Fink, G. R., & Hicks, J. B. (1987) *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Shortle, D. (1989) *J. Biol. Chem. 264*, 5315−5318.
Tanford, C. (1962) *J. Am. Chem. Soc. 84*, 4240−4247.
Yamamoto, T., Moerschell, R. P., Wakem, L. P., Ferguson, D., & Sherman, F. (1992) *Yeast 8*, 935−948.
Zamyatnin, A. A. (1972) *Prog. Biophys. Mol. Biol. 24*, 107−123.

BI9415659