Structure, Function, and Temperature Sensitivity of Directed, Random Mutants at Proline 76 and Glycine 77 in Ω-Loop D of Yeast Iso-1-cytochrome *c*[†]

Jacquelyn S. Fetrow,* Jennifer S. Spitzer, Brad M. Gilden, Scott J. Mellender, Thomas J. Begley, Brian J. Haas, and Terry L. Boose

Department of Biological Sciences, Center for Biochemistry and Biophysical Chemistry, State University of New York at Albany, Albany, New York 12222

Received September 12, 1997; Revised Manuscript Received December 9, 1997

ABSTRACT: Residues 75−78 form a tight turn within Ω-loop D in *Saccharomyces cerevisiae* iso-1-cytochrome *c*. Directed, random mutagenesis of invariant residues proline 76 and glycine 77 in this turn were analyzed for the in vivo functionality and level of protein within the cell. All proteins, except Pro76Val, also exhibit a significant decrease in intracellular cytochrome *c* levels, ranging from 15% to 80% of wild type. Furthermore, all isolated mutant strains, except the one expressing Pro76Val, exhibit a significant decrease in growth on lactate medium, suggesting that the variant cytochromes are much less functional than wild type. This requirement for protein function is clearly the cause for the strict invariance of these residues in eukaryotic cytochromes *c*. Seven proteins with mutations just at Pro76 were purified and studied by circular dichroism spectroscopy. All proteins with mutations at Pro76 exhibit melting temperatures about 7 °C less than that of the wild-type protein, suggesting that mutation of Pro76 affects the entropy of the denatured state. It is proposed that the functional significance of Pro76 and Gly77 is the requirement for a type II ($\beta_{\gamma L}$) β-turn in this loop, the conformation of which requires a glycine at the third position, and that a change occurs in this turn conformation upon a change in the redox state of the protein.

De novo protein design and engineering requires an understanding of the physical properties that govern protein structure and stability. Numerous studies of peptides and proteins have led to some understanding of the rules that govern helix structure and stability, and more recently, a number of groups have begun to study β-sheet structure and stability. Few studies have focused on three- and four-residue turns. Studies of structure and stability of longer loops are even more rare. Approximately one-quarter of the structure of all proteins is composed of these longer loops; thus, an understanding of these structures is critical to the further progress of protein design and engineering. Our research has focused on studying the nonregular secondary structures, the loops and turns, and the role that they play in protein structure, stability, folding, and function.

Many researchers regard loops and turns as simple connections between the well-recognized regular secondary structures and consider these structures as polypeptide connectors that have little structural significance. On the other hand, it has long been suggested that some loops and turns may be important in folding and stability (*1*−*3*). Recent data have shown that changes in these structures can markedly affect protein stability. Some analyses have started to elucidate intrinsic and extrinsic effects (*4*−*6*). These seemingly opposing descriptions of loops, simple connections versus structurally significant connections, may both be accurate in describing different loops and turns, but our understanding of how these structures affect the structure and stability of the main body of the protein is lacking. To help rectify this lack, we have been studying the loops in iso-1-cytochrome *c*. In this paper, we describe the complete analysis of a β-turn located in an Ω-loop in this protein.

Turns are small structures, generally three, four, or five residues in length. As turns were originally defined, a hydrogen bond is found between the carbonyl oxygen of residue i and the amide nitrogen of residue $i + 3$ (*7*). This original classification scheme has been expanded and subdivided to include various types of turns based on the backbone dihedral angles of residues $i + 1$ and $i + 2$ (*8*−*11*). Comprehensive reanalysis of these structures and their amino acid preferences has been published (*12*, *13*). Subcategories of turns based on their flanking regular secondary structures have also been described. β-Hairpins are turns found between β-strands (*14*−*16*). α−α hairpins connect two α-helices (*17*, *18*).

Although the structures of turns have been well described, their role in determining protein structure and stability is not well understood. Turns between regular secondary structure elements have been studied experimentally. Some turns appear to be quite variable without much effect on protein structure, while substitutions in other turns affect protein stability. For example, random mutagenesis of a turn between two α-helices of the four-helix bundle protein cytochrome *b*562 suggests that this turn plays little, if any, role in determining the final protein structure (*19*). Likewise,

[†] This work was supported by NIH Grant GM44829 to J.S.F.

* To whom correspondence should be addressed at Department of Biological Sciences, University at Albany, SUNY, 1400 Washington Ave., Albany, NY 12222; phone 518-442-4389; email jacque@isadora.albany.edu.

S0006-2960(97)02279-4 CCC: $15.00
Published on Web 02/04/1998

Table 1: Loop D Sequence of Iso-1-cytochrome *c* and the Sequences of Residues 75−78 in 106 Sequenced Eukaryotic Cytochromes *c*[a]

	70				75		78		80				84
Iso-1 cyt *c*	Asn	Pro	Tml[2]	Lys	Tyr	Ile	Pro	Gly	Thr	Lys	Met	Ile	Phe	Val	Gly
Crithidia oncopelti						Met									
Crithidia fasciculata						Met									
Chlamydomonas reinhardtii						Met			Asn						
Euglena gracilis						Val									
Euglena viridis						Val									
Tetrahymena pyriformis						Val									

[a] Ninety-six sequences were originally compiled by Pettigrew and Moore (*77*); 10 additional sequences have since been added (*78*). Superscript 2 indicates this lysine side chain is trimethylated in *Saccharomyces cerevisiae* iso-1-cytochrome *c*.

the sequence of a turn in the four-helix bundle protein Rop can vary greatly with little effect on protein structure (*20*), but these substitutions have a large effect on protein stability (*4*). Computational analysis suggests that turn−helix interactions are important for stabilizing the structure of these four-helix bundle proteins (*21*). When a four-residue type I′ β-turn between two β-strands from staphylococcal nuclease was replaced by a five-residue turn from concanavalin A, the hybrid protein was less stable than the wild-type nuclease and the turn retained the conformation found in the concanavalin protein, a type I turn followed by a bulge (*22*). Other mutations in turns have a larger effect on turn formation (*23*, *24*) and this suggests that these turns may play a role in protein folding (*3*). For example, a five-residue peptide corresponding to a turn connecting helices G and H in myoglobin was found to adopt a reverse turn conformation in solution, a structure that is lost when the proline and glycine in this turn are replaced by two alanines (*25*).

As opposed to turns, the term "loop" is generally applied to connecting segments longer than turns. These are inherently more difficult to classify because of the range of conformations that their backbone atoms can assume. Classes of loop conformations have been published for specific families of proteins, including the immunoglobulins (*26*−*28*) and the α/β barrel proteins (*29*−*31*). A general class of loop, the Ω-loop, was first described in 1986 (*1*). This structure, which is six or more residues in length, is classified on the basis of the proximity of the end points of the peptide segment rather than on the specific conformation of the peptide backbone. This classification of Ω-loops, peptide segments whose end points are close in space, can be subdivided into a class also called Ω-loops, in which the backbone is relatively planar, and ζ-loops, in which the peptide backbone is not planar (*32*). A third class, the strap loop, in which the segment end points are not close in space, was also described (*32*). (For a review, see ref *33*.) In this research, the term Ω-loop is used as originally defined (*1*), and this definition includes both Ω- and ζ-loops described by Ring et al. (*32*). Another general classification of short, three- to eight-residue turns/short loops subdivided by size and by flanking secondary structures and then clustered on the basis of backbone rms differences was also recently published (*34*) and others have classified longer loops on the basis of flanking secondary structures (*35*, *36*).

As described above, much of the mutagenesis work on turns in proteins has been done on turns that link segments of regular secondary structures, particularly turns between helices; however, many β-turns are located within longer protein loops and these turns have not been studied. Furthermore, little work has been done to study those particular forces that stabilize these larger loop structures or that control their dynamics and flexibility. Because we are particularly interested in studying the role of the longer loops in protein folding, stability, function, and dynamics, we wish to understand the role of the reverse turns in stabilizing these structures. We further want to analyze how the characteristics of turns in loop regions compare to those of turns found between regular secondary structures.

To this end, two residues in a reverse turn located in an Ω-loop in the protein iso-1-cytochrome *c* from the yeast *Saccharomyces cerevisiae* have been randomly mutated (*37*). Because cytochrome *c* from yeast has been the subject of intense research for many years, it is an ideal protein for studying the role of loops and turns in protein folding, stability, and structure, both in vivo and in vitro. First, the yeast organism itself is easy to manipulate genetically (*38*, *39*) and large amounts of mutation data, including data from directed, random mutagenesis experiments, are available for this protein (*37*, *40*−*42*). The crystal structures of cytochromes *c* from six eukaryotic organisms have been solved, including iso-1- [reduced, 1ycc (*43*); oxidized, 2ycc (*44*)], iso-2- [reduced, 1yea (*45*)], horse [oxidized, 1hrc (*46*)], albacore [reduced, 5cyt (*47*); oxidized, 3cyt (*48*)], rice [1ccr (*49*)], and bonito cytochromes *c* [1cyc (*50*)], and nuclear magnetic resonance (NMR) solution structures of horse (*51*−*53*) and yeast iso-1-cytochromes *c* (*54*, *55*) have been reported. Crystal structures of a number of mutant cytochromes *c* have also been determined (*56*−*62*). Interesting NMR experiments for structural analysis of iso-1-cytochrome *c* have been performed (*63*−*71*) and much progress toward understanding the folding pathway of this protein has been accomplished (*72*−*76*). All of this information can be used to dissect the importance of our chosen reverse turn.

In Ω-loop D, residues 70−84 in iso-1-cytochrome *c* (Table 1), a hydrogen bond is formed between the carbonyl oxygen of Ile75 and the amide nitrogen of Thr78 (Figure 1) (*43*). In the crystal structure of ferrocytochrome *c* (*43*), the hydrogen-bond nitrogen−oxygen distance is 2.94 Å and the backbone dihedral angles of Pro76 and Gly77 form a type II [$\beta_{\gamma L}$ in the nomenclature of Wilmot and Thornton (*12*)] β-turn (Table 2). In the structure of the oxidized protein (*44*), the hydrogen bond is slightly longer (2.98 Å), though still within hydrogen-bonding distance, but the backbone dihedral angles of Pro76 and Gly77 no longer assume a type II β-turn conformation (Table 2), and in fact do not fit neatly into any of the turn categories described by Thornton and colleagues (*12*, *13*).

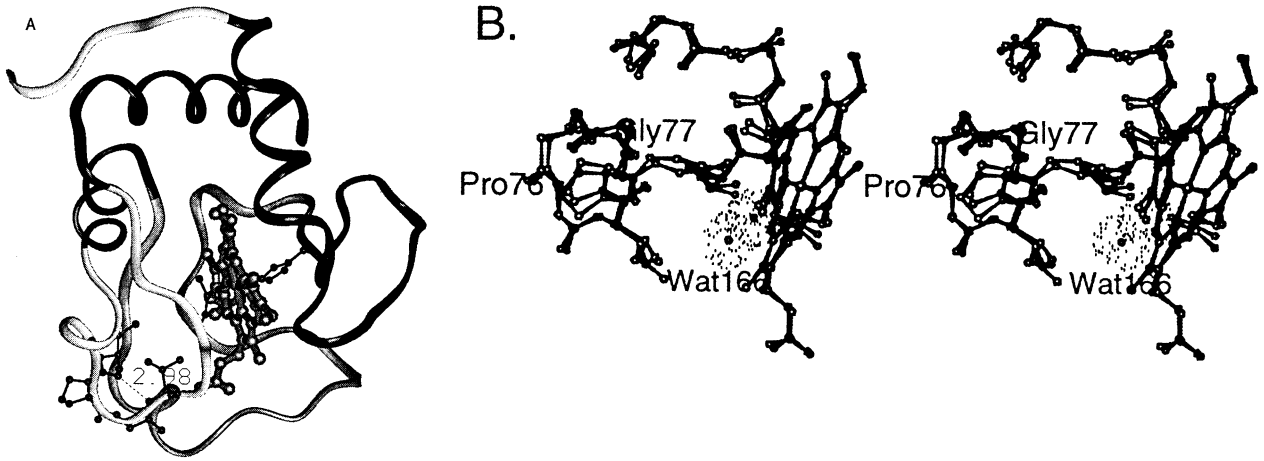

FIGURE 1: (A) Ribbon structure of iso-1-ferrocytochrome *c* (*43*). Loop D is the lightest gray structure and is shown with the side chains of Pro76 and Gly77 as black ball-and-stick models. The hydrogen bond between Ile75 O and Thr78 N shown as a dotted line. (B) Stereoview of the backbone of loop D, residues 70−84, taken from the crystal structures of oxidized (*44*) and reduced (*43*) iso-1-cytochrome *c*. The two structures were superimposed by doing a least-squares fit of the loop backbone atoms. The side chains of Pro76 and Gly77 are drawn. Wat166, the conserved water molecule, is shown with a van der Waals surface drawn.

Table 2: Backbone Conformations of the Turn, Residues 75−78, Found in the Cytochrome *c* Crystal Structures from Five Eukaryotes

		proline 76 (deg)			glycine 77 (deg)			
species	redox state	φ	ψ	ω	φ	ψ	ω	H-bond distance (Å)
S. cerevisiae (iso-1)	ox	−62.6	115.2	−177.5	138.1	−34.1	176.3	2.98
	red	−66.3	128.8	−178.0	99.7	−1.5	177.1	2.94
horse (NMR)	ox	−60.3	114.7	−176.7	97.2	5.7	178	2.75
	red	−79.7	14.8	−178.3	−95.1	−33.3	−179.4	3.80
tuna	ox	−56.5	128.5	−178.6	94.7	−0.68	178.3	3.02
	red	−66.8	132.2	−174.9	110.1	−17.0	−167.4	3.13
rice (P84,G85)	ox	−58.2	134.4	−179.9	93.5	−1.51	176.1	3.05
S. cerevisiae (iso-2)	red	−59.1	130.1	179.5	104.5	−24.7	179.9	2.99

However, the polypeptide chain still reverses direction in space. Although the hydrogen bond is not observed in the solution NMR structure of reduced iso-1-cytochrome *c* reported recently (*54*), the Thr78 amide nitrogen proton exhibits measurable protection factors in ferro-iso-1-cytochrome *c* but has exchanged out by the first time point (6.55 h) in ferri-iso-1-cytochrome *c* (*67*). These data suggest that a conformational change occurs in the turn between oxidized and reduced forms of the protein.

The loop D sequence is the most highly conserved segment in eukaryotic cytochrome *c* evolution (*37*). Residues $i + 1$ and $i + 2$ in the turn, Pro76 and Gly77, are invariant in the sequence alignment of the 106 eukaryotic cytochromes *c* (Table 1) (*77*, *78*). Similarly, residues *i* and $i + 3$, Ile75 and Thr78, are very highly conserved. The residue homologous to Ile75 is a methionine in *Crithidia oncopelti*, *Crithidia fasciculata*, and *Chlamydomonas reinhardtii* and by valine in *Euglena gracilis*, *Euglena viridis*, and *Tetrahymena pyriformis*, while residue 78 is not a threonine in just one organism, *C. reinhardtii* (Table 1).

To better understand how this conserved turn structure located within an Ω-loop affects the structure of the loop and the overall structure and function of the protein, we performed site-directed, random mutagenesis of Pro76 and Gly77 (*37*). Forty-nine independent mutant transformants were obtained and sequenced, producing 45 different sequence combinations at residues 76 and 77 (*37*). We have now completed analysis of the amount and level of function of cytochrome *c* in these yeast strains at 30 °C and determination of the temperature sensitivity of the yeast strains expressing these variants. In addition, several proteins with mutations only at Pro76 were purified and their thermal stability was observed by circular dichroism (CD) spectroscopy. These data are discussed with respect to the role of this turn in the structure and function of cytochrome *c*.

MATERIALS AND METHODS

Yeast Strain and Protein Nomenclature. C93 is the standard, "wild-type" yeast strain used in these studies and all other yeast strains are isogenic with this strain, except at the CYC1 locus. The C93 genotype is *MAT*a, *CYC1*-C102T *cyc7*−68::*CYH2 his3*-Δ1 *ura3*-52 *leu2*-3 112 *trp1*-289 *can1*-100 *cyh2* LEU2/cytc (*37*). Strain C15 is the "negative control" strain that was produced from C93 and contains a complete deletion of the CYC1 locus. Mutant strains were produced by directed, random mutagenesis in which a strain containing a nonfunctional cytochrome *c* is transformed with a randomized oligonucleotide and transformants are screened for a functional cytochrome *c*, as previously described (*37*, *79*). Strains produced in this fashion with random mutations at Pro76 and Gly77 are designated as 3R or 4R followed by the strain number, as 4R-1. All strains are isogenic with C93 and lack the iso-2-cytochrome *c* isozyme. Diploid strains, made by crossing these strains with strain B2111 (*80*), are denoted by a small d following the strain number, as C93d.

Variant proteins are designated by the one-letter amino acid abbreviation of the wild-type residue, followed by the residue number, followed by the name of the variant residue

substituted at this site. Thus, P76Y indicates a protein in which proline 76 was replaced by a tyrosine. Multiple changes in the same protein are indicated in the same fashion, with residues separated by a comma, as P76A,G77E.

The standard eukaryotic sequence alignment for numbering the yeast cytochrome *c* sequences is used; thus, yeast cytochrome *c* extends from Thr(−5) to Glu103. All of the cytochromes *c* used in this study, including the "wild-type" protein found in strain C93, contain threonine rather than a cysteine at position 102 and are thus C102T. This amino acid substitution makes in vitro analysis easier by preventing the slow dimerization through disulfide bond formation of the cysteine residues; the C102T protein has been shown to exhibit characteristics similar to those of the wild-type protein (*81*). Because all of the proteins used in this study contain this mutation, it is assumed, but not explicitly specified, in the protein nomenclature.

Growth Curves in Lactate. To test the level of function of the mutant cytochromes *c* within the cell, yeast strains were grown at 30 °C in liquid lactate medium. A functional cytochrome *c* is required for growth on this nonfermentable carbon source and the amount of growth is a measure of the level of functionality of the cytochrome *c* in isogenic strains. Haploid yeast were first made diploid by crossing with strain B2111 (*80*) to counteract interference from any background mutations in other genes (*82*). Experiments were performed as previously described (*37*). Cell growth was observed by monitoring culture turbidity over a period of 30 h. The percentage of growth of each mutant strain was determined by comparison of its culture density to the density of strains C93 (wild-type) and C15 (cyc1 deletion) at 30 h [(density$_{mutant}$ − density$_{C15}$)/(density$_{C93}$ − density$_{C15}$)]. Each reported result is the average of at least three experiments. Repeated experiments have shown the error in these measurements to be approximately 3−10%.

Whole Cell, Low-Temperature Difference Spectroscopy. To determine the amount of cytochrome *c* present in whole yeast cells, low-temperature difference spectroscopy was performed (*83*, *84*). We have not been successful in obtaining consistent results using a paste of cells and normalizing the quantity of cytochrome *c* by measuring the ratio of cytochrome *c* to cytochrome *b* (*85*); thus we quantify the number of cells in the cuvette prior to spectroscopic measurements and perform the experiment as previously described (*37*). Prior to the experiment, strains are grown at 30 °C on YPD medium (1% yeast extract, 2% peptone, and 3% dextrose), a fermentable carbon source for which a functional cytochrome *c* is not required, so that cytochrome *c* proteins that are nonfunctional can be visualized and quantitated. Briefly, a known quantity of cells is divided into two aliquots. One sample is oxidized with $K_3Fe(CN)_6$; the other is reduced with dithionite. Saturated sucrose is added, the samples are mixed, and the cuvettes are frozen in liquid nitrogen. A visible difference spectrum from 500 to 700 nm is then acquired. The amount of protein in the cell was determined by comparison of the absorbance at 548 nm to that measured for an equal quantity of cells from yeast strains C93 and C15 [(absorbance$_{mutant}$ − absorbance$_{C15}$)/(absorbance$_{C93}$ − absorbance$_{C15}$)]. Reported results are the average of at least three experiments. Results from repeated experiments show that the error in these measurements is about 5−12%.

Temperature Sensitivity Analysis. To determine the temperature sensitivity of strains containing mutations at Pro76 and Gly77, growth on solid medium was monitored at three different temperatures. Yeast strains were grown in 1 mL of liquid YPD medium overnight. The culture (70 μL) was centrifuged and the cell pellet was washed twice with sterile distilled water. The cells were resuspended in 1 mL of sterile water. Aliquots (5 μL) of this cell suspension were then plated on YPD, YPG (1% yeast extract, 2% peptone, and 3% glycerol), and YPL (1% yeast extract, 2% peptone, and 1% lactate) plates. One of each type of media plate was then grown at ambient temperature (approximately 25 °C), 30 °C, and 37 °C and the relative growth was recorded for 3 days following plating.

On each day, growth was scored on a scale of 0−3, with the wild-type strain (C93) being scored as 3 and the cytochrome *c* deletion strain (C15) being scored as 0. Strains were then classified as nts (not temperature-sensitive), sts (slightly temperature-sensitive), ts (temperature-sensitive), bg (barely grew), sg (slow grower), or nwt (not wild type). Further descriptions of these classifications are presented in Results.

Circular Dichroism Spectroscopy and Thermal Stability Measurements. Proteins from strains C93, 3R4, 3R7, 3R17, 3R22, 3R25, 4R8, and 4R20 were purified by a modification of the standard protocol (*38*), as previously described (*5*). During the purification, proteins were oxidized with $K_3Fe(CN)_6$, which was removed prior to the second ion-exchange column. Lyophilized proteins were subjected to high-performance liquid chromatography (HPLC)−electron spray ionization (ESI) mass spectroscopy to confirm their molecular weight. SDS−PAGE (*86*) of each protein showed a single band indicating that the proteins were pure. Circular dichroism spectroscopy was accomplished using an Aviv 62DS spectropolarimeter essentially as previously described (*5*). Briefly, lyophilized protein was dissolved in 100 mM acetate buffer, pH 4.6, to a concentration of about 9 μM. Exact concentrations were calculated from the visible absorbance at 410 nm using the extinction coefficient of 106.1 cm^{-1} mM^{-1} (*87*). Degassed, oxidized samples were assayed in 1-mm quartz cuvettes. Spectra were acquired from 200 to 250 nm, every nanometer with a 1-nm bandwidth. The averaging time per data point was 15−30 s, depending on the experiment. To visualize the spectrum of each protein, mean residue ellipticity was calculated (*88*) and plotted versus wavelength.

Thermal stability was assessed by monitoring the CD spectrum at 222 nm from 5 to 80 °C, with a 30−90 s equilibration time, and a 30 s averaging time per data point, essentially as previously described (*5*). Mean residue ellipticity was calculated at each temperature (*88*), and data were processed by a nonlinear least-squares fitting procedure (KaleidaGraph) to the equations of ref *89*, essentially as previously described (*5*, *90*). The equation fits the mean residue ellipticity and temperature data to native and denatured baseline points, ΔH_m, and T_m. $K_{D,T}$ was calculated from these data using the modified Gibbs−Helmholtz equation (*89*) and $\Delta G_{D,298}$ was then calculated using the Gibbs equation, as previously described (*5*, *90*). During the calculations, ΔC_p was held constant at 1.37 kcal mol^{-1} K^{-1}, the value determined for wild-type C102T (*90*). Errors in these calculations were estimated using the error in ΔC_p

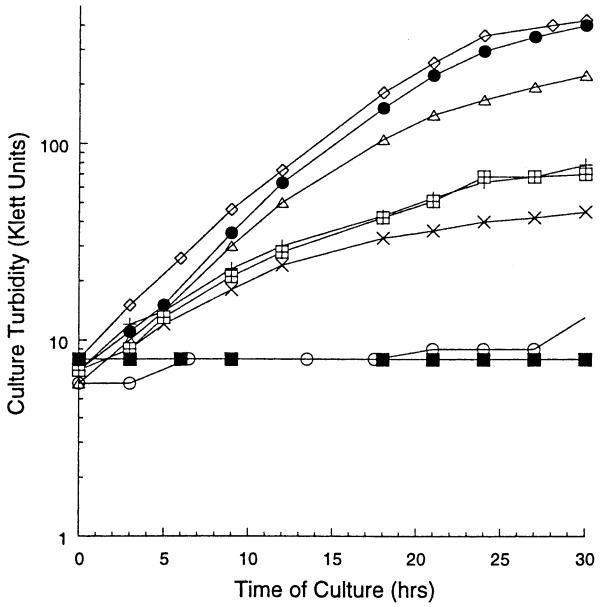

FIGURE 2: Representative growth curves on liquid lactate medium of diploid yeast strains expressing the Pro76,Gly77 mutations. Control strains: C93, ●; C15, ■. Experimental strains: 3R4, ○; 3R7, ◇; 3R22, ×; 3R25, +; 4R8, △; 4R20, plus sign enclosed in a square. A summary of all data can be found in Table 3.

published by Cohen and Pielak (*90*) and a standard analysis of error. Reported results are the averages of two or three experiments.

To apply this type of analysis, the unfolding reaction must be two-state and reversible. To show reversibility, a complete scan from 205 to 240 nm was collected at 5 °C before melting and another complete scan was acquired at 5 °C after quickly cooling the sample following denaturation to 80 °C. Scans acquired before and after thermal denaturation were superimposable for all samples, demonstrating reversibility of the unfolding reaction.

RESULTS

Relative Function of Pro76,Gly77 Mutants. To determine the level of functionality of the Pro76,Gly77 mutants, diploid strains were placed in liquid lactate medium and growth was monitored for 30 h at 30 °C. Yeast strains growing on this nonfermentable carbon source require a functional cytochrome *c*; thus, the level of growth in this medium is a measure of cytochrome *c* functionality in isogenic strains. Results are shown in Figure 2 and summarized in Table 3.

The most striking result of this experiment is the significant decrease in function of all yeast strains, except two. Strain 3R8, which produces the wild-type protein, grows as expected for a strain containing the wild-type sequence. Strain 3R7, which expresses the protein P76V, grows just as well as the strains producing the wild-type protein. All other yeast strains grow to less than 60% of the level of the wild-type strain, as measured by culture turbidity at 30 h in liquid lactate cultures grown at 30 °C.

At 30 h, all but 10 strains have grown to less than 15% of the level of wild-type C93. For these 10 strains, the common denominator is that they maintain a glycine at position 77; thus, each of these 10 strains integrated an oligonucleotide that modified only Pro76. Thus, mutations at Pro76 and Gly77, but particularly Gly77, substantially decrease strain growth on lactate medium. This result is an indication that cytochrome *c* function has been severely compromised in these strains or that these strains are producing much less holocytochrome *c* protein than the wild-type strain.

Amount of Holocytochrome c Present in Yeast Strains Expressing Pro76,Gly77 Mutants. Next we wanted to ask whether the diminished growth was due to a decrease in the amount of protein in the cell. To determine the amount of holocytochrome *c* present in vivo, low-temperature difference spectroscopy was performed on intact yeast cells (*83*, *84*). The peak at 550 nm is due to a $\pi \rightarrow \pi^*$ electronic transition in the heme group of cytochrome *c* (*91*); thus, any cytochrome *c* observed by this technique has entered the mitochondria, bound heme, and folded into a cytochrome *c*-like conformation. The yeast cells were first grown on medium containing dextrose so that even if the variant cytochrome *c* is nonfunctional, the cells would still grow and the amount of holoprotein present within the cell could be measured. The results of these experiments are shown in Figure 3 and summarized in Table 3.

Strains containing random mutations at Pro76 and Gly77 produce various amounts of protein, ranging from 8% to 115%, depending on the exact mutation. The results show that, in parallel with the growth curve assay, strains containing a mutation only at Pro76 while maintaining the wild-type glycine at position 77 generally produce higher levels of protein than strains containing mutations at both positions. Most of these strains produce greater than 70% of the wild-type level of protein. Only one other strain produces more than 70% of the amount of protein found in the wild-type strain: 3R6, which expresses P76A,G77E (Table 3). The majority of the mutant strains produce between 20% and 60% of the level of protein found in C93 (Table 3). Previous work has shown that the growth of strains producing 50% of the wild-type protein is almost indistinguishable from strains producing normal levels of wild-type protein (*92*, *93*). Most of our strains express 40−60% of the normal protein level but grow at less than 20% of the wild-type level on nonfermentable carbon sources. Clearly, the significant decrease in the function of these proteins is, in most cases, not due to a decrease in the amount of holocytochrome *c* present in the cell.

Temperature Sensitivity of Pro76,Gly77 Mutants. The temperature sensitivity experiments were done to determine if the mutant yeast strains were sensitive to temperature, a result which would suggest that the variant proteins are less stable to heat than the wild-type protein and thus possibly degrade faster in the cell. Furthermore, we wished to determine qualitatively if the strains behave differently on the nonfermentable carbon sources, glycerol and lactate, which are metabolized differently in yeast. Temperature sensitivity of yeast strains containing mutations at Pro76 and Gly77 was measured by monitoring growth at ambient temperature (about 25 °C), 30 °C, and 37 °C on dextrose, glycerol, and lactate media plates over the course of 3 days. Strains were classified as follows: nts, not temperature-sensitive, strain was essentially identical with wild type at all temperatures on all days; sts, slightly temperature-sensitive, strain was like wild type at 25 and 30 °C, but grew less well than wild-type at 37 °C; ts, temperature-sensitive, strain grew like wild-type at 25 and 30 °C but grew very little, if at all, at 37 °C; bg, barely grew, strain barely grew

Table 3: Summary of Strain Names, Protein Names, Sequences, and in Vivo Results of Strains Containing Directed, Random Mutations in Loop D

yeast strain	protein name[a]	relative growth[b] (%)	relative amt cyt *c*[c] (%)	temp sen (YPL)[d]	temp sen (YPG)[d]
C93	wild type	100	100	WT	WT
C15	del cyt *c*	0	0		
3R8		99	104	nts	nts
3R7	P76V	100	115	nts	nts
3R5	P76R	40	70	nts	nts
3R17	P76R	33	73	nts	nts
3R11	P76G,N70D,K86E	18	73	nts	nts
3R25	P76G	18	68	nts	nts
4R11	P76G	18	78	nts	nts
4R15	P76W	16	73	nts	nts
3R4	P76W	14	78	nts	nts
3R9	P76L,G77R,K99E	5	47	nts	nts
3R6	P76A,G77E	12	75	nts	nts
3R13	P76Q,G77W	4	42	nts	nts
3R14	P76D,G77A	3	41	nts	nts
3R15	P76Q,G77N	10	74	nts	nts
4R9	P76Q,G77K	5	50	nts	nts
4R20	P76Y	19	66	nts	nwt
3R22	P76L,K72M	6	61	nts	nwt
3R2	P76G,G77F	5	49	nts	nwt
4R43	P76G,G77F	1	55	nts	nwt
3R23	P76G,G77A	4	55	nts	nwt
4R19	P76S,G77L	2	34	nts	nwt
4R37	P76R,G77R	2	49	nts	nwt
3R1	P76W,G77D	9	40	nts	nwt
4R2	P76Q,G77S	2	21	nts	nwt
4R1	P76Q,G77R	4	52	nts	sts, nwt
4R5	P76T,G77R,K79T	2	32	nts	ts, nwt
4R4	P76I,G77S	5	16	nts	ts, nwt
4R36	P76A,G77V	2	28	nts	ts, nwt
4R3	P76R,G77V	2	17	ts	ts, nwt
3R29	P76V,G77R	2	24	ts	ts, nwt
3R12	P76R,G77E	14	57	nts	ts
3R16	P76R,G77L	2	38	nts	ts
3R18	P76R,G77L	4	34	nts	ts
3R20	P76G,G77S	3	65	sts	sts
3R21	P76G,G77Q	3	59	sts	sts
3R27	P76R,G77S	4	62	nts	bg
3R3	P76S,G77W	6	25	sts	bg
4R7	P76G,G77Y	2	53	sts	bg
3R24	P76G,G77R	3	44	ts	bg
3R10	P76G,G77R,T102S	5	49	ts	bg
3R26	P76W,G77K	2	48	ts	bg
4R24	P76W,G77A,K79T	2	30	ts	bg
4R10	P76E,G77V	2	21	ts	bg
4R6	P76V,G77T	3	43	ts, sg	bg
3R28	P76A,G77L	3	33	sts	ts, sg
4R14	P76F,G77A	4	35	sg	ts, nwt
4R8	P76S	59	63	sg	nwt

[a] Names and amino acid sequences are taken from ref *5*. [b] Relative growth is the cell density of each yeast strain after growth for 30 h at 30 °C, as determined by growth curves in liquid lactate medium, normalized to the growth of the wild-type strain (C93, 100%) and the *cyc1* deletion strain (C15, 0%). As the strains are isogenic, except at the *cyc1* locus, this number represents the functionality of the cytochrome *c* protein present in the strain. [c] Relative amount of cyt *c* is the amount of holocytochrome *c* present in the yeast cell at 30 °C, as determined by the absorbance peak at 548 nm from the low-temperature spectrum. The number presented is a percentage, normalized to the amount of protein present in the wild-type strain (C93, 100%) and the *cyc1* deletion strain (C15, 0%). [d] Temperature sensitivity on glycerol (YPG) and lactate (YPL) is indicated as follows: nts = not temperature sensitive; sts = slightly temperature sensitive; ts = temperature sensitive; bg = barely grew; sg = slow grower; nwt = not wild-type (for more detail about these classifications, see the text).

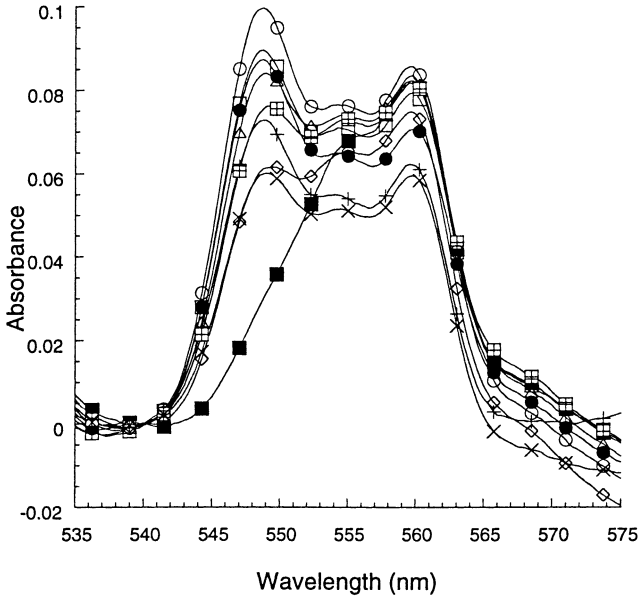

FIGURE 3: Representative low-temperature difference spectra of diploid strains expressing the Pro76,Gly77 mutations. Strains were grown in medium containing dextrose. Control strains: C93, ●; C15, ■. Experimental strains: 3R7, ○; 3R8, □; 3R4, △; 3R6, plus sign enclosed in a square; 3R11, +; 3R9, ◇; 3R10, ×. A summary of all data can be found in Table 3.

on any day at any temperature; sg, slow grower, strain grew more slowly at all temperatures over the course of 3 days; nwt, not wild type, strain did not grow quite as well as wild type on any day or temperature but could not be classified into the above groups. The results are presented in Table 3.

Fifteen strains are not temperature-sensitive by this assay on either glycerol or lactate. As expected from the growth curve and low-temperature experiments, one of these random mutants contains the wild-type sequence and eight more retain Gly77. The remaining six non-temperature-sensitive mutants contain large, hydrophilic or charged residues at either position 76 or 77 or, most often, both (Table 3).

Nine strains are classified as nwt on glycerol. Nine other mutants are classified as ts or ts,nwt on glycerol (Table 3). Fifty percent of the strains that are ts on glycerol contain an arginine at either position 76 or 77. The growth on lactate of most of these strains is unaffected, suggesting that these mutations have interfered with glycerol metabolism only.

Two strains, 3R20 and 3R21, expressing P76G,G77S and P76G,G77Q, respectively, are slightly temperature-sensitive on both media types. Most likely, these mutants are more thermolabile than wild-type protein. Nine strains were classified as bg on glycerol and exhibited varying phenotypes on lactate. Because of the very minimal amount of growth on glycerol displayed by these strains, it is likely that these mutations are exhibiting a real difference in interaction with the redox partners within the cell. Finally, three strains were not like any of the other strains and were classified sg on either glycerol or lactate. Cytochrome *c* mutations in these strains could have affected mitochondrial import, folding, stability, or function.

Circular Dichroism Spectroscopy of Gly77 Mutants. To determine if mutations at Pro76 were destabilizing the protein enough to cause the observed decrease in the amount of protein in vivo, proteins from the 3R and 4R strains containing an amino acid mutation only at Pro76 were

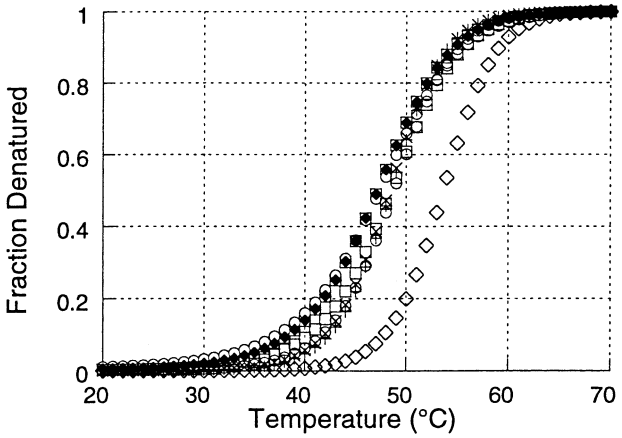

FIGURE 4: Representative thermal denaturation curves of proteins containing mutations just at Pro76 isolated from the 3R and 4R yeast strains. The CD signal at 222 nm was monitored from 4 to 80 °C; the fraction of denatured protein at each temperature was calculated as described in Materials and Methods and is plotted here against temperature. The data shown are from one typical experiment for each protein. Symbols: C102T, ◇; P76Y, ○; P76R, □; P76V, ×; P76G, +; P76W, ◆; P76L,K72M, △; P76S, plus sign enclosed in a square. To enhance visualization, only the central part (20−70 °C) of the denaturation curves is shown.

Table 4: Summary of in Vitro Analysis of Pro76 Mutants

yeast strain	residues 76 and 77	T_m[a] (°C)	$\Delta G_{D,298}$ (kcal/mol)
C93	Pro Gly	53.5 (± 0.36)	5.08
3R22	Leu Gly[b]	48.6 (± 0.76)	1.64
3R25	Gly Gly	48.5 (± 0.32)	4.30
3R7	Val Gly	48.3 (± 0.01)	4.01
3R17	Arg Gly	48.6 (± 0.34)	4.11
4R8	Ser Gly	47.8 (± 0.45)	3.19
4R20	Tyr Gly	46.7 (± 0.79	3.01
3R4	Trp Gly	47.0 (± 0.13)	2.70

[a] T_m is the thermal denaturation temperature determined by monitoring the change in mean residue ellipticity at 222 nm. T_m is the average melting temperature of at least two (usually three) experiments (± standard error). [b] This protein contains a second-site mutation (K73M; see Table 3).

purified. Far UV CD spectroscopy was performed to analyze the secondary structure and thermal denaturation properties of these Pro76 variants. For each variant, the shape of the far UV CD spectrum is similar to that of the wild-type protein (data not shown), indicating that the structure is not grossly perturbed and the regular secondary structure content is similar to that found in the wild-type protein.

The thermal denaturation of the seven Pro76 variants was monitored by following the mean residue ellipticity at 222 nm from 4 to 80 °C. The fraction denatured, melting temperature (T_m), and free energy at 25 °C ($\Delta G_{D,298}$) was calculated as described in Materials and Methods and the data are reported in Figure 4 and Table 4. The thermal denaturation curves of these seven proteins show an interesting property: all of these proteins denature at 47−48 °C, 5 °C below the melting temperature of the wild-type protein, which is 53 °C (*5*, *90*). The two variants with bulky, aromatic residues at position 76 (P76W and P76Y) have the lowest T_m (47.0 and 47.4 °C, respectively), but these T_ms are only about 1.5 °C lower than P76G, which exhibits a T_m of 48.5 °C. Thus, residue size at position 76 does not significantly contribute to the decrease in stability. Because the stability changes are consistent in most of the variants, it is likely that mutation of the imino acid proline at position 76 to an amino acid has increased the entropy of the denatured state, resulting in an apparent decrease in the stability of the native state.

DISCUSSION

Summary of Results. Pro76 and Gly77, the central two residues in a β-turn located in Ω-loop D of iso-1-cytochrome *c* (Figure 1), are completely invariant in the 106 eukaryotic cytochromes *c* that have been sequenced (Table 1). Directed, random mutagenesis of these two residues showed that a large number of amino acids can be substituted at these positions and at least a minimally functional cytochrome *c* was obtained (*37*). We have now analyzed the level of function and amount of structure in vivo in the yeast strains expressing these mutant proteins. In all cases except one, the amount of protein and its level of function are greatly decreased in these strains, which neatly explains why Pro76 and Gly77 are invariant during the evolution of this molecule. Now, we wish to ask why these two residues are so significant in the function of cytochrome *c*. Are these residues important for docking to one of the redox partners? Are they important for the stability of the protein structure? Are they important in the process of electron transfer itself?

Role of Iso-1-cytochrome c in Yeast Metabolism. Cytochrome *c* is essential in the aerobic metabolism of yeast. Its major role is in the transport of electrons from complex III (cytochrome bc_1 complex or cytochrome *c* reductase) to complex IV (cytochrome aa_3 complex or cytochrome *c* oxidase) in the electron transport chain. Complex IV ultimately passes the electrons to the terminal acceptor, oxygen, to form water (*77*). The electrons transported through complex III to cytochrome *c* ultimately come from the oxidation of NADH and $FADH_2$, small molecules that are produced by various pathways during oxidative metabolism. In yeast, iso-1-cytochrome *c* has four other physiological redox partners: lactate dehydrogenase (cytochrome b_2), cytochrome *c* peroxidase, cytochrome bc_1 complex (cytochrome *c* reductase), and cytochrome aa_3 complex (cytochrome *c* oxidase) (*77*). For the purposes of this work, we are particularly interested in the bc_1 complex, which is necessary for metabolism of glycerol, and lactate dehydrogenase, a molecule that oxidizes lactate to pyruvate, passes the electrons directly to cytochrome *c*, and is important for utilization of lactate as a carbon source.

Importance of Pro76 and Gly77 to the Function of Iso-1-cytochrome c. Any substitution at Pro76 or Gly77, except P76V, causes a significant decrease in the ability of yeast to grow on liquid lactate medium (Table 3). A trivial explanation for this result is that these mutations simply cause a decrease in the amount of protein present in the cell; however, our results refute this explanation because the function of the mutant proteins does not correlate with the amount of protein present in vivo (Figure 5). Earlier work has shown that the growth on lactate of a control strain producing as little as 50% of the wild-type protein is barely distinguishable from strains expressing the normal amount of wild-type protein (see the symbol **X** in Figure 5) (*92*, *93*). Furthermore, a strain producing only 11% of the level of wild-type protein can grow 20% as well as the wild-type strain (see symbol **Z** in Figure 5) (*80*). In contrast, in the

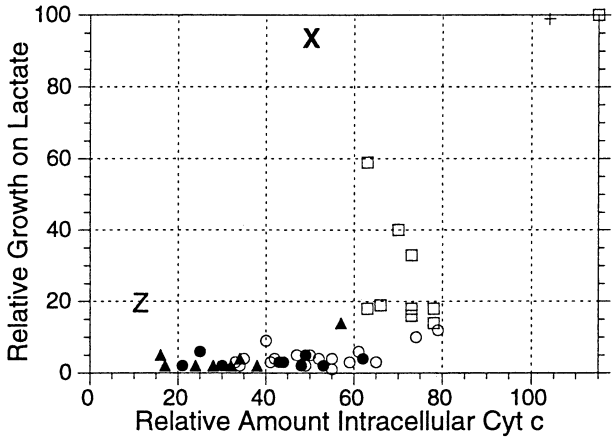

FIGURE 5: Correlation of the strain growth in liquid lactate medium versus the amount of variant cytochrome *c* present in vivo. (+) Strain 3R8, a random mutant expressing the wild-type Pro76,Gly77 sequence; (□) strains expressing variant cytochromes *c* containing mutations only at Pro76; (○) for other strains that are not temperature-sensitive on either YPG or YPL; (●) data for strains that are temperature-sensitive on solid glycerol medium; (▲) strains that barely grew on glycerol medium. The X is a data point taken from earlier work on a control strain expressing 50% of the wild-type protein (*92*, *93*); the Z is a data point taken from earlier work on a control strain expressing 11% of the wild-type protein (*80*). Protein function was determined by growth curves in liquid lactate medium at 30 °C over the course of 30 h. The amount of protein was calculated from low-temperature difference spectroscopy of whole cells grown on dextrose medium. Data are taken from Table 3.

current study, mutant strains producing 80−100% of the normal amount of cytochrome *c* are, except in one case, less than 60% functional (Table 3, Figure 5).

Another possible explanation for the decrease in function is that mutations at these positions change the redox potential of the protein. This explanation is also unlikely. Solvent accessibility and heme cavity polarity are the main determinants of the redox potential of cytochrome *c* heme. Consequently, changes at residues in the heme cleft cause significant changes in the redox potential of the protein; however, such changes are not necessarily correlated with the function of the protein (see, for example, refs *57*−*60* and *94*). Pro76 and Gly77 are not in contact with the heme, so a change in the side chain will not directly change the heme environment. Mutations at Pro76 and Gly77 might perturb the structure of loop D (part of which is in contact with the heme, Figure 1), thus indirectly causing a change in the heme environment; however, Pro76 and Gly77 are among the most exposed residues in iso-1-cytochrome *c*. Solvent accessibility calculations on the oxidized iso-1-C102T crystal structure (*44*) using the Access program (*95*) show that only 29% of the Pro76 and only 47% of the Gly77 residues are excluded from access to a 1.4 Å radius probe in the protein. Not including the N- and C-terminal residues, only 13 residues expose a larger percentage of their surface area than Pro76 and only 10 additional residues are more exposed than Gly77. Thus, changes in the side chains of these proteins should be easily accommodated at the surface of the protein without perturbing the loop D structure and affecting the redox potential of the heme group.

On the other hand, substitution at Gly77 might cause a change in the local backbone structure of loop D. In the crystal structures of both reduced (*43*) and oxidized (*44*) cytochrome *c*, the backbone dihedral angles of Gly77 lie in a region of the Ramachandran map that is inaccessible to any residue with a β-carbon (Table 2). Substitution of any other residue will require local rearrangement of the backbone, which could propagate through loop D and cause a change in the heme environment. However, in the NMR solution structure of horse cytochrome *c* (*51*), the backbone conformation of Gly77 falls into an allowed region in the Ramachandran plot (Table 2). So it is not clear that the Gly77 conformation in the crystal structures is the same as that found in solution or if the conformation in the crystal is the result of crystal packing interactions. Even so, it is unlikely that such a local perturbation could change the redox potential significantly enough to cause the changes in function that have been observed.

A third explanation for the large decrease in protein function which we observed upon substitution at Pro76 and Gly77 is that these residues are important for docking to the physiological redox partners. This is also unlikely. In a large number of protection and lysine-labeling studies, loop D lysines 72, 73, and 79 appear to be located only peripherally to the binding sites, if at all (reviewed in ref *77*). In the cocrystal structure of cytochrome *c* with cytochrome *c* peroxidase (*96*), loop D is at the periphery of the binding site; thus, changes at these residues should not have the huge functional effect observed for the mutations described here.

A final possible explanation for the large effect on cytochrome *c* function is that the unique conformation of Gly77 is important for redox partner docking or a switch in the conformation of Gly77 is responsible for redox partner undocking. Except for the NMR solution structure of horse cytochrome *c*, there is a clear switch in the conformation of Gly77 in going from the oxidized to the reduced form (Figure 1, Table 2). In all oxidized forms of the protein, except iso-1-cytochrome *c*, Pro76 and Gly77 are in a type II ($\beta_{\gamma L}$) β-turn conformation; in all reduced forms except iso-1, the backbone conformation of Gly77, but not Pro76, changes so that it no longer falls into one of the common turn categories previously described (*12*, *13*). In the oxidized and reduced crystal structures of iso-1-cytochrome *c*, the same two conformations are found, but the reduced protein exhibits the type II turn and the oxidized protein contains the unclassifiable turn (Figure 1, Table 2). Neither of the two conformations observed for Gly77 in the crystal structures is allowed for residues containing a β-carbon. Given the rather large effect of substituting any residue at Gly77 (Table 3, Figure 5), we suggest that the odd conformation of Gly77 might be important for docking of a redox partner or that the switch between the two conformations is necessary for docking or undocking of the redox partners. In support of this suggestion, Pielak and co-workers observed a difference in the protection factors of the Thr78 amide proton. This amide proton had exchanged out by the first data point (6.55 h) in the oxidized protein but exhibited significant protection in the reduced form of the protein (*67*). Others have also found that protons in this structure are more easily exchangeable in the oxidized rather than in the reduced form (*55*).

The conformation of these residues in the NMR solution structure of horse cytochrome *c* (*51*) does not support this suggestion because the conformation of Gly77 in this structure falls in an allowed region of the Ramachandran

plot (Table 2). It is not possible to determine if this solution conformation is real or if there were simply not enough constraints in the NMR data to accurately determine the structure in this region.

Differences in the Metabolism of Glycerol and Lactate. The temperature sensitivity studies suggest that some mutant strains are more sensitive to growth on glycerol than to growth on lactate (Table 3). There are several possible explanations for observing different behavior on the different media. First, the mutation could be reporting a difference in how cytochrome *c* is interacting with its various redox partners. In lactate metabolism in yeast, lactate dehydrogenase catalyzes the oxidation of lactate and concomitant reduction of cytochrome *c* (*97*), whereas all electrons produced by the metabolism of glycerol go through the cytochrome bc_1 complex to which cytochrome *c* binds (*77*). Electrons passed to cytochrome *c* by either glycerol or lactate metabolism are then passed to complex IV. A decrease in the functionality on glycerol, but not lactate, would suggest that the mutation has interfered with cytochrome *c* docking to the bc_1 complex but not to lactate dehydrogenase. An alternative, more trivial explanation is found in the composition of the medium. Glycerol is not a buffer, but lactate is. We have observed that the pH of unbuffered, synthetic liquid glycerol medium can drop precipitously during log phase growth of large fermentations of yeast, enough to inhibit the growth before the culture is saturated (T.L.B. and J.S.F., unpublished observations). Growth of cytochrome *c* on glycerol plates could cause a slight drop in pH around the yeast colonies that might destabilize the mutant protein enough to see inhibition in the colony growth, although we have not observed this pH drop on solid YPG medium.

If a pH change on glycerol medium was affecting the growth, one might expect that a temperature increase to 37 °C might exacerbate this effect, but this was not observed in the nine mutants classified as nts on lactate and nwt on glycerol, a result that suggests the pH drop is not occurring on the solid medium or pH is not affecting these strains. However, 11 other mutants were classified as nts on lactate and ts or ts,nwt on glycerol (Table 3). One interpretation of the results for these strains is that these variants are especially sensitive to pH and small pH changes are affecting their function.

On the other hand, the difference between growth on glycerol and on lactate is large for some strains and is clearly not attributable to any possible small pH changes. This includes one strain that is indistinguishable from wild type on lactate but barely grows on glycerol medium (Table 3). Some mutations at Pro76 and Gly77 have affected growth on glycerol (and consequent docking or interaction of cytochrome *c* to the bc_1 complex) more than they have affected growth on lactate (and docking to lactate dehydrogenase). Resolution of this possibility requires more quantitative study of in vivo growth in glycerol medium and in vitro binding studies to physiological redox partners.

Why Does P76V Function as Well as Wild-Type Protein and Why Is This Substitution Not Observed in Nature? The strain 3R7 expressing P76V produces as much protein as wild-type C93 and grows just as well on nonfermentable carbon sources (Table 3). Valine is the side chain whose volume most closely matches the volume of the proline side chain (*26*); thus, valine should not drastically interfere during docking with physiological redox partners. On the other hand, in a survey of turn types, it was found that valine is strongly disfavored at position $i + 1$ in a type II β-turn. The residue most often found in that position is proline; the second most often occurring residue is lysine (*13*). One interpretation of these data is that valine is not typically found at position $i + 1$ because it destabilizes type II turns. If so, then the functionality of the P76V variant argues that the type II turn is not important for the function of the protein, as we suggested above. However, the lack of valines in type II turns in the database of protein crystal structures could be a result of other factors. For instance, turns are often found at the protein surface. Valine might be disfavored here because of its hydrophobicity. In fact, the lack of valine at position $i + 1$ is also found to be statistically significant in type I, I′, and II′ β-turns (*13*).

It has recently been shown that valine does destabilize a turn between two helices in the four-helix bundle protein Rop by 0.4 kcal/mol (*4*). Removal of neighboring side chains caused the stability to decrease further, showing that there are both intrinsic and context-dependent effects upon substitution of valine into a turn (*4*). In vitro evolution of stable turns has demonstrated that turn sequences favored in the database are the most thermodynamically favorable (*6*). Therefore, if the single substitutions at Pro76 were selected for thermodynamic reasons, we would expect that the amount of protein present in the cell would correlate with intrinsic turn preferences and that lysine, not valine, would be most favorable at this position. No correlation could be found between the published type II β-turn preferences (*13*) and the amount of holocytochrome *c* found in vivo, further suggesting that substitution of valine at Pro76 was not chosen for thermodynamic reasons.

It is interesting to note that results similar to those found for Pro76 were also found for Pro71 in iso-1-cytochrome *c*. This highly conserved residue was replaced by leucine, isoleucine, valine, serine, and threonine. It was found that all mutations except valine resulted in a significant decrease in the in vivo function of the protein; the amount of intracellular protein was decreased by about 50% in all cases (*93*).

It is possible that the Val76 variant is not observed in natural eukaryotic sequences (Table 1) because this event is unlikely to occur during the natural selection process of evolution. The proline-to-valine mutation requires two nucleotide changes within the codon. The simultaneous, random replacement of two consecutive nucleotides would be an extremely rare event. The more likely event evolutionarily would be to change one nucleotide and then in a later event change the second nucleotide. However, our data demonstrate that most any single-nucleotide change will cause a decrease in the cytochrome *c* function (Table 3). Such an event would put the organism at a disadvantage; thus, the single-nucleotide mutation is unlikely to become fixed in the population.

Conclusions. We have shown that the reason for evolutionary invariance of Pro76 and Gly77 in the sequences of eukaryotic cytochromes *c* is caused by a large decrease in the function of the protein upon replacement of these residues. This decrease in function is not due to a concerted decrease in the amount of protein present in the cell. Instead, the data suggest that the mutations interfere with redox

partner docking or with an important conformational change that occurs at Gly77 upon change of the redox state. No real conclusions can be drawn from the identity of amino acids that might be important for redox partner binding. Such information awaits in vitro binding and kinetic studies on these mutant proteins. Most of the mutations at Pro76 and Gly77 cause a decrease in the amount of holoprotein in the cell. This decrease is not correlated with the thermostability of the variant protein produced by the strains.

ACKNOWLEDGMENT

We thank Gary Pielak for the alignment of the 106 eukaryotic cytochromes *c* and for critical discussions. We also thank Robert Stack and Charles Hauer of the Wadsworth Research Laboratories, NYS Department of Health, for performing the mass spectrometry experiments. J.S.F. thanks Jeffrey Skolnick for support during her sabbatical and Susan Wynant for proofreading the manuscript.

REFERENCES

1. Leszczynski, J. F., and Rose, G. D. (1986) *Science 234*, 849−855.
2. Ptitsyn, O. B. (1981) *FEBS Lett. 131*, 197−201.
3. Wright, P. E., Dyson, H. J., and Lerner, R. A. (1988) *Biochemistry 27*, 7167−7175.
4. Predki, P. F., Agrawal, V., Brünger, A. T., and Regan, L. (1996) *Nat. Struct. Biol. 3*, 54−58.
5. Fetrow, J. S., Horner, S. R., Oehrl, W., Schaak, D. L., Boose, T. L., and Burton, R. E. (1997) *Protein Sci. 6*, 195−208.
6. Zhou, H. X., Hoes, R. H., and DeGrado, W. F. (1996) *Nat. Struct. Biol. 3*, 446−451.
7. Venkatachalam, C. M. (1968) *Biopolymers 6*, 1425−1436.
8. Lewis, P. N., Momany, F. A., and Scheraga, H. A. (1973) *Biochim. Biophys. Acta 303*, 211−229.
9. Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−339.
10. Rose, G. D., Gierasch, L. M., and Smith, J. A. (1985) *Adv. Protein Chem. 37*, 1−109.
11. Chou, P. Y., and Fasman, G. D. (1977) *J. Mol. Biol. 115*, 135−175.
12. Wilmot, C. M., and Thornton, J. M. (1990) *Protein Eng. 3*, 479−493.
13. Hutchinson, E. G., and Thornton, J. M. (1994) *Protein Sci. 3*, 2207−2216.
14. Milner-White, E. J., and Poet, R. (1986) *Biochem. J. 240*, 289−292.
15. Sibanda, B. L., and Thornton, J. M. (1985) *Nature 316*, 170−174.
16. Sibanda, B. L., Blundell, T. L., and Thornton, J. M. (1989) *J. Mol. Biol. 206*, 759−778.
17. Efimov, A. V. (1991) *Protein Eng. 4*, 245−250.
18. Wintjens, R. T., Rooman, M. J., and Wodak, S. J. (1996) *J. Mol. Biol. 255*, 235−253.
19. Brunet, A. P., Huang, E. S., Huffine, M. E., Loeb, J. E., Weltman, R. J., and Hecht, M. H. (1993) *Nature 364*, 355−358.
20. Predki, P. F., and Regan, L. (1995) *Biochemistry 34*, 9834−9839.
21. Chou, K.-C., Maggiora, G. M., and Scheraga, H. A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7315−7319.
22. Hynes, T. R., Kautz, R. A., Goodman, M. A., Gill, J. F., and Fox, R. O. (1989) *Nature 339*, 73−76.
23. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., and Wright, P. E. (1988) *J. Mol. Biol. 201*, 161−200.
24. Milburn, P. J., Konishi, Y., Meinwald, Y. C., and Scheraga, H. A. (1987) *J. Am. Chem. Soc. 109*, 4486−4496.
25. Shin, H.-C., Merutka, G., Waltho, J. P., Wright, P. E., and Dyson, H. J. (1993) *Biochemistry 32*, 6348−6355.
26. Chothia, C., and Lesk, A. M. (1987) *J. Mol. Biol. 196*, 901−917.
27. Tramontano, A., Chothia, C., and Lesk, A. M. (1989) *Proteins: Struct., Funct., Genet. 6*, 382−394.
28. Tramontano, A., and Lesk, A. M. (1992) *Proteins: Struct., Funct., Genet. 13*, 231−245.
29. Urfer, R., and Kirschner, K. (1992) *Protein Sci. 1*, 31−45.
30. Scheerlinck, J.-P. Y., Lasters, I., Claessens, M., DeMaeyer, M., Pio, F., Delhaise, P., and Wodak, S. J. (1992) *Proteins: Struct., Funct., Genet. 12*, 299−313.
31. Edwards, M., Sternberg, M., and Thornton, J. (1987) *Protein Eng. 1*, 173−181.
32. Ring, C. S., Kneller, D. G., Langridge, R., and Cohen, F. E. (1992) *J. Mol. Biol. 224*, 685−699.
33. Fetrow, J. S. (1995) *FASEB J. 9*, 708−717.
34. Kwasigroch, J.-M., Chomilier, J., and Mornon, J.-P. (1996) *J. Mol. Biol. 259*, 855−872.
35. Donate, L. E., Rufino, S. D., Canard, L. H. J., and Blundell, T. L. (1996) *Protein Sci. 5*, 2600−2616.
36. Oliva, B., Bates, P. A., Querol, E., Avilés, F. X., and Sternberg, M. J. E. (1997) *J. Mol. Biol. 266*, 814−830.
37. Mulligan-Pullyblank, P., Spitzer, J. S., Gilden, B. M., and Fetrow, J. S. (1996) *J. Biol. Chem. 271*, 8633−8645.
38. Sherman, F., Stewart, J. W., Parker, J. H., Inhaber, E., Shipman, N. A., Putterman, G. J., Gardishy, R. L., and Margoliash, E. (1968) *J. Biol. Chem. 243*, 5446−5456.
39. Sherman, F., Stewart, J. W., Jackson, M., Gilmore, R. A., and Parker, J. H. (1974) *Genetics 77*, 255−284.
40. Hampsey, D. M., Das, G., and Sherman, F. (1986) *J. Biol. Chem. 261*, 3259−3271.
41. Auld, D. S., and Pielak, G. J. (1991) *Biochemistry 30*, 8684−8690.
42. Fredericks, Z. L., and Pielak, G. J. (1993) *Biochemistry 32*, 929−936.
43. Louie, G. V., and Brayer, G. D. (1990) *J. Mol. Biol. 214*, 527−555.
44. Berghuis, A. M., and Brayer, G. D. (1992) *J. Mol. Biol. 223*, 959−976.
45. Murphy, M. E. P., Nall, B. T., and Brayer, G. D. (1992) *J. Mol. Biol. 227*, 160−176.
46. Bushnell, G. W., Louie, G. V., and Brayer, G. D. (1990) *J. Mol. Biol. 214*, 585−595.
47. Takano, T., and Dickerson, R. E. (1981) *J. Mol. Biol. 153*, 79−94.
48. Takano, T., and Dickerson, R. E. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 6371−6375.
49. Ochi, H., Hata, V., Tanaka, N., Kakudo, M., Sakurai, T., Aihara, S., and Morita, Y. (1983) *J. Mol. Biol. 166*, 407−418.
50. Tanaka, N., Yamane, T., Tsukihara, T., Ashida, T., and Kakudo, M. (1975) *J. Biochem. (Tokyo) 77*, 147.
51. Qi, P. X., DiStefano, D. L., and Wand, A. J. (1994) *Biochemistry 33*, 6408−6417.
52. Qi, P. X., Beckman, R. A., and Wand, A. J. (1996) *Biochemistry 35*, 12275−12286.
53. Banci, L., Bertini, I., Gray, H. B., Luchinat, C., Reddig, T., Rosato, A., and Turano, P. (1997) *Biochemistry 36*, 9867−9877.
54. Baistrocchi, P., Banci, L., Vertini, I., Turano, P., Bren, K. L., and Gray, H. B. (1996) *Biochemistry 35*, 13788−13796.
55. Banci, L., Bertini, I., Bren, K. L., Gray, H. B., Sompornpisut, P., and Turano, P. (1997) *Biochemistry 36*, 8992−9001.
56. Berghuis, A. M., Guillemette, J. G., McLendon, G., Sherman, F., Smith, M., and Brayer, G. D. (1994) *J. Mol. Biol. 236*, 786−799.
57. Berghuis, A. M., Guillemette, J. G., Smith, M., and Brayer, G. D. (1994) *J. Mol. Biol. 235*, 1326−1341.
58. Lo, T. P., Murphy, M. E. P., Guillemette, J. G., Smith, M., and Brayer, G. D. (1995) *Protein Sci. 4*, 198−208.
59. Lo, T. P., Guillemette, J. G., Louie, G. V., Smith, M., and Brayer, G. D. (1995) *Biochemistry 34*, 163−171.
60. Louie, G. V., Pielak, G. J., Smith, M., and Brayer, G. D. (1988) *Biochemistry 27*, 7870−7876.
61. Louie, G. V., and Brayer, B. D. (1989) *J. Mol. Biol. 210*, 313−322.

62. Murphy, M. E. P., Fetrow, J. S., Burton, R. E., and Brayer, G. D. (1993) *Protein Sci. 2*, 1429−1440.
63. Chau, M.-H., Cai, M. L., and Timkovich, R. (1990) *Biochemistry 29*, 5076−5087.
64. Gao, Y., Boyd, J., Williams, R. J. P., and Pielak, G. J. (1990) *Biochemistry 29*, 6994−7003.
65. Gao, Y., Boyd, J., Pielak, G. J., and Williams, R. J. P. (1991) *Biochemistry 30*, 1928−1934.
66. Hong, X., and Dixon, D. W. (1989) *FEBS Lett. 246*, 105−108.
67. Marmorino, J. L., Auld, D. S., Betz, S. F., Doyle, D. F., Young, G. B., and Pielak, G. J. (1993) *Protein Sci. 2*, 1966−1974.
68. Pielak, G. J., Boyd, J., Moore, G. R., and Williams, R. J. P. (1988) *Eur. J. Biochem. 177*, 167−177.
69. Baxter, S. M., Boose, T. L., and Fetrow, J. S. (1997) *J. Am. Chem. Soc. 119*, 9899−9900.
70. Bai, Y., Sosnick, T. R., Mayne, L., and Englander, S. W. (1995) *Science 269*, 192−197.
71. Bai, Y., and Englander, S. W. (1996) *Proteins: Struct., Funct., Genet. 24*, 145−151.
72. Roder, H., Elove, G. A., and Englander, S. W. (1988) *Nature 335*, 700−704.
73. Colón, W., Elöve, G. A., Wakem, L. P., Sherman, F., and Roder, H. (1996) *Biochemistry 35*, 5538−5549.
74. Elöve, G. A., Bhuyan, A. K., and Roder, H. (1994) *Biochemistry 33*, 6925−6935.
75. Qu, K., Vaughn, J. L., Sienkiewicz, A., Scholes, C. P., and Fetrow, J. S. (1997) *Biochemistry 36*, 2884−2897.
76. Pierce, M. M., and Nall, B. T. (1997) *Protein Sci. 6*, 618−627.
77. Pettigrew, G. W., and Moore, G. R. (1987) *Cytochromes c Biological Aspects*, Springer-Verlag, Berlin.
78. Pielak, G. J., Auld, D. S., Beasley, J. R., Betz, S. F., Cohen, D. S., Doyle, D. F., Finger, S. A., Fredericks, Z. L., Hilgen-Willis, S., Saunders, A. J., and Trojak, S. K. (1995) *Biochemistry 34*, 3268−3276.
79. Fumo, G., Spitzer, J. S., and Fetrow, J. S. (1995) *Gene 164*, 33−39.
80. Fetrow, J. S., Cardillo, T. S., and Sherman, F. (1989) *Proteins: Struct., Funct., Genet. 6*, 372−381.
81. Cutler, R. L., Pielak, G. J., Mauk, A. G., and Smith, M. (1987) *Protein Eng. 1*, 95−99.
82. Schweingruber, M. E., Sherman, F., and Stewart, J. W. (1977) *J. Biol. Chem. 252*, 6577−6580.
83. Sherman, F., and Slonimski, P. P. (1964) *Biochim. Biophys. Acta 90*, 1−15.
84. Cottrell, S. F., Rabinowitz, M., and Getz, G. S. (1975) *J. Biol. Chem. 250*, 4087−4094.
85. Linske-O'Connell, L. I., Sherman, F., and McLendon, G. (1995) *Biochemistry 34*, 7094−7102.
86. Schagger, H., and von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
87. Margoliash, E., and Frohwirt, N. (1959) *Biochem. J. 71*, 570−572.
88. Strickland, E. H. (1974) *CRC Crit. Rev. Biochem. 2*, 113−174.
89. Elwell, M. L., and Schellman, J. A. (1977) *Biochim. Biophys. Acta 494*, 367−383.
90. Cohen, D. S., and Pielak, G. J. (1994) *Protein Sci. 3*, 1253−1260.
91. Moore, G. R., and Pettigrew, G. W. (1990) *Cytochromes c. Evolutionary, Structural and Physicochemical Aspects*, Springer-Verlag: Berlin.
92. Schweingruber, M. E., Stewart, J. W., and Sherman, F. (1979) *J. Biol. Chem. 254*, 4132−4143.
93. Ernst, J. F., Hampsey, D. M., Stewart, J. W., Rackovsky, S., Goldstein, D., and Sherman, F. (1985) *J. Biol. Chem. 260*, 13225−13236.
94. Luntz, T. L., Schejter, A., Barber, E. A. E., and Margoliash, E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3524−3528.
95. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.
96. Pelletier, H., and Kraut, J. (1992) *Science 258*, 1748−1755.
97. Appleby, C. A., and Morton, R. K. (1954) *Nature (London) 173*, 749−752.

BI972279A