Review Letter

Yeast iso-l-cytochrome c: genetic analysis of structural requirements

D. Michael Hampsey, Goutam Das[†] and Fred Sherman[†]

Departments of Biophysics and [†]Biochemistry, University of Rochester Medical Center, Rochester, NY 14642, USA

We describe the use of classical and molecular genetic techniques to investigate the folding, stability, and enzymatic requirements of iso-l-cytochrome c from the yeast *Saccharomyces cerevisiae*. Interpretation of the defects associated with an extensive series of altered forms of iso-l-cytochrome c was facilitated by the recently resolved three dimensional structure of iso-l-cytochrome c [(1987) J. Mol. Biol. 199, 295–314], and by comparison with the phylogenetic series of eukaryotic cytochromes c. Residue replacements that abolish iso-l-cytochrome c function appear to do so by affecting either heme attachment or protein stability; no replacements that abolish electron transfer function without affecting protein structure were uncovered. Most nonfunctional forms retained at least partial covalent attachment to the heme moiety; heme attachment was abolished only by replacements of Cys19 and Cys22, which are required for thioether linkage, and His23, a heme ligand. Replacements were uncovered that retain function at varying levels, including replacements at evolutionarily conserved positions, some of which were structurally and functionally indistinguishable from wild type iso-l-cytochrome c.

1. INTRODUCTION

Since its discovery in 1930 by Keilin [2], cytochrome c has been a preferred heme-protein for the investigation of a wide range of biological phenomena, from the mechanism of electron transport to the evolution of proteins. Because of its ease of purification, cytochrome c has been subjected to numerous studies designed to characterize its structural, physical, and enzymatic properties. To date the complete primary sequence of 92 eukaryotic cytochromes c has been reported, providing more comparative sequence information than for any other protein [3]. Moreover, the tertiary structure of cytochrome c has been solved from five eukaryotic organisms [1,4–9], with refinement to 1.5 Å resolution in the case of reduced tuna cytochrome c. This wealth of information provides a unique opportunity for

investigating the structural requirements of specific amino acid residues in cytochrome c.

2. YEAST ISO-1-CYTOCHROME c

We have been systematically investigating the role of specific amino acid residues for mitochondrial import, heme attachment, protein folding, and enzymatic activity of cytochrome c from the yeast *Saccharomyces cerevisiae*. Yeast cytochrome c is ideally suited for such studies because of the extensive repertoire of genetic and biochemical techniques that can be exploited with this organism. Accordingly, altered forms of yeast cytochrome c, consisting simply of single amino acid substitutions, can be readily generated.

There are two isozymes of cytochrome c in *S. cerevisiae*, iso-1-cytochrome c and iso-2-cytochrome c, which are encoded by the *CYC1* and *CYC7* genes, respectively; iso-1-cytochrome c is the predominant isozyme, accounting for approx. 95% of the total yeast complement of cytochrome

c under normal physiological conditions. Early investigations of iso-1-cytochrome *c* were particularly successful as a result of the development of convenient genetic techniques for selecting nonfunctional as well as functional forms of the protein [10], and for estimating the sites of the alterations by genetic fine-structure mapping [11]. Currently, molecular genetic techniques permit the rapid retrieval and DNA sequence analysis of mutations in the genome of yeast [3,12] so that altered forms of iso-1-cytochrome *c* can be readily defined. In addition, specific amino acid replacements can be made, either by site-directed mutagenesis with bacterial vectors [13,14] or by transforming yeast directly with synthetic oligonucleotides [15]. An extensive collection of amino acid replacements in iso-1-cytochrome *c* has been uncovered using these various techniques. Furthermore, the tertiary structure of iso-1-cytochrome *c* was recently solved to 2.8 Å resolution [1], which now enables us to rigorously interpret these amino acid replacements.

3. DELETERIOUS AMINO ACID REPLACEMENTS

Approximately 500 *cyc1* mutants either lacking iso-1-cytochrome *c* or containing nonfunctional forms of iso-1-cytochrome *c* have been isolated, primarily by a method that uses selection on chlorolactate medium [10]. The degree of deficiency of iso-1-cytochrome *c* or the degree of its activity was estimated by the amount of growth on lactate medium. Genetic analyses of these mutants, and DNA sequencing of representative alleles, have exposed an array of missense mutants, i.e., mutations that produce amino acid replacements. So far a total of 44 *cyc1* missense mutants, representing 30 single amino acid replacements at 16 different residue positions in iso-1-cytochrome *c*, have been characterized. One additional mutant was the result of a 14 residue deletion encompassing amino acids 45–58. Each of these single amino acid replacements was correlated to the intracellular level of iso-1-cytochrome *c* and its degree of low-level function. The alterations and the phenotypes of representative *cyc1* mutants are summarized in table 1. All of the *cyc1* missense mutants either completely lack or contain reduced amounts of iso-1-cytochrome *c*, as determined by

spectral analysis of whole cells, especially when grown at the optimum temperature of 30°C rather than at the reduced temperature of 22°C. These results indicate that the altered iso-1-cytochromes *c* are thermolabile or their synthesis is thermolabile. Furthermore, several of the nonfunctional iso-1-cytochromes *c* were labile in vitro. In fact, none of the nonfunctional iso-1-cytochromes *c* was present at the normal cellular level or retained normal stability in vitro. Therefore these replacements appear to be affecting either heme attachment or stability of the protein. Surprisingly, we have never found an amino acid replacement that abolishes the electron transfer function of iso-1-cytochrome *c*, while otherwise retaining normal stability.

3.1. Replacements of residues required for heme attachment

Low-temperature (−196°C) spectroscopic examination of intact cells conveniently reveals the C_α-band, which is dependent upon the covalent attachment of the heme moiety to the apo-protein of iso-1-cytochrome *c*. Using the spectral properties of cytochrome *c*, the nonfunctional missense mutants were categorized as either retaining or lacking the heme group. Unexpectedly, most amino acid replacements that abolish function still allow at least some degree of heme attachment (table 1). The only amino acid replacements uncovered that completely prevent heme attachment to the apo-protein are replacements of Cys19 and Cys22, which are required for covalent heme attachment, and His23, which is one of the two heme axial ligands. Even amino acid replacements of Met85, the second heme axial ligand, result in clearly detectable, albeit low, levels of iso-1-cytochrome *c*.

Apo-cytochrome *c* is encoded in the nucleus, synthesized in the cytoplasm, and subsequently translocated to the mitochondrial intermembrane space where cytochrome *c* heme lyase covalently attaches heme to yield holo-cytochrome *c* [16,17]. Hennig et al. demonstrated that the heme analog, deuterohemin, prevented both heme attachment and import of the apo-cytochrome *c* into mitochondria, thereby indicating that these two processes are coupled [16]. It is reasonable to expect that Cys19, Cys22 and His23 play a key role in import as well as heme attachment.

Table 1

Alterations of the missense mutants, the corresponding amounts and activities of iso-1-cytochrome c, and the probable defect

Normal	Mutant	Allele	% Iso-1-cyt.c^a	Growth on lactate[b]	Probable defect
Gly11 (GGT)	Ser11 (AGT)	cyc1-25	70	4	side-chain size
	Cys11 (TGT)	cyc1-164	70	2	
	Asp11 (GAT)	cyc1-523	45	3	
Leu14 (CTT)	Pro14 (CCT)	cyc1-115	30	1	torsion angle
Cys19 (TGT)	Tyr19 (TAT)	cyc1-81	<5	0	heme attachment
	Phe19 (TTT)	cyc1-88	<5	0	
	Trp19 (TGG)	cyc1-304	<5	0	
Cys22 (TGC)	Tyr22 (TAC)	cyc1-34	<5	0	heme attachment
	Phe22 (TTC)	cyc1-20	<5	0	
His23 (CAC)	Tyr23 (TAC)	cyc1-14	<5	0	heme attachment
	Asn23 (AAC)	cyc1-613	<5	0	
Gly34 (GGT)	Ser34 (AGT)	cyc1-73	60	0	side-chain size
	Asn34 (AAT)	cyc1-188	45	0	
	Asp34 (GAT)	cyc1-56	30	0	
Pro35 (CCA)	Leu35 (TTA)	cyc1-21	60	2	torsion angle
	Leu35 (CTA)	cyc1-59	80	1.5	
Leu37 (TTG)	Ser37 (TCG)	cyc1-63	50	0	hydrophobic heme pocket
	Trp37 (TGG)	cyc1-522	50	0	
His38 (CAT)	Pro38 (CCT)	cyc1-190	70	2	torsion angle
Trp64 (TGG)	Gly64 (GGG)	cyc1-526	60	0	hydrophobic heme pocket
	Leu64 (TTG)	cyc1-690	75	7	
	Ser64 (TCG)	cyc1-166	60	0	
	Cys64 (TGT)	cyc1-189	80	1.5	
Tyr72 (TAC)	Asp72 (GAC)	cyc1-521	60	0	hydrophobic heme pocket
Leu73 (TTG)	Ser73 (TCG)	cyc1-57	60	2	hydrophobic heme pocket
Pro76 (CCA)	Leu76 (TTA)	cyc1-24	90	0	side-chain size
	Leu76 (CTA)	cyc1-58	90	0	
Met85 (ATG)	Arg85 (AGG)	cyc1-42	20	0	heme axial ligand
	Thr85 (ACG)	cyc1-654	20	0	
	Ile85 (ATT)	cyc1-620	20	1	
Leu99 (TTA)	Ser99 (TCA)	cyc1-107	90	3	hydrophobic heme pocket
Leu103 (TTG)	Ser103 (TCG)	cyc1-62	90	4	hydrophobic heme pocket

[a] Iso-1-cytochrome c content after growth at 22°C relative to the amount (100%) present in the D311-3A strain containing the normal protein at the normal level

[b] Level of growth of the cyc1-363/cyc1-x diploid strain after 4 days on lactate medium at 22°C. The cyc1-363 allele contains a deletion of the entire CYC1 locus; cyc1-363/cyc1-x represents a series of hemizygous strains that produced clearer results for the assignment of the degree of growth. The values 0 through 7 represent an arbitrary scale where 0 denotes no detectable growth and 7 denotes a degree of growth only slightly below that observed with the cyc1-363/CYC1+ control strain. (From Hampsey et al. [3] and Hampsey, D.M., Das, G. and Sherman, F., unpublished results.)

3.2. *Replacements of residues involved in formation of the hydrophobic heme pocket*

In addition to replacements of residues involved in direct attachment of the apo-protein to heme, replacements were found at other residues that appear to interact with the heme group to disrupt folding or function of iso-1-cytochrome c. These include replacements of the normal Leu37, Trp64,

Tyr72, Leu73, Leu99, and Leu103 (table 1). The tertiary structure of iso-1-cytochrome c reveals that each of these hydrophobic residues directly constitutes part of the lining of the heme pocket [1]. Nonfunctional replacements of Leu14 and Pro35, which are other residues in direct heme contact, were also found. Loss of function by the Leu14 $\longrightarrow$ Pro14 and Pro35 $\longrightarrow$ Leu35 replacements, however, are more likely attributable to the effects of proline on the torsion angle of the protein backbone, rather than to the effects of these replacements on the hydrophobicity of the heme pocket. Proline replacements are discussed below.

3.3. Replacements of residues that facilitate overall protein folding

Loss of function of the remaining altered iso-1-cytochromes c can be attributed to amino acid replacements that disrupt proper folding of the protein; these fall into two classes: (i) those that replace normal glycine residues; and (ii) those that involve substitutions either to or from a proline residue.

The normal Gly11 and Gly34 residues are located in extremely electron dense regions of the protein. Loss of function by replacements of these residues can therefore be attributed to the structural distortion resulting from the inability of the protein to accommodate a side chain in these regions that is larger than the glycine proton. Gly11 is invariant and Gly34 is highly conserved in the phylogenetic series (fig.1). Presumably these residues are strictly conserved, not because of their contribution to directing the folding of the protein backbone, but simply because they are the only

residues small enough to fit into their respective regions of the mature protein.

We have uncovered four replacements involving proline substitutions. Two of these are replacements of normal residues by proline, Leu14 $\longrightarrow$ Pro14 and His38 $\longrightarrow$ Pro38; the other two are replacements of normal prolines, Pro35 $\longrightarrow$ Leu35 and Pro76 $\longrightarrow$ Leu76. With the exception of the Pro76 replacement, we attribute loss of function of each to the unique structural conformations imposed on the protein backbone by proline residues; loss of function resulting from the Pro76 $\longrightarrow$ Leu76 replacement is attributed to the leucine side chain, and is discussed below. Notably, the Leu14 and His38 substitutions are the only replacements in nonfunctional iso-1-cytochromes c that are not at evolutionarily conserved positions.

4. DISPENSABLE AMINO ACID RESIDUES

Previous results describing amino acid replacements in iso-1-cytochrome c clearly established the dispensability of many residues throughout the molecule. By selecting $cyc1$ revertants on lactate medium [10], arrays of single and multiple replacements that at least partially restore the function of iso-1-cytochrome c can be generated. These replacements, compiled in fig.1 and summarized in fig.2, represent iso-1-cytochromes c with at least 20% of the normal specific activity. The majority of these replacements occur at evolutionarily variant residue positions, although several, as described below, occur at invariant or highly conserved positions, some of which are nearly indistinguishable from the normal protein.

Fig.1. Composite of amino acid sequences of eukaryotic cytochromes c and of mutationally altered iso-1-cytochromes c. The normal iso-1-cytochrome c sequence is presented as a continuous sequence, with the amino-terminal residue assigned position 1. (In the vertebrate numbering system, Gly6 is assigned position 1; therefore, conversion to the vertebrate numbering system requires subtraction of 5.) The other residues found at each position in 91 other species are listed above the normal iso-1-cytochrome c sequence; the amino acid replacements in mutationally altered forms of iso-1-cytochrome c are listed below the line. The positions of amino termini in the phylogenetic series and in the mutants are indicated, respectively, by arrows above and below the normal sequence; Δ denotes deletion of the particular amino acid. The single residue replacements uncovered in the nonfunctional mutants described in table 1 are denoted by underlines and flanking vertical bars; those replacements abolishing function, yet retaining detectable levels of iso-1-cytochrome c, are denoted by single underlines, whereas replacements at positions 19, 22 and 23, associated with the absence of iso-1-cytochrome c, are denoted by double underlines. The two deletions in the $cyc1$-453 mutant are denoted by underlines within the normal sequence; loss of function is attributed to the Ser45-Ile58 deletion (see text). The phylogenetic series and the amino acid replacements of iso-1-cytochrome c appear in [3], with the exception of the Thr85 nonfunctional replacement and the Lys53, Leu81, Leu96 and Gly96 functional replacements, which are reported here.

```
                →                              Pro                      Met
                Gln                            Lys                      Phe
                → Glu →                        Glx Phe                  Lys
        Pro Pro Asn Pro              Pro       Glu Glx Arg  Arg         Glu Gln      Gln
        → Lys Phe Asp Gln Tyr Pro Lys          Asn Glu Asn Lys Arg      Asp Glu      Glu
        → Lys Glu Arg Ser Asx Ile Glu Gln      Asp Asp Thr Gln Lys      Thr Asn      Asn
        → Pro Glu Thr Glu Val Thr Leu Asn Glu  Glx Ser Thr Ser Glu Asn Thr  Ser Ser      Ser Glu
        → Pro Lys Ser Ser Ala Ser Val Ser Asp  Asn Ile Ser Val Asp Asp Ile  Ile Ala Lys  Ala Thr
        Gly Lys Ala Ala Ala Gly Gly Gly Ala Ala Ser → Asp Val Ala Ala Gly Ser Val  Val Gly Gln Ala Gly Ala
                                1          5                10               15          20
        Thr-Glu-Phe-Lys-Ala-Gly-Ser-Ala-Lys-Lys-Gly-Ala-Thr-Leu-Phe-Lys-Thr-Arg-Cys-Leu-Gln-Cys-His-
Thr Thr Glu Lys Ala Gly Gly Ala Ala Ser Ala Ala Gly Leu Ala |Ser|Ser Asn Ala      Ile      |Phe|Gln Leu |Phe||Asn|
→ Glu Phe Phe Val Ala Leu Leu Ile Thr Val Leu Val Ser Glu |Asp|Thr        Ser            |Tyr|      Ser |Tyr||Tyr|
          →     Met Leu Leu Ser Ile Glu Gln Leu Ile Leu Thr |Cys|          Thr            |Trp|      Lys
                → Ile Ile Thr Ser Gln Lys Thr Thr Phe Glu            Asn                         Tyr
                  Thr Asn Asp Thr Arg His Lys Glu Tyr Gln            His
                  Lys Glu Asn Asn Met Phe Arg Phe Pro His            Phe
                  Arg Met Gln Glu        Pro → Cys     Arg           |Pro|
                  Met → Lys Lys        →       Met     Phe
                  →     Phe →        →         Pro     Tyr
                        Tyr →                  →       Trp
                        →                              →
```

```
                      Δ                              Trp                                 Pro
    Cys     Arg       Cys                            Tyr                                 Lys
    Tyr     Gln       Lys                            Phe                       Arg       Gln
    Glu Gln Glu       Glu       Gln                  Gln           Δ           Lys       Glu Asn
    Ile Asn Asx Δ Δ Ile Gln     Glu                  Asn       Phe Tyr Δ       Gln Cys   Thr Asp          Lys
Ser Leu Asp Asn Lys Glu Leu Asx Thr Thr      Ser     Asp       Leu Ile Asn     Glx Gln   Lys Ser Ser      Thr
Ala Ala Ala Val Asn Leu Ala Asn Ser Ile      Val     Ser       Val Leu Ser     Thr Thr   Thr Ile Val Δ Pro Val
Gly Gly Gly Ala Ala Ala Gly Ser Gly Ala Ala  Ala     Gly       Ala Val Val     Ala Ala   Ser Val Ala Ser Phe Ala
    25              30                  35             40               45               50
Thr-Val-Glu-Lys-Gly-Gly-Pro-His-Lys-Val-Gly-Pro-Asn-Leu-His-Gly-Ile-Phe-Gly-Arg-His-Ser-Gly-Gln-Ala-Glu-Gly-Tyr-Ser-
    Ile Leu Glu Val Ala      Asn Leu      |Ser||Leu|      |Ser|Leu                                      Val Lys
        Gln                  Gln          |Asp|           |Trp|Thr
        Tyr                  Tyr          |Asn|                Ser
                             Trp                                Ala
                                                               |Pro|
```

```
                                        Lys
                                        Gln
                          Arg           Glu       Lys       Pro
                          Gln           Asn       Gln Pro Lys           Met
            Pro           Glu           Asp       Glu Tyr Gln           Tyr Lys
            Lys           Asn Met       Thr       Asn Lys Glu His       Phe Asp
            Glu           Asp Arg       Ile       Thr Gln Asp Asp       Arg Glx           Lys
            Asn       Arg Ser Asn Thr Gln Val     Ser Asn Ser Thr       His Ile           Glx
            Thr       Met Lys Ala Ser Ala Ile Ala Ala Asp Ala Ser Phe Asp Val         Glu Asx       Phe Met
    Ser Ala Gly Asx Ala Gly Ala Gly Ala Gly     Gly Ala Gly Val Leu Leu Ala Phe     Leu Asp Ser His Val
    55              60                  65               70               75          80
Tyr-Thr-Asp-Ala-Asn-Ile-Lys-Lys-Asn-Val-Leu-Trp-Asp-Glu-Asn-Asn-Met-Ser-Glu-Tyr-Leu-Thr-Asn-Pro-Lys-Lys-Tyr-Ile-Pro-
Lys                 Ile Ser Asn      Thr Δ Lys Leu Ala Leu Asp Ser     Leu|Asp||Ser|     Val Arg Glu Leu      Leu
                    Arg                   Phe     Gln                   Ser               Ile     Gln
                                          Tyr     Lys                   Gln               Ser     Lys
                                          |Gly|   Tyr                   Lys               Thr     Asn
                                          |Leu|                         Tyr              |Leu|
                                          |Ser|                         Trp
                                          |Cys|
```

```
                                                                      His
                                                  Gln                 Arg Gln       Δ
                                          Pro Glu             Lys      His Glu       Lys  Δ
                                          Lys Asn             Gln      Glx Asn Lys Δ Gln Lys
                                          Glx Asp             Glu      Glu Asp Glu Lys Asn Gln
                          Pro             Asn Asp Thr         Thr      Thr Thr Asn Asp Asp Glu
                Ser       Thr       Met   Asx Thr Ser Gln     Val          Lys Leu Ser Ser Thr Asn Ser Glu
                Ile       Val   Phe Pro Ser Ser Ala Glu       Ala His Ile Leu Ser Val Ile Leu Ser Ala Leu  Δ
    Lys Asn     Val       Ala   Ile Ser Ala Ala Gly Asx       Gly Asn Val Val Ala Phe Met Ala Ala Val Ala Gly Ala Glu
            85                90                  95               100              105              108
Gly-Thr-Lys-Met-Ala-Phe-Gly-Gly-Leu-Lys-Lys-Glu-Lys-Asp-Arg-Asn-Asp-Leu-Ile-Thr-Tyr-Leu-Lys-Lys-Ala-Cys-Glu
            |Ile|           Ala                Leu           Leu      |Ser|     Leu|Ser|
            |Arg|                              Lys           Gly               Gln
            |Thr|                              Tyr
```

279

```
 1              15             30             45
 |    1         |   5          |  10          |  15
ATG ACT GAA TTC AAG GCC GGT TCT GCT AAG AAA GGT GCT ACA CTT TTC AAG ACT
(Met)Thr-Glu-Phe-Lys-Ala-Gly-Ser-Ala-Lys-Lys-Gly-Ala-Thr-Leu-Phe-Lys-Thr-


     60             75             90            105            120
     |  20          |  25          |  30          |  35          |  40
AGA TGT CTA CAA TGC CAC ACC GTG GAA AAG GGT GGC CCA CAT AAG GTT GGT CCA AAC TTG CAT GGT ATC
Arg-Cys-Leu-Gln-Cys-His-Thr-Val-Glu-Lys-Gly-Gly-Pro-His-Lys-Val-Gly-Pro-Asn-Leu-His-Gly-Ile-


         135            150            165            180
         |  45          |  50          |  55          |  60
TTT GGC AGA CAC TCT GGT CAA GCT GAA GGG TAT TCG TAC ACA GAT GCC AAT ATC AAG AAA AAC GTG TTG
Phe-Gly-Arg-His-Ser-Gly-Gln-Ala-Glu-Gly-Tyr-Ser-Tyr-Thr-Asp-Ala-Asn-Ile-Lys-Lys-Asn-Val-Leu-


 195            210            225            240            255
 |  65          |  70          |  75          |  80          |  85
TGG GAC GAA AAT AAC ATG TCA GAG TAC TTG ACT AAC CCA AAG AAA TAT ATT CCT GGT ACC AAG ATG GCC
Trp-Asp-Glu-Asn-Asn-Met-Ser-Glu-Tyr-Leu-Thr-Asn-Pro-Lys-Lys-Tyr-Ile-Pro-Gly-Thr-Lys-Met-Ala-


     270            285            300            315            330
     |  90          |  95          |  100         |  105      108 |
TTT GGT GGG TTG AAG AAG GAA AAA GAC AGA AAC GAC TTA ATT ACC TAC TTG AAA AAA GCC TGT GAG TAA
Phe-Gly-Gly-Leu-Lys-Lys-Glu-Lys-Asp-Arg-Asn-Asp-Leu-Ile-Thr-Tyr-Leu-Lys-Lys-Ala-Cys-Glu
```
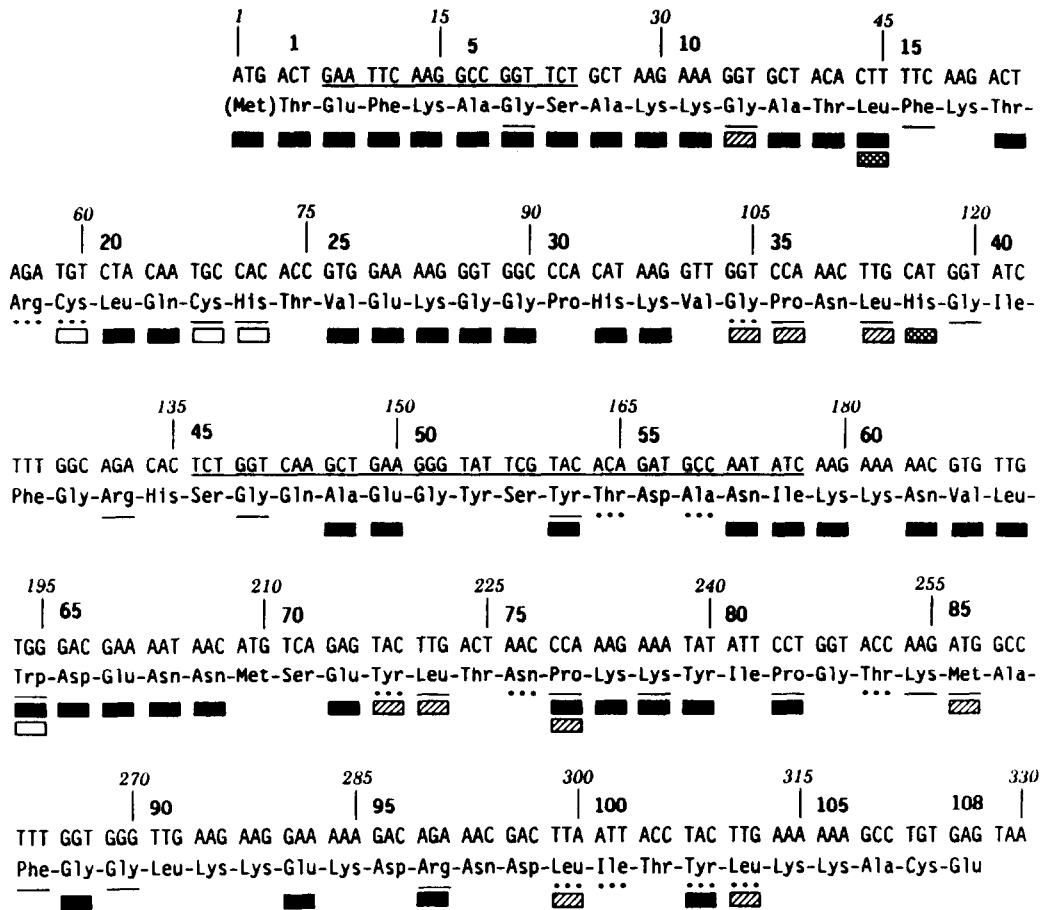
Fig.2. Correlation of functional and nonfunctional amino acid replacements in iso-1-cytochrome c to evolutionarily invariant, conserved and variant residue positions. The nucleotide sequence of the CYC1 structural gene and the corresponding amino acid sequence of iso-1-cytochrome c are shown with the numbering system beginning at, respectively, the ATG initiation codon and the N-terminal threonine. Evolutionarily invariant (———) and conserved (···) positions are indicated immediately below the appropriate residues. Filled boxes indicate positions in iso-1-cytochrome c where functional replacements have been found; open boxes indicate positions of replacements that cause loss of function and that do not allow heme attachment; single-hatched boxes indicate the position of replacements that cause loss of function, but still allow at least partial heme attachment; cross-hatched boxes at positions 14 and 38 indicate proline replacements that cause loss of function, but still allow partial heme attachment. The two lines over amino acids 2–7 and 45–58 indicate the deletions in the cyc1-453 mutant.

Early experiments designed to deduce the DNA sequence of the 5'-end of the CYC1 gene by determining the amino acid replacements in revertants of mutations mapping to this region of the gene exposed an extensive number of functional amino acid replacements in the N-terminus of the protein. In all, 85 amino acid replacements have been found within the first 9 amino acid residues, including iso-1-cytochromes c having both extended and truncated N-termini (fig.1). The functional replacements of Gly6, an evolutionarily invariant

residue, are especially noteworthy. Presumably Gly6 serves a specific function within the cytochrome c molecule for it to be maintained through evolution; however, this requirement apparently is minimal and is not understood. In contrast, Gly11 and Gly34 appear to be absolutely required, even though they may serve no function other than meeting a minimal structural requirement defined simply as the absence of an amino acid side chain.

As shown in fig.2, each of the amino acid

substitutions that abolishes the function of iso-1-cytochrome c, with the exception of the two residues replaced by prolines, occurs at evolutionarily invariant or highly conserved positions. Considering the large number of mutants, this observation suggests that those residue positions in cytochrome c that are not at least highly conserved through evolution are capable of substantial residue variation while maintaining proper folding and overall structural requirements so that function is retained.

5. FUNCTIONAL REPLACEMENTS AT CONSERVED RESIDUES

Our collection of amino acid replacements that abolish function of iso-1-cytochrome c, combined with the convenient selection procedure for generating functional revertants of cyc1 mutants, represent an excellent opportunity to investigate the precise structural requirements at each of these conserved residue positions. Ernst et al. used this approach to show that iso-1-cytochrome c is functional with valine, threonine, isoleucine and serine replacing the evolutionarily invariant Pro76 residue (fig.1) [18]. Moreover, the iso-1-cytochrome c containing the Val76 replacement was nearly indistinguishable from the normal protein with respect to structure in vitro [19] and function in vivo [18]. Although the local backbone conformation imposed by a proline residue at this position would be expected to be different from that imposed by valine, this structural difference is negligible with respect to the overall structure of iso-1-cytochrome c, at least as we are able to measure this parameter in the laboratory [19]. The replacements of Pro76 causing either loss of function or retention of at least partial function of iso-1-cytochrome c have been explained by a side chain size requirement [18] and by the requirement for residues that break α-helices [19].

Schweingruber et al. previously isolated mutants encoding functional amino acid replacements of the invariant Trp64 residue, including phenylalanine, tyrosine and leucine replacements [20]. Furthermore, second site revertants were isolated in which the mutation that abolished function, either Cys64 or Ser64, was retained, while function was restored by a Ser45 ⟶ Phe45 replacement. The tertiary structure of iso-1-cytochrome c reveals

that Trp64 and Ser45 are immediately proximal to each other [1]. Apparently the requirement served by the normal Trp64 is directly compensated for by the Phe45 replacement. Although the crystal structure shows hydrogen bonding between the Trp64 side chain and propionate moiety on tetrapyrrole ring A of the heme [1], this bonding cannot be critical for folding, stability, or function of iso-1-cytochrome c since this residue can be replaced by residues incapable of hydrogen bonding to the propionate. Trp64 is probably required to contribute to the hydrophobic lining of the heme crevice, whereas hydrogen bonding to propionate apparently is dispensable. In fact, NMR data from the solution structure of tuna cytochrome c suggest that the Trp64-propionate hydrogen bond may not exist in solution, but occurs as a result of crystal packing [21].

Other functional replacements at evolutionarily invariant or conserved residue positions also were uncovered by selection on lactate medium. Lysine was found to functionally replace Tyr53; glutamate to replace Lys78; leucine to replace Pro81; leucine and glycine to replace Arg96; and leucine to replace Tyr102. Furthermore, recent site-directed mutagenesis experiments of the CYC1 gene indicate that additional highly conserved residues can be functionally replaced. Serine, tyrosine, and glycine were shown to functionally replace the invariant Phe87 in vitro [22], and arginine can functionally replace the conserved Lys77 residue both in vitro and in vivo [23]. The Arg77 replacement is especially interesting since Lys77 is trimethylated in fungal and plant species and has been suggested to be involved in both mitochondrial import and association of iso-1-cytochrome c with the cytochrome c oxidase complex. The Arg77 replacement clearly shows not only that trimethylation of Lys77 is not required for processing or function of iso-1-cytochrome c, but also that the structural requirement for this highly conserved lysine residue is not absolute. (Only serine is found in addition to lysine at this position in the phylogenetic series (fig.1).)

The recognition that certain amino acids can functionally replace evolutionarily invariant residues, sometimes at levels indistinguishable from wild type, indicates that evolutionary invariance does not necessarily imply functional invariance. These residues may be conserved through

evolution because cytochromes c containing them confer a particular advantage to the organism in the wild that we are unable to recognize in the laboratory. Other explanations are also conceivable. For example, it may be that even a slight difference in function conferred by the normal residue is strongly selected for over the course of the evolutionary time scale. Another interesting possibility is that cytochrome c may have evolved to minimize the effect of point mutations on function, as was suggested to explain functional replacements of conserved residues in the α-interferon gene family [24].

6. DELETION OF A 'LOOP' STRUCTURE

Sherman et al. isolated approximately 100 deletions encompassing various segments of the CYC1 gene by a technique that relied on the sporulation of certain diploid strains and on the selection for loss of cytochrome c function on chlorolactate medium [11]. Apparently cyc1 deletions arise after the meiotic interaction of certain CYC1 alleles. All but one were completely deficient in iso-1-cytochrome c. The exception, designated cyc1-453, contained a nonfunctional iso-1-cytochrome c at approx. 20% of the normal level. This mutant arose after sporulation of a diploid strain containing the CYC1-31-K and CYC1-183-AD alleles; CYC1-31-K and CYC1-183-AD encode functional iso-1-cytochromes c with deletions encompassing residues Glu2 through Ser7 and Ala8 through Ala12, respectively [11]. DNA sequence analysis of cyc1-453 established that this allele encodes the Glu2 through Ser7 deletion of CYC1-31-K, and encodes a second deletion corresponding to residues Ser45 through Ile58; loss of iso-1-cytochrome c function is attributed to Ser45-Ile58 deletion since the Glu2-Ser7 deletion does not abolish function.

The cyc1-453 mutation is especially interesting for two reasons. Firstly, this deletion is within the segment of cytochrome c that is absent in the 'S' type bacterial cytochromes c [25]. For example, cytochromes c_{551} are lacking amino acids corresponding to iso-1-cytochrome c residues 42 through 64. Secondly, the sequence from residues 40–54 in tuna cytochrome c (corresponding to residues 45–59 in iso-1-cytochrome c) constitutes a region of secondary structure defined as a 'loop' by Leszczynski and Rose [26]. Although the

iso-1-cytochrome c from the cyc1-453 mutant is nonfunctional, it retains sufficient structural integrity such that a substantial level of the holoprotein is formed, suggesting that the Ser45-Ile58 region is not required for heme attachment or mitochondrial import.

7. CONCLUSION

We have described the use of genetic selection procedures to isolate an extensive series of altered forms of yeast iso-1-cytochrome c. The amino acid replacements in each of these proteins were defined as an approach to determining the role of specific residues for the maturation and function of iso-1-cytochrome c. Interpretation of these data was facilitated by the recently available crystal structure of iso-1-cytochrome c, as well as the extensive phylogenetic series of cytochromes c. We are optimistic that the wealth of cytochrome c structural information summarized here and elsewhere [27], along with available techniques for continued investigation, will soon lead to a detailed understanding of the structural and functional requirements of cytochrome c.

REFERENCES

[1] Louie, G.V., Hutcheon, W.L.B. and Brayer, G.D. (1988) J. Mol. Biol. 199, 295–314.
[2] Keilin, D. (1930) Proc. R. Soc. London B106, 418–444.
[3] Hampsey, D.M., Das, G. and Sherman, F. (1986) J. Biol. Chem. 261, 3259–3271.
[4] Dickerson, R.E. and Timkovich, R. (1975) in: The Enzymes (Boyer, P.D. ed.) vol.11, pp.397–547, Academic Press, New York.
[5] Takano, T. and Dickerson, R.E. (1981) J. Mol. Biol. 153, 79–94.
[6] Takano, T. and Dickerson, R.E. (1981) J. Mol. Biol. 153, 95–115.
[7] Ashida, T., Tanaka, N., Yamane, T., Tsukihara, T. and Kakudo, M. (1973) J. Biochem. (Tokyo) 73, 463–465.

[8] Matsuura, Y., Hata, Y., Yamaguchi, T., Tanaka, N. and Kakudo, M. (1979) J. Biochem. (Tokyo) 85, 729–737.

[9] Ochi, H., Hata, Y., Tanaka, N. and Kakudo, M. (1983) J. Mol. Biol. 166, 407–418.

[10] Sherman, F., Stewart, J.W., Jackson, M., Gilmore, R.A. and Parker, J.H. (1974) Genetics 77, 255–284.

[11] Sherman, F., Jackson, M., Liebman, S.W., Schweingruber, A.M. and Stewart, J.W. (1975) Genetics 81, 51–73.

[12] Huibregtse, J.M. and Engelke, D.R. (1986) Gene 44, 151–158.

[13] Zoller, M.J. and Smith, M. (1984) DNA 3, 479–488.

[14] Kunkel, T.A., Roberts, J.D. and Zakour, R.A. (1987) Methods Enzymol. 154, 367–382.

[15] Moerschell, R.P., Tsunasawa, S. and Sherman, F. (1988) Proc. Natl. Acad. Sci. USA 85, 524–528.

[16] Hennig, B., Koehler, H. and Neupert, W. (1983) Proc. Natl. Acad. Sci. USA 80, 4963–4967.

[17] Dumont, M.E., Ernst, J.F., Hampsey, D.M. and Sherman, F. (1987) EMBO J. 6, 235–241.

[18] Ernst, J.F., Hampsey, D.M., Stewart, J.W., Rackovsky, S., Goldstein, D. and Sherman, F. (1985) J. Biol. Chem. 260, 13225–13236.

[19] Ramdas, L., Sherman, F. and Nall, B.T. (1986) Biochemistry 25, 6952–6958.

[20] Schweingruber, M.E., Stewart, J.W. and Sherman, F. (1979) J. Biol. Chem. 254, 4132–4143.

[21] Williams, G., Clayden, N.J., Moore, G.R. and Williams, R.J.P. (1985) J. Mol. Biol. 183, 447–460.

[22] Pielak, G.J., Mauk, A.G. and Smith, M. (1985) Nature 313, 152–154.

[23] Holzschu, D., Principio, L., Conklin, K.T., Hickey, D.R., Short, J., Rao, R., McLendon, G. and Sherman, F. (1987) J. Biol. Chem. 262, 7125–7131.

[24] Valenzuela, D., Weber, H. and Weissmann, C. (1985) Nature 313, 698–700.

[25] Dickerson, R.E. (1981) Sci. Am. 244, 137–153.

[26] Leszczynski, J.F. and Rose, G.D. (1986) Science 234, 849–855.

[27] Pielak, G.J., Concar, D.W., Moore, G.R. and Williams, R.J.P. (1987) Prot. Eng. 1, 83–88.