Introduction

Cytochrome c oxidase (COX) is the terminal, and, possibly, the rate-limiting (Erecinska and Wilson 1982; Poyton et al. 1988; Poyton and McEwen 1996) enzyme complex in the mitochondrial electron transport chain. These complexes have been designated I–V (Hatfield 1985). All but complex II have components encoded by both mitochondrial and nuclear DNA. Mammalian COX (complex IV) is a dimeric enzyme; each monomer consists of 13 subunits, 3 of which are encoded in the mitochondrial genome and perform the known catalytic functions of COX. Relatively little is known of the role of the nuclear-encoded subunits, although they have been proposed to function in the assembly and regulation of the holoenzyme (Kadenbach, Kuhn–Nentwig, and Buge 1987; Capaldi 1990; Poyton and McEwen 1996).

Among the nuclear-DNA-encoded COX subunits, VIa is one of three that exhibit tissue-specific isoforms in mammalian species (Smith et al. 1991; Fabrizi et al. 1992; Mell, Seibel, and Kadenbach 1994; Wan and Moody 1995). The COX VIa heart isoform (COX VIa-H) is found only in contractile muscle, whereas the liver isoform (COX VIa-L) is present in most tissues. Furthermore, the tissue specificity follows a developmental program in heart and muscle to yield prenatal expression of predominantly COX6AL, followed by a postnatal transition to predominantly COX6AH expression (Ewart, Zhang, and Capaldi 1991; Taanman et al. 1992; Bonne et al. 1993; Grossman et al. 1995; Parsons et al. 1996). In addition, isoforms of COX are suspected of being deficient in some diseases that exhibit a tissue-specific reduction in COX activity (DiMauro et al. 1990; Kennaway et al. 1990; Van Beeumen et al. 1990).

One potential isoform function is suggested by studies with bovine COX, which implicate the heart isoform in governing the coupling between oxidation and phosphorylation (Anthony, Reimann, and Kadenbach 1993; Rohdich and Kadenbach 1993; Kadenbach et al. 1995; Frank and Kadenbach 1996). The regulation of the efficiency of energy transduction by ADP was proposed to have the physiological role of modulating cardiac heat production at different workloads (Anthony, Reimann, and Kadenbach 1993). This is based on the stimulation of uncoupled respiration of reconstituted heart but not liver enzyme by ADP (Rohdich and Kadenbach 1993) and the stimulation of the respiratory control ratio of the heart enzyme (Rohdich and Kadenbach 1993). The proposed interaction of bovine COX VIa-H with ADP has recently been confirmed by protein crystallography (Tsukihara et al. 1996), which identified a matrix cholate (ADP) site that is salt-bridged with the phosphates of Arg14 and Arg17, and also hydrogen-bond ed to the main chain of Phe21.

Crystallography of bovine COX has also identified a 10-amino-acid region at the N-terminus of COX VIa-H that interacts with COX III of the opposite monomer of the dimeric COX enzyme and may play an important role in stabilization of the dimer (Tsukihara et al. 1996). We show that this 10-residue region is highly divergent in paralogous comparisons, and thus identify a second or alternative isoform function.

An acceleration of evolutionary rate after gene duplication has been proposed as a mechanism to acquire new function in globin evolution (Goodman 1981) and calmodulin evolution (Baba et al. 1984). To see if a similar pattern occurred in COX6A evolution, we examine evolutionary rates for the heart (H) and liver (L) isoform genes using a more extensive data set than previously available (Saccone, Pesole, and Kadenbach...
Materials and Methods

Isolation of the Mouse Cox6ah Gene

A partial cDNA of the mouse Cox6ah gene (Grossman et al. 1995) was used to screen 10 million phages from a 129/Sv mouse genomic library (a gift from Dr. A. Petersen) by standard methods. Bacteriophage DNA was prepared from three positive clones (Donovan, Lu, and Nagarajan 1993). Restriction maps of the cloned inserts were generated with selected restriction enzymes. A unique 5.5-kb Sac I fragment from ACOX6AH-11 that hybridized to the cDNA was subcloned into Sac I-digested pKS(-) vector (Stratagene). Deletion subclones and subclones of individual restriction fragments were subjected to bidirectional sequencing with an Applied Biosystems, Inc. model 373A automated DNA sequencer (Center for Molecular Medicine and Genetics, Wayne State University).

Alignment of Gene and Protein Sequence

A total of eight cDNA or nuclear DNA sequences of COX VIa proteins are now available for both isoforms from human, mouse, rat, and cow (Fabrizi et al. 1992; Smith et al. 1991; Mell, Seibel, and Kadenbach 1994; Grossman et al. 1995; Wan and Moreadith 1995). These sequences were aligned for phylogenetic analysis at both the nucleotide and the amino acid levels. The analyses were restricted to the coding sequences of three exons and two small introns (186 and 76 bp, respectively). Evolutionary analysis of the aligned sequences was carried out using Phylogenetic Analysis Using Parsimony (PAUP) version 3.1.1 (Swofford 1993). These analyses used unordered and unweighted character states and utilized the branch-and-bound search algorithm. In order to estimate the relative strengths of nodes within the phylogeny, bootstrap replicates (n = 500) and Bremer support indices were calculated (Felsenstein 1985; Bremer 1988; Hedges 1992; Kallersjo et al. 1992). Phylogenetic analyses were rooted with the yeast COX6A gene sequence. Minimum mutational distances were calculated from the aligned amino acid sequences (Fitch and Margoliash 1967) to estimate the probability of a common evolutionary origin of the yeast and mammalian COX VIa sequences with the method of Moore and Goodmam (1977). It does so by estimating whether the minimum mutational distance between two amino acid sequences is less than would be expected for two unrelated sequences.

Maximum-parsimony analysis of the aligned sequences was carried out using Phylogenetic Analysis Using Parsimony (PAUP) version 3.1.1 (Swofford 1993). These analyses used unordered and unweighted character states and utilized the branch-and-bound search algorithm. In order to estimate the relative strengths of nodes within the phylogeny, bootstrap replicates (n = 500) and Bremer support indices were calculated (Felsenstein 1985; Bremer 1988; Hedges 1992; Kallersjo et al. 1992). Phylogenetic analyses were rooted with the yeast COX6A gene sequence. Minimum mutational distances were calculated from the aligned amino acid sequences (Fitch and Margoliash 1967) to estimate the probability of a common evolutionary origin of the yeast and mammalian COX VIa sequences with the method of Moore and Goodmam (1977). It does so by estimating whether the minimum mutational distance between two amino acid sequences is less than would be expected for two unrelated sequences.

Results and Discussion

Isolation and Characterization of the Mouse Cox6ah Gene

The mouse Cox6ah gene spans 689 bp, consisting of three exons and two small introns (186 and 76 bp,
Alignment of Gene and Protein Sequences

The complete amino acid sequences of the mature COX6A protein (fig. 1) and the nucleotide sequences (not shown) for COX6A genes from mammals and yeast were aligned. Comparisons between the H and L isoforms exhibited previously identified patterns of divergence (Saccone, Pesole, and Kadenbach 1991). The N-terminal regions of the two isoforms are divergent and difficult to align at both the nucleotide and the amino acid levels. In contrast, the cytosolic region is highly conserved. Similarity of the yeast COX VIa protein with mammalian COX VIa isoforms is limited to a 7-residue section of the transmembrane region and a 22-residue portion of the intermembrane region. These regions are among those conserved in comparisons between isoforms of mammalian COX6A.

Maximum-Parsimony Analysis

Putative species phylogenies of COX VIa isoform sequences of the major mammalian lineages included in the analysis (rodent, bovine, and primate) were evaluated on the basis of parsimony. With either nucleotide or inferred amino acid sequences, the most parsimonious branching pattern first joined mouse and rat together, next joined cow to this rodent branch, and then joined human to the cow‐rodent branch (fig. 2). The phylogeny of the genes included in the analysis showed support at most nodes at greater than 80% of bootstrap replicates for either the nucleotide or amino acid sequences. Character state optimization methods (ACCTRAN and DELTRAN) did not differ to any great degree, primarily because consistency indices (a measure of homoplasy) were 0.82 and 0.94 for the nucleotide and amino acid analyses, respectively. Results presented here are from the DELTRAN optimization, which provided a more stringent test of hypotheses by placing more of the ambiguous character state changes along mammalian branches of the evolutionary tree. Minimum mutational distances between the yeast COX6A gene and mammalian COX6A isoform genes ranged from 69 to 74 over the 83-residue region (excluding indels), providing evidence that the yeast gene has a common evolutionary origin with the mammalian genes (P < 0.01).

Rate Variation in COX6A Isoform Evolution

To evaluate the overall rates of N and S nucleotide substitutions to the H and L isoforms of COX6A (fig. 3), sequence divergences were added together for the time since divergence of the species included in the analysis. Branches leading from the ancestral COX6A sequence to the species included in the analysis were not included to mitigate the potential randomizing effect of the highly divergent yeast sequence. Total N sequence divergences for the H and L isoforms were 15.70 ± 13.64% and 76.91 ± 5.15%, respectively. Exon 1 encodes the 5′ untranslated (UTR) region, presequence, and the hydrophilic N-terminal domain of the mature protein. Exon 2 encodes 46 amino acids, including the transmembrane domain. Exon 3 encodes 27 amino acids, including the hydrophilic C-terminal plus 112 bp of 3′ UTR sequence present in the cDNA. While this work was in progress, a sequence of the mouse Cox6ah gene was presented by Wan and Moreadith (1995). Our sequence, which is identical in the region of overlap, includes an additional 413 bp in the 5′ region and is deposited in GenBank under accession number U63716.
These results indicate that the difference in N substitution rate is statistically significant ($t = 2.308, P < 0.05$), whereas that in the S substitution rate is not ($t = 0.173, P > 0.25$). These results indicate that the N substitution rate of the COX6AH gene was more rapid than that of the ubiquitously expressed COX6AL gene, consistent with the hypothesis that genes with tissue-specific or developmentally specific functions evolve more rapidly than do more ubiquitously expressed genes (Toniolo, Persico, and Alcalay 1988; Sugawara et al. 1990; McCarrey 1994; Hastings 1996). This also indicates that the higher N substitution rate of the H isoform gene cannot be accounted for by a higher background mutation rate and suggests that it could result from selective pressures.

Molecular evidence from other genes has suggested that primates and bovines are more closely related to each other than to rodents (Easteal 1990; Graur, Huide, and Li 1991; Graur et al. 1992). This topology was examined in addition to the most parsimonious tree to determine the potential effect on evolutionary rates of the COX Vila isoform genes. This topology is four steps longer in the amino acid analysis and five steps longer in the nucleotide sequence analysis. For this topology, total N sequence divergences for the H and L isoforms were $15.42 \pm 2.87\%$ and $8.56 \pm 2.09\%$, whereas respective S divergences were $96.88 \pm 13.42\%$ and $84.01 \pm 12.65\%$. The difference in the N substitution rate remains statistically significant ($t = 1.932, P < 0.05$), and the S substitution rate does not ($t = 0.698, P > 0.25$). These results indicate that the difference in N substitution rate for the two isoforms is independent of topology.

Intramolecular Substitution Heterogeneity of Functional Regions

Regions of COX Vila proteins were evaluated for amino acid sequence divergence among orthologous H or L isoforms and between paralogous genes of the two isoforms to examine whether divergence is related to function (fig. 4). Initially, amino acid sequence divergence was evaluated over the mature protein by using a sliding window of 20 residues. When the paralogous H and L genes are compared, sequence divergence is high at both the N-terminus (fig. 4, region A) and the transmembrane region (fig. 4, region B). Lower amino acid sequence divergences are located near the end of the transmembrane region and at the C-terminus (fig. 4, region C). The difference in divergence is from approximately 90% to 10%. Sequence divergence for orthologous and paralogous comparisons contrast most in the N-terminal region, where paralogous divergence is high (~90%) but orthologous divergence is low (~10%). The magnitude of the peak orthologous divergence is modest compared to the N-terminal peak paralogous divergence of ~90%.

Saccom, Pescoll, and Kadenbach (1991) suggested that positive natural selection on the N-terminal residues was involved in the acquisition of the tissue-specific function of COX Vila isoforms. To test this hypothesis, we divided COX6A evolution into two time frames: (I) after duplication but prior to the divergence of the eutherian mammals included here, and (II) after divergence of the eutherian mammals. We examined the N/S ratios in these time periods for each of the three functional regions of the gene (fig. 5). For the complete gene, N/S ratios for time periods I and II were $0.41 \pm 0.10$ and $0.13 \pm 0.02$, respectively. The difference in N/S is statistically significant ($t = 2.746, P < 0.05$). To evaluate whether any subregion of interest contributed particularly to the overall N/S difference, each subregion was examined individually. The N-terminal matrix (region A) N/S ratios for time periods I and II were $1.25 \pm 0.50$ and $0.12 \pm 0.07$, a difference that was also statistically significant ($t = 2.238, P < 0.05$). No other regions of the gene produced statistically significant differences between the two time frames. These results suggest that selective pressures on nucleotide changes...
to the region encoding the residues located in the mitochondrial matrix differed relative to other regions and time frames. One explanation for the disparity in substitution rates of the different domains is that a "burst" of nonsynonymous substitutions occurred in this region after duplication of COX6A genes. Subsequent acquisition of a novel (e.g., tissue-specific) gene function may have led to purifying selection that retarded the accumulation of N sequence divergence in the matrix domain. Although there is evidence of differing selective pressures on the N-terminal sequences, the hypothesis of positive natural selection is premature. The N/S of this region (1.25 ± 0.50) is not statistically greater than 1.0 (t = 0.572, P > 0.25). More divergent isoforms of COX6A (if they exist) are needed to evaluate the hypothesis of positive selection.

**Tissue-specific Function**

Stimulation of bovine heart COX activity by ADP has led to a proposal for ADP-modulated tissue-specific regulation of COX activity (Anthony, Reimann, and Kadenbach 1993; Rohdich and Kadenbach 1993; Kadenbach et al. 1995; Frank and Kadenbach 1996). Protein crystallography of bovine COX VIa-H thus far supports this hypothesis by showing the presence of a cholate molecule, whose electron density map agrees with ADI (Tsukihara et al. 1996). The ADP molecule interacted with Arg14, Arg17, and Phe21 of COX VIa-H. These amino acids are present in the taxa examined, except for the replacement in mouse and rat of Arg at position 14 with Asn. Whether or not rodent COX VIa-H binds ADP has yet to be determined. Interestingly, both Arg14 and Phe21 residues are also conserved in the L isoform, whereas Arg17 is replaced by Lys. Although COX6AL is not stimulated by ADP (Anthony, Reimann, and Kadenbach 1993; Rohdich and Kadenbach 1993; Kadenbach et al. 1995), crystallographic examination of COX VIa-L has not yet taken place.

The 10 residues at the N-terminus of the matrix COX VIa protein are unique in interacting with COX III of the opposite monomer of the dimeric COX holo enzyme, and could function in stabilization of COX dimer formation (Tsukihara et al. 1996). Interestingly, these residues (like all of the N-terminal region) are mostly conserved in comparisons of orthologous isoforms in the species examined (with only two substitutions in the H isoform and none in the L isoform). In contrast, the two paralogous COX6A genes are highly divergent, with only four residues in common. If stabilization of COX dimer formation is the function of these residues, then the lack of similarity between the two isoforms suggests that homodimer stabilization could be affected by tissue specificity. This suggests a second mechanism through which COX activity is regulated differentially in tissues.

**Age of Divergence of COX6A Isoform Genes**

Sacccone, Pesole, and Kadenbach (1991) estimated the time of COX6A gene duplication at 240 ± 90 MYA. This estimate assumed a constant rate of evolution for each of the two isoforms of COX VIa, since sequence for only one H and two L isoform genes were available at the time. Analysis of a broader database shows a more rapid rate of substitution in COX6AH than in COX6AL and extensive intramolecular heterogeneity in substitution rates. Since the origin of the two isoforms appears to predate the origin of mammals, an assumption about the constancy of the substitution rate also may not hold for nonmammalian taxa or ancestors. Obtaining a more reliable date for the origin of the isoforms depends on sampling of metatherian mammals and other nonmammalian groups, such as bony fishes, amphibians, reptiles, and birds. The suggestion that COX VIa-H functions in heat regulation is noteworthy in light of the best time estimate for gene duplication corresponding to the time when the evolution of homeothermy is thought to have occurred in mammalian ancestors (Crompton, Taylor and Jagger 1978; Hillenius 1994). Data on the presence of the COX6A gene duplication in nonplacental mammals and other vertebrate groups could help identify the importance of the COX6A gene duplication in the evolution of homeothermy in mammals.

**Acknowledgments**

Support for this work was provided by grants from NIH (GM48800 and GM48517). We thank N. Rosenthal.
for experimental help and Dr. G. Tromp for assistance with the statistical analyses.

LITERATURE CITED


KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park.


SACCONE, C., G. PESOLE, and B. KADENBACH. 1991. Evolutionary analysis of the nucleus-encoded subunits of mam-
Molecular Evolution of COX Vla Isoforms


Rodney L. Honeycutt, reviewing editor

Accepted February 12, 1997