# Evolutionary analysis of Quinone Reductases 1 and 2 suggest that NQO2 evolved to function as a pseudoenzyme

**Faiza Islam, Nicoletta Basilone, Eric Ball, and ‡Brian Shilton**

Department of Biochemistry, University of Western Ontario

‡To whom correspondence should be addressed: Brian H. Shilton, Department of Biochemistry, The University of Western Ontario, 1151 Richmond St., London Ontario Canada N6A 5C1. Tel.: 519-661-4124; Fax: 519-661-3175; Email bshilton@uwo.ca

## Abstract

Quinone reductases 1 and 2 (NQO1 and NQO2) are paralogous FAD-linked enzymes found in all amniotes. NQO1 and NQO2 have similar structures and can both catalyze reduction of quinones and other electrophiles. The two enzymes differ in their cosubstrate specificity, with NQO1 using cellular redox couples NAD(H) and NADP(H), while NQO2 is almost completely inactive with these cosubstrates, and instead uses dihydronicotinamide riboside (NRH) and small synthetic cofactors such as *N*-benzyl-dihydronicotinamide (BNAH). Ancient sequence reconstruction was used to investigate the catalytic properties of a predicted common ancestor and 2 additional ancestors from each of the evolutionary pathways to extant NQO1 and NQO2. We found that in all cases, the small nicotinamide cofactors NRH and BNAH were good cosubstrates for the common ancestor and the enzymes along the NQO1 and NQO2 lineages. In the case of NADH, however, extant NQO1 evolved a catalytic efficiency 100x higher than the common ancestor, while NQO2 evolved a catalytic efficiency 1000x lower than the common ancestor. These results suggest a selective pressure for evolution of NQO1 towards greater efficiency with NADH, and for NQO2 towards extremely low efficiency with NADH. These divergent trajectories have implications for the cellular functions of both enzymes, but particularly for NQO2 whose cellular functions are only beginning to be uncovered.

# Introduction

Quinone reductases comprise an ancient family of flavoenzymes that afford the two-electron reduction of quinones and other electrophiles as a detoxification strategy. These enzymes use a catalytic FAD or FMN to receive a hydride from reducing cosubstrates NAD(P)H and then transfer the hydride to the quinone substrate or another electrophile (Deller, Macheroux, and Sollner 2008; Vasiliou, Ross, and Nebert 2006). Humans have two paralogous quinone reductase genes: Quinone Reductase 1, *nqo1* and Quinone Reductase 2, *nqo2* (Jaiswal et al. 1990; Zhao et al. 1997). The two proteins share 47% sequence identity and a strikingly similar overall structure. NQO1 has numerous roles and its regulation as part of the Keap1-Nrf2 pathway provide a clear function in the cellular response to oxidative stress (Lind, Hochstein, and Ernster 1982; Sihvola and Levonen 2017; Ross and Siegel 2021).

The cellular functions of NQO2 are still somewhat of a mystery (S Liao, Dulaney, and Williams-Ashman 1962; Vella et al. 2005), but there has been intriguing progress in recent years. In the brain, NQO2 was found to be overexpressed in a model system for age-related memory impairment and a pharmacological model of amnesia (Brouillette et al. 2007), and NQO2-knockout mice demonstrate enhanced learning (Benoit et al. 2010). Roles for NQO2 in learning and memory, along with insight into the cellular mechanisms, continue to be investigated by Rosenblum and co-workers (Rappaport et al. 2015; Gould et al. 2020; 2021). Our own interest in NQO2 began with the observation that it was an off-target interactor with inhibitors of the kinase CK2 (Duncan et al. 2008; Kevin K. K. Leung and Shilton 2015). At the time, NQO2 was known

to be an off-target interactor with two other kinase-targeted therapeutics, imatinib, targeted to the Abl kinase, and bisindolylmaleimide inhibitors targeted towards PKC (Bantscheff et al. 2007; Rix et al. 2007; Winger et al. 2009; Brehmer et al. 2004). The broad extent of NQO2 off-target interactions with kinase inhibitors was demonstrated in a comprehensive screen of over 200 clinically used kinase inhibitors, where NQO2 was found as an off-target interactor with over 30, by far the most frequent non-kinase interactor discovered (Klaeger et al. 2017). This finding suggests that NQO2 may play a role in the cellular effects of these drugs, such that interaction with NQO2 is selected for during drug screening and optimization. It is notable that unlike NQO2, NQO1 has not been picked up in off-target screens for kinase-targeted drugs. NQO2 has also been found to interact with antimalarials (Graves et al. 2002; Kwiek, Haystead, and Rudolph 2004). Until the cellular functions of NQO2 are understood in greater detail, it is difficult to know how its interaction with these various drugs might affect cell metabolism, signalling, and viability.

NQO2 is unusual because it cannot efficiently use the cellular redox couples NAD(H) and NADP(H) (Shutsung Liao and Williams-Ashman 1961; S Liao, Dulaney, and Williams-Ashman 1962; Chen, Wu, and Knox 2000; Islam et al. 2022). Intriguingly, NQO2 can efficiently use various smaller dihydronicotinamide cosubstrates such as dihydronicotinamide riboside (NRH) and *N*-propyl-dihydronicotinamide (Shutsung Liao and Williams-Ashman 1961; S Liao, Dulaney, and Williams-Ashman 1962), *N*-benzyl-dihydronicotinamide (BNAH), *N*-methyl-dihydronicotinamide and others (Knox et al. 2000) as sources of electrons. Of these smaller reducing cosubstrates, only NRH, and its oxidized form, nicotinamide riboside (NR) function in cells where they contribute to salvage pathways to regenerate $NAD^+$ (Bieganowski and Brenner

2004; Yang et al. 2019; 2020; Giroud-Gerbetant et al. 2019). There is no currently no evidence that NR and NRH could support enzymatic activity of NQO2 by functioning as a cellular redox couple similar to NAD(H) and NADP(H).

Why would an enzyme evolve such an unconventional cosubstrate specificity? The cosubstrate specificity of human NQO2 is conserved in other mammals (S Liao, Dulaney, and Williams-Ashman 1962; Zhao et al. 1997; Celli et al. 2006). More recently, the inability of NQO2 to efficiently use NAD(P)H was demonstrated for reptile (*Alligator mississipiensis*) and bird (*Anas platyrhynchos*) NQO2, indicating that this unusual property of NQO2 is conserved throughout the amniotes (Islam et al. 2022). A phylogenetic analysis of NQO1 and NQO2 sequences pointed to a gene duplication event in a vertebrate ancestor before the split of the ray-finned and lobe-finned fish, approximately 460 MYA (Islam et al. 2022; Hedges 2009). Gene duplication is a significant force for functional diversity in protein families (Ohno 1970; Zhang 2003). The most common fate of duplicated genes is the degeneration and silencing of one of the pair to a non-functioning pseudogene (Lynch and Conery 2000; McGrath et al. 2014). In less common cases, both copies of the gene could be maintained if they confer a survival benefit (Walsh 2003). In these cases, there are two possible fates for the duplicated genes: they could develop sub-functionalizations, where the original function is partitioned across both copies; alternatively there could be neo-functionalization, where one copy acquires a new function (Zhang 2003). The difference in cosubstrate specificity observed between NQO1 and NQO2 could indicate either sub- or neo-functionalization. In this regard, an important question is whether NQO2 evolved to use NRH efficiently, or whether it evolved specifically *not* to use NAD(P)H.

We have used ancestral sequence reconstruction to investigate the evolution of cosubstrate specificity in NQO1 and NQO2. Ancestral sequence reconstruction has been used to trace the molecular mechanisms of substrate or inhibitor specificity for many enzyme families (Merkl and Sterner 2016; Wilson et al. 2015; Hobbs et al. 2012; Harris et al. 2022). Following from these studies, we have reconstructed the predicted common ancestor of NQO1 and NQO2, along with two ancestors of each enzyme along their evolutionary paths from the common ancestor. The reconstructed NQO enzymes are catalytically active with the common quinone substrate, menadione. However, kinetic analyses show that after duplication from the common ancestor, NQO2 lost the ability to efficiently use NADH as a cosubstrate, whereas the catalytic efficiency of NQO1 improved compared to the common ancestor. In contrast to the consistent changes towards increased (NQO1) or decreased (NQO2) catalytic efficiency with NADH, all of the enzymes were able to used BNAH or NRH with moderate to high efficiency. This investigation into the evolution of the two enzymes indicates that NQO2 has evolved to avoid efficient reduction by NADH and NADPH rather than use NRH efficiently, and has provided insight into the molecular determinants of the difference in cosubstrate specificity between NQO1 and NQO2.

# Materials and Methods

## Reagents

Ampicillin, Trypsin, FAD, NADH, and NADPH were purchased from Sigma-Aldrich. 1-Benzyl-1,4-dihydro-3-pyridinecarboxamide (BNAH) was purchased from TCI Chemicals. Nicotinamide riboside (NR) chloride was purchased from Chromadex (Los Angeles, CA) as "Tru Niagen" nutritional supplements. The contents of each capsule (263 mg of nicotinamide riboside) were dissolved in 10 mL of 1.2 M $NaHCO_3$, pH 8.1, and the solution was filtered to remove insoluble microcrystalline cellulose. Following the basic protocol outlined in Zarei *et al.* (Zarei et al. 2021) dihydronicotinamideriboside (NRH) was prepared by degassing the NR solution, adding 2.7 equivalents of $Na_2S_2O_4$ and incubating under nitrogen for 3 hours. NRH was purified by reversed-phase HPLC using a semi-preparative Zorbax XD8-C18 column (9.4 x 250 mm; Agilent).

## Evolutionary Analysis

The evolutionary analysis consisted of multiple sequence alignments of NQO1 and NQO2 sequences that were found using human NQO1 and NQO2 protein sequences as bait to BLASTp against genomes in the NCBI database. Whether a given sequence is NQO1 or NQO2 was recognized by the annotation suggested by NCBI, by the percentage identity, and by the presence or absence of the 43-residue C-terminal domain of NQO1. A total of 172 sequences were collected. The multiple sequence alignment was calculated with MUSCLE and sequences could be aligned with minimal gaps or insertions. The next stage of the evolutionary analysis was ascertaining the evolutionary relationship between the protein sequences by constructing a phylogenetic tree. The

phylogenetic tree was constructed by maximum likelihood with the CAT-JTT evolutionary model. The phylogenetic tree for the protein sequences matched the species-level phylogenetic tree closely. The multiple sequence alignment and the phylogenetic tree construction were done using Phylobot. http://www.phylobot.com (25). Finally, the multiple sequence alignment and the phylogenetic tree were used to infer the amino acids of the sequences at the nodes of the phylogenetic tree. The predicted protein sequences from Phylobot at nodes of interest were selected for analysis. The resurrected sequence had the sequences with the highest probability for each position. The amino acid sequences were reverse-translated to yield the corresponding DNA sequence.

## Cloning of constructs and protein expression

Ancestral DNA sequences were ordered as genes from IDT (Integrated DNA Technologies, Iowa). The sequences were codon-optimized for expression in bacteria using IDT-provided tools. Two restriction sites were engineered into the sequences, KasI/Ehe1 at the N-terminal, and HindIII at the C-terminal. The ancestral genes were cloned into the pProEx-HTA (Invitrogen) plasmid in a way that fuses 6XHis tag followed by a TEV cleavage site upstream of the genes HHHHHHDYDIPTTENLYFQ/GA. Proteolytic cleavage with the TEV protease cuts out the 6XHis tag and leaves two extra GA residues on the N-terminal, unlikely to affect protein structure.

The chimeric sequences were constructed from CA2NQO2 and CA2NQO1 by Gibson Assembly according to previously published protocol (26). The coding sequences for CA2QR2 and CA2QR1 were PCR amplified and inserted into a modified pET vector, pEBC plasmid, via

Gibson Assembly. The modified pEBC plasmid contained a cellulose-binding module (CBM) in the N-terminal. BamHI and XhoI cut sites were used to open the plasmid vector. Primers for the target sequences were designed to incorporate the cut sites and some upstream sequences matching the vector allowing for overlapping ends during Gibson Assembly.

Cloning of the NQO2 mutants was also done with the same Gibson Assembly strategy as the chimeras. PCR primers were ordered to either make a short NQO2 mutant without the C-terminal tail or a mutant with the longer 43-residue C-terminal of CA2NQO1. The sequence of all constructs was verified with Sanger sequencing by the Robarts Research Institute.

Protein expression was done in *Escherichia coli* BL21-DE3 strains under ampicillin selection. For large-scale protein production, 4-litre cultures were inoculated with a single colony starter culture and grown overnight at 37 °C in Studier's ZYP-5052 auto-induction media (26). For screening the chimeric constructs, a colony was picked and grown overnight at 37 °C in 2.5 mL of Terrific Broth and protein expression was induced with 0.2 mM IPTG.

## Protein Purification

The ancestral proteins were purified by previously described methods (10, 28). Bacterial cell pellets were thawed, treated with DNase and lysozyme, and dispersed by the Dounce homogenizer and then the French press. The cell lysates were supplemented with 50 mM NaP, pH 7.4 and 1 M NaCl and 25 mM imidazole. The cells were ultracentrifuged at 4°C at 100,000 x g for 1 hour and 20 minutes to remove the cell membranes. After centrifugation, the pellet was discarded. The supernatant was flown through a $Ni^{2+}$ affinity (IMAC Sepharose Cytiva)

equilibrated with the same buffer. The 6XHis tagged protein was eluted with 250 mM imidazole in the same base buffer. After nickel column purification, the protein was dialyzed to remove the excess imidazole in 100 mM Tris-HCl, pH 8.2 and 0.5 M NaCl, 5 mM EDTA. Then, the protein was digested with the Tobacco Etch Virus protease, TEV, to completely cleave off the 6X His tag in 5 mM DTT added directly to the protein.

For anion-exchange chromatography, the proteins were dialyzed against 10 mM Tris, pH 8.4 and 1 mM EDTA, and 0.5 mM TCEP to prepare for anion exchange. The protein was applied to a 1.6 X 15-cm Q-Sepharose HP (Cytiva) anion exchange column equilibrated with 20 mM Tris and 1 mM EDTA at pH 8.4. They were eluted with a salt gradient from 0 to 500 mM at a 2 mL/min flow rate with a fraction of 8 mL. For cation-exchange chromatography, the protein was dialyzed against 20 mM NaP, pH 5.0 and 0.5 mM TCEP, and 1 mM EDTA. Then, it was applied to the S-Sepharose HP (Cytiva) cation exchange column equilibrated with 20 mM NaP, pH 5.0 and 1 mM EDTA, and 0.5 mM TCEP. The protein was eluted against a salt gradient (0 – 500 mM) at a 2 mL/min flow rate with a fraction of 8 mL. After ion-exchange chromatography, the protein was pooled and concentrated by ultrafiltration (Amicon) for gel-filtration chromatography. For FAD reconstitution, the concentrated protein samples were incubated with 1-3 M Guanidine-HCl, and 10 mM FAD, ZnCl, and DTT while sitting on an ice bucket. After gentle mixing, the protein samples were centrifuged to remove any precipitate and then applied to a 2.6 X 65-cm column of Superdex-200 preparatory grade resin (Cytiva), previously equilibrated with 50 mM Tris, pH 8.0 and 150 mM NaCl, 10 μM FAD, 10 μM $Zn^{2+}$. After gel filtration, NQO2 was dialyzed against 50

mM HEPES, pH 7.4 and 100 mM NaCl concentrated with Amico Ultra-centrifugal filter units, flash-frozen in liquid nitrogen, and stored at -80 °C.

Protein expression of the chimeric constructs was also done accordingly to previously described protocols (26). Overnight bacterial cultures were pelleted and resuspended in 10 mM Tris, pH 7.5. The cells were lysed by two rounds of freeze-thaw and a brief five-second sonication. The samples were centrifuged at 20,000 x g for 15 minutes to separate cell debris. The cell extract was added to a six-well plate and brought up to 2 mL in 10 mM Tris, pH 7.5 and 1 mM EDTA. For protein binding, Whatman #1 filter paper was used to hole-punch paper disks, and they were incubated overnight at 4°C on the extract. After protein binding, the disks were washed with the same buffer over a sintered glass-filter vacuum. To elute the protein off the disk, we added 100 µg/mL of TEV protease and incubated it at room temperature for 1 hour. The protein was collected from the disk by centrifugation.

## Thermal Shift Assay

To examine the stability of the constructs, the melting temperature ($T_m$) of the proteins was determined with the Applied Biosystems Real-Time PCR System using the QuantStudio™ 3 Real-Time PCR System. The thermal shift assay works by monitoring the change in fluorescence of the Sypro Orange Dye™, with an excitation maximum of 472 nm and an emission maximum of 570 nm. The PCR system thermally denatures proteins by changing the temperature in a controlled manner. The fluorescence emission from Sypro Orange™ dye increases when it binds to the exposed hydrophobic regions of unfolded proteins. Proteins at concentrations of 0.2 to 0.3 mg/mL were diluted into 100 mM HEPES, pH 7.4 and 200 mM NaCl, 5 mM DTT and incubated with 1X

Sypro Orange Dye™. Each protein was investigated in at least four replicates in a total volume of 20 mL in a 96-well plate. As negative controls, the fluorescence of dye only and buffer only were measured. As FAD excitation and emission overlap with Sypro Orange™, the change in fluorescence of FAD from the protein alone was measured as a control. There was no comparable signal from the FAD alone at the protein concentration, so it was likely not to affect the rest of the assay.

The 96-well plate was incubated on ice before starting the thermal denaturation assay. To generate the melt curves, the temperature was increased to 25°C at a rate of 1.6 °C/s and then held at 25°C for 2 minutes. Then, fluorescence data collection was started, and the temperature was increased to 99.0 °C at a rate of 0.05 °C/s. The temperature was then held at 99.0 °C for 2 minutes, after which the assay was finished. The Protein Thermal Shift™ Software v1.0 analyzed the melt curve data.

## Steady-State Kinetics

We used a previously described protocol to determine the enzymatic activity of all five ancestral proteins (7). For kinetic assay with BNAH and menadione, the stock BNAH and menadione solutions in methanol were prepared fresh every day and kept on ice during the duration of the kinetic experiments. The concentration of each reagent was determined spectroscopically using the extinction coefficient $\varepsilon_{355\ nm}$ of 7220 $M^{-1} \bullet cm^{-1}$ for BNAH and $\varepsilon_{333\ nm}$ of 2450 $M^{-1} \bullet cm^{-1}$ for menadione in the Cary 100 spectrophotometer. The kinetic reactions were monitored spectroscopically by directly measuring the oxidation rate of the reducing cosubstrate or the decrease in absorbance at 355 nm for BNAH. The final enzymatic reactions were carried out for 2

minutes in 1 mL volumes in 100 mM HEPES, pH 7.4 at 25 °C under constant stirring with a small magnetic stir bar. Enzymatic activity was measured given BNAH (500 μM - 5 μM) and menadione (100 μM - 2 μM) and the protein concentration varied from (0.5 μM – to 0.5 nM). Each reaction was carried out in triplicate or more.

For the kinetic assays with NADH, the same protocol was carried out. Stock NADH was dissolved in 100 mM Tris-HCl, pH 8.0, and prepared freshly before use. NADH (2 mM – 10 μM) oxidation was measured by monitoring the change in absorbance at a wavelength of 360 nm, $\varepsilon_{360\ nm}$ of 4147 $M^{-1} \bullet cm^{-1}$ or of 340 nm $\varepsilon_{340\ nm}$ of 6220 $M^{-1} \bullet cm^{-1}$. For the kinetic assays with NRH, NRH ( 2 mM – 10 μM) was dissolved in 100 mM Tris-HCl, pH 8.0. The reaction progress was monitored by measuring the change in absorbance at a wavelength of 360 nm, $\varepsilon_{360\ nm}$ of 4360 $M^{-1} \bullet cm^{-1}$ or at a wavelength of 380 nm, $\varepsilon_{380\ nm}$ of 1070 $M^{-1} \bullet cm^{-1}$. The initial rates were globally fit to a ping-pong bi-substrate Michaelis Menten equation (Equation 1, below) to obtain $K_M$ for both substrates and $K_I$ for menadione using Prism (www.graphpad.com).

$$rate = \frac{k_{cat}[NADH][menadione]}{K_{NADH}[menadione](1+\frac{[menadione]}{K_{I,menadione}})+K_{menadione}[NADH]+[NADH][menadione]}$$

# Results

## Conservation of NQO2 co-substrate specificity among amniotes

As part of the Keap1-Nrf pathway, NQO1 has clear roles in the reduction of quinones and other electrophiles in response to oxidative and xenobiotic stress. In contrast, the cellular functions of NQO2 are not well-characterized, and its activity as an enzyme for quinone reduction is restricted because, unlike NQO1 and its orthologues, NQO2 is unable to efficiently use the reducing co-substrates NADH and NADPH. NQO2 is able to use dihydronicotinamide riboside (NRH) and a variety of small synthetic dihydronicotinamide cofactors (S Liao, Dulaney, and Williams-Ashman 1962; Knox et al. 2000), but it appears unlikely that sufficient amounts NRH are present in cells to support catalysis. To better understand the relationship between NQO2 and its paralogue, NQO1, we have analyzed their evolution.

An initial evolutionary analysis of NQO1 and NQO2 showed that NQO1 and NQO2 orthologs were present in all tetrapod genomes, except for some amphibians (Islam et al. 2022). A phylogeny based on extant NQO1 and NQO2 sequences indicated a gene duplication event in an ancestral fish approximately 460 MYA, prior to the divergence between the ray-finned and lobe-finned fishes; an abbreviated phylogenetic tree is illustrated in Figure 1A with the full phylogenetic tree in the Supplemental Figure 1. Our interest in the co-substrate specificity of NQO2 prompted us to find whether this unusual characteristic was conserved beyond the mammals. To this end, analysis of NQO2 proteins from an extant bird (Mallard Duck, *A. platyrhynchos*) and reptile (American Alligator, *A. mississippiensis*) showed that the inability to use NAD(P)H efficiently

was conserved in these organisms (Islam et al. 2022). On this basis, we concluded that the co-substrate specificity of NQO2 was conserved among all amniotes. This observation raises two possibilities for natural selection and evolution of NQO2. The first is that efficient use of NRH was subject to natural selection, and during this process the ability of NQO2 to use NAD(P)H was lost. A second alternative is that NQO2 has evolved to be unable to use NAD(P)H efficiently; that is, NQO2 may gained a non-enzymatic cellular function that is incompatible with an ability to accept electrons from NADH and NADPH. Deciding between these two alternatives will help to elucidate the cellular functions of NQO2.

## Resurrection of ancestral NQO1 and NQO2 proteins

To decide between these alternative scenarios for the natural selection of NQO2, we have extended the evolutionary analysis of NQO1 and NQO2 by reconstructing their common ancestor and two additional enzymes along each of the evolutionary pathways to extant NQO1 and NQO2. The ancestral sequence reconstruction method relies on the multiple sequence alignment of extant related protein sequences and their phylogenetic relationship to reconstruct the ancestral gene at the nodes of the phylogenetic tree. Ancestral NQO1 and NQO2 sequences were predicted using 172 extant NQO1 and NQO2 sequences and a combined optimization of the sequence alignment and phylogeny by maximum likelihood using PhyloBot (Figure 1). This phylogenetic tree closely matched the accepted species-level tree of life (Hedges 2009). The NQO1 and NQO2 sequences were in separate branches, as expected from the gene duplication results. The phylogenetic relationship of NQO sequences is consistent with a previous analysis of quinone reductase

sequences, with NQO1 and NQO2 arising from a gene duplication and forming separate branches of the tree (Vasiliou, Ross, and Nebert 2006).

The multiple sequence alignment and the phylogenetic relationships were then used to infer 171 ancestral sequences for the nodes of the phylogenetic tree. To characterize the changes in NQO1 and NQO2 from a common ancestor to the extant proteins, 5 predicted sequences were chosen for reconstruction (Figure 1A). Three of the sequences included the common ancestor (CA) and the two nodes (CA1NQO1 and CA1NQO2) immediately following the duplication. An additional two nodes were chosen at the top of the phylogenetic trees for the amniotes (CA2NQO1 and CA2NQO2. For a given reconstructed ancestral sequence, each amino acyl residue had a particular posterior probability that gave the likelihood of the residue at that position; these results are summarized in Figure 1C for the 5 sequences we have reconstructed, and the complete data are provided in the supplemental material. The resulting 5 sequences were reliably predicted, with 60% of the residues in the common ancestor (CA) having a probability of 0.95 to 1.0, rising to 90% for the second common ancestor of NQO2 ("CA2NQO2", Figure 1C). It is notable that the only gap in the alignment of the NQO2 sequences occurs in the common ancestor.

## Purification and Characterization of the Ancient Proteins

The predicted protein sequences were reverse-translated to their corresponding nucleotide sequences for expression in *Escherichia coli*. Protein expression and purification followed our established method for extant NQO2 sequences (Kevin Ka Ki Leung, Litchfield, and Shilton 2012). All of the ancestral sequences were efficiently expressed in *E. coli* and purified to homogeneity using Ni-NTA affinity chromatography, removal of the affinity tag, and then ion-

exchange and gel filtration chromatography. The last step included reconstitution with FAD since our experience indicated that NQO2 produced in *E. coli* contained sub-stoichiometric amounts of FAD (Kevin Ka Ki Leung, Litchfield, and Shilton 2012). The final protein preparations were pure, as visualized by SDS-PAGE (Figure 3A). UV/Vis absorbance spectra of the ancestral proteins indicated maxima at 260 nm - 274 nm and two local maxima at wavelengths of 455 nm and 380 nm for the isoalloxazine ring (Figure 4).

To assess the stability of the ancestral enzyme constructs, we determined their melting temperature by thermal shiftassay , which measures temperature-dependent protein unfolding using a fluorescent dye. The melting temperatures were derived from the fluorescence derivative plot and non-linear fitting to Boltzmann sigmoidal curve. The melting temperature from the Boltzmann equation and the derivative of the fluorescence plot were closely correlated. The mean $T_M$ (n = 3) from the fluorescence derivative plot is summarized in the table below. Generally, the ancestral sequences were more thermostable by at least 10 C° than the corresponding extant proteins. The increased thermostability of the ancient constructs can be attributed to using the consensus sequences (30). Proteins in the NQO2 branch had a higher melting temperature than the proteins in the NQO1 branch (Table 1).

Interestingly, we noted some proteolysis of the C-terminal tail of CA after purification. The sequence of the C-terminal tail for CA had some ambiguously predicted amino acids. We hypothesized that the C-terminal tail of CA did not fold stably as it was ambiguously predicted in the ancestral state reconstructions. It was probably subject to proteolysis by the trace amounts of active proteases in the purified protein preparations. We confirmed this by conducting limited

proteolysis experiments with different concentrations of trypsin to find that the C-terminal tail was very sensitive to proteolysis by proteases (Figure 3B). Also, alternative amino acid predictions of the same ancestor did not include the C-terminal. Ambiguously positioned residues in the C-terminal tail could lead to improper folding and greater disorder, probably making this region prone to proteolytic cleavage. For CA1NQO1, it was clear again that the tail was subjected to some proteolysis although posterior probabilities were higher for this sequence.

**Table 1. Thermal stability of ancient and extant quinone reductases.**

| Protein | $T_m$ (C°) | 95% CI |
|---|---|---|
| CA | 68.3 | (67.7 – 68.9) |
| CA1NQO1 | 69.1 | (64.7 – 73.5) |
| CA2NQO1 | 69.2 | (68.7 – 69.7) |
| NQO1 | 58.1 | (56.0 – 60.2) |
| CA1NQO2 | 78.9 | (78.4 – 79.3) |
| CA2NQO2 | 74.7 | (74.0 – 75.5) |
| NQO2 | 64.2 | (62.7 – 65.7) |

## Evolution the catalytic efficiency of NQO1 and NQO2 with NADH

To trace the molecular evolution of NQO2 cosubstrate specificity, we conducted steady-state kinetic experiments of the ancestral proteins with different reducing cosubstrates. We previously reported that extant NQO2 catalytic efficiency with conventional reducing cosubstrates, NAD(P)H was about a million-fold lower than NQO1. NQO2 was incredibly inefficient with NADH, the catalytic efficiency of 3.0 ± 1.8 $M^{-1}$•$s^{-1}$, low $k_{cat}$ 0.0792 ± 0.018 $s^{-1}$ and a high $K_M$ of 26800 ± 10000 μM (Islam *et al.* 2022). In contrast, human NQO1 could use NADH with a catalytic efficiency of $8.3 \cdot 10^5$ $M^{-1}$ • $s^{-1}$, a high $k_{cat}$ of 240 ± 12 $s^{-1}$, $K_M$ of 290 ± 48 μM in its reaction with menadione.

The kinetic investigation of the ancestral proteins highlighted evidence for functional divergence following gene duplication. We found that NQO1 and NQO2 lineage evolved very different substrate specificities.

The most ancient protein, CA had a low turnover of NADH at 5.95 $s^{-1}$ with a 95% confidence interval at (4.50-8.64) $s^{-1}$, the catalytic efficiency of 9554 $M^{-1}$• $s^{-1}$. Then, going down the NQO1 lineage, the catalytic efficiency with NADH increased 1000-fold up to CA2NQO1. In contrast, the catalytic efficiency of the NQO2 lineage with NADH decreased from CA 10-fold and it remained low down the NQO2 lineage. In contrast, the CA1NQO2 was 200-fold less efficient with NADH and got worse down the line. The inability to use NADH by NQO2 enzymes was reflected by its high $K_M$ in the millimolar range and low turnover. Extant NQO1 can use both NADH and NADPH almost equally efficiently. We also tested the ability of the ancestral proteins to use NADPH as a reducing cosubstrate to find that CA and NQO1 ancestors could use NADPH equally efficiently.

At the same time, NQO2 enzymes lost the ability to use NADPH. Interestingly, all the ancestral proteins were conserved in their ability to reduce menadione, and the $K_M$ for menadione was consistent at ~20 μM throughout the ancestors. However, menadione was found to be a substrate inhibitor for the reducing cosubstrate, and the $K_I$ for menadione varied between the five ancestors.

**Table 2. Kinetics parameters of quinone reductases with NADH**

| Protein | $k_{cat, NADH}$ ($s^{-1}$)[a] | $K_{M, NADH}$ (μM)[a] | $k_{cat}/K_{M, NADH}$ ($M^{-1} \cdot s^{-1}$) |
|---|---|---|---|
| CA | 5.95 (4.50-8.64) | 623 (414 – 1007) | 9554 |
| CA1NQO1 | 59.8 (45.4 – 84.4) | 811 (493 – 1435) | $7.4 \cdot 10^4$ |
| CA2NQO1 | 1903 (1470 – 2664) | 608 (387 – 942) | $3.1 \cdot 10^6$ |
| CA1NQO2 | 0.156 (0.109 – 0.286) | 3481[b] (2096 – 7839) | 44.9 |
| CA2NQO2 | 0.299 (0.215 – 0.482) | 5785[b] (3421 – 11302) | 51.6 |
| NQO2[c] | 0.0792± 0.018 | 26800± 10000 | 3.0 |
| NQO1[c] | 240± 12 | 290± 48 | $8.3 \cdot 10^5$ |

[a]Indicated are the upper and lower limits of asymmetric 95% confidence intervals

[b] For these measurements, NQO2 could not be saturated with the co-substrate, and the lack of plateau in the rate at high concentrations of the cosubstrate estimated rate parameters and errors difficult

[c]The kinetic parameters for NQO1 and NQO2 were determined in a previous publication (Islam *et al* 2022)

## Effective use of NRH and BNAH by NQO1, NQO2 and ancestral proteins

Although extant NQO2 cannot use conventional nicotinamide cosubstrates NAD(P)H efficiently, it can catalyze quinone reduction very efficiently given smaller nicotinamide cosubstrates such as NRH, NMEH or BNAH. To test the hypothesis that NQO2 may have evolved to specifically rely on smaller nicotinamide cosubstrates, given menadione as the quinone substrate, we measured the oxidation of BNAH by the ancestral proteins by monitoring the decrease in absorbance at 355 nm. We found that all the ancestral enzymes could use the synthetic smaller cosubstrate BNAH highly efficiently in the range of $10^6 – 10^7$ $M^{-1} \bullet s^{-1}$. The $K_M$ for BNAH ranged from 20 mM to 200 mM, and the $k_{cat}$ was from 1557 $s^{-1}$ to 157 $s^{-1,}$ with no clear evidence for evolutionary selection on either protein lineage. While BNAH is an exogenous cosubstrate not found in cells, there has been some evidence for the existence of NRH under physiological conditions. We tested the hypothesis that NQO2 evolved specifically to use NRH as a reducing cosubstrate as opposed to NAD(P)H by measuring the ability of the ancestral NQO1 and NQO2 proteins to use NRH given menadione as quinone substrate. We found that all the ancestral sequences could use NRH as a reducing cosubstrate as they could with BNAH.

## Table 3 Kinetics of ancient enzymes with BNAH as cosubstrate

| Protein | $k_{cat, BNAH}$ ($s^{-1}$)[a] | $K_{M, BNAH}$ (mM)[a] | $k_{cat}/K_{M, BNAH}$ ($M^{-1} \cdot s^{-1}$) | $K_{M, Menadione}$ (mM)[a] | $K_{I, Menadione}$ (mM)[a] |
|---|---|---|---|---|---|
| CA | 661 (550 – 811) | 162 (104 – 249) | $4.1 \cdot 10^6$ ($5.3 \cdot 10^6 – 7.8 \cdot 10^6$) | 9.04 (5.69 – 14.0) | ND[b] |
| CA1NQO1 | 1178 (990 – 1448) | 44.6 (23.8 – 70.2) | $2.6 \cdot 10^7$ ($2.1 \cdot 10^7 – 4.2 \cdot 10^7$) | 19.5 (14.2 – 27.7) | 26.5 (11.7 – 50.4) |
| CA2NQO1 | 829 (635 – 1201) | 213 (137 – 341) | $3.9 \cdot 10^6$ ($4.6 \cdot 10^6 – 8.8 \cdot 10^6$) | 21.3 (12.9 – 38.5) | 192 (69.5 – 6519) |
| CA1NQO2 | 158 (143 – 178) | 28.3 (19.0 – 39.0) | $5.6 \cdot 10^6$ ($3.7 \cdot 10^6 – 9.4 \cdot 10^6$) | 28.3 (13.4 – 23.7) | ND[b] |
| CA2NQO2 | 1557 (1224 – 2113) | 219 (144 – 353) | $7.1 \cdot 10^6$ ($1.5 \cdot 10^7 – 3.5 \cdot 10^6$) | 28.1 (43.8 – 18.8) | 78.8 (34.1 – 221) |

[a] Indicated are the upper and lower limits of asymmetric 95% confidence intervals
[b] For these measurements we did not observe any substrate inhibition with menadione

## Table 4. Kinetics of quinone reductases with NRH as cosubstrate

| Protein | $k_{cat, NRH}$ ($s^{-1}$)[a] | $K_{M, NRH}$ (mM)[a] | $k_{cat}/K_{M, NRH}$ ($M^{-1} \cdot s^{-1}$) |
|---|---|---|---|
| CA | 194 (160 – 249) | 855 (593 – 1299) | $2.3 \cdot 10^5$ ($2.7 \cdot 10^5 – 1.9 \cdot 10^5$) |
| CA1NQO1 | 130 (113 – 154) | 290 (219 – 394) | $4.5 \cdot 10^5$ ($2.9 \cdot 10^5 – 7.1 \cdot 10^5$) |
| CA2NQO1 | 952 (844 – 1091) | 405 (313 – 532) | $2.4 \cdot 10^6$ ($1.6 \cdot 10^6 – 7.1 \cdot 10^5$) |
| CA1NQO2 | 70.3 (66.4 – 74.6) | 115 (101 – 132) | $6.1 \cdot 10^5$ ($5.0 \cdot 10^5 – 3.5 \cdot 10^6$) |
| CA2NQO2 | 1776 (1607 – 1969) | 158 (117 – 212) | $1.1 \cdot 10^7$ ($7.6 \cdot 10^6 – 1.7 \cdot 10^7$) |
| NQO2 | 101.2 (78.41 – 144.8) | 629 (382 – 1150) | $1.6 \cdot 10^5$ ($3.8 \cdot 10^5 – 6.8 \cdot 10^4$) |
| NQO1 | 17.36 (12.97 – 27.04) | 1150 (2284 – 674.6) | $1.5 \cdot 10^4$ ($4.0 \cdot 10^4 – 5.7 \cdot 10^3$) |

[a] Indicated are the upper and lower limits of asymmetric 95% confidence intervals

**Homology modelling and NQO1 NQO2 chimeras revealed that NQO2 accumulated deleterious mutations all over the protein sequence to lose its ability to use NAD(P)H efficiently.**

Kinetic investigation of the ancestral NQO1 and NQO2 proteins revealed that the two proteins probably underwent functional divergence as NQO1 became more efficient with NAD(P)H. At the same time, NQO2 lost the ability to use those cosubstrates. We hypothesized that structural comparison of the ancestral proteins could highlight the molecular details behind poor NAD(P)H use. Changes in the amino acid sequence which emerged and took place down the NQO1 lineage increased enzymatic efficiency with NAD(P)H, while mutations in the NQO2 lineage had the opposite effect. We did homology modelling with the Swiss Model to infer the tertiary structure of the most recent ancestral proteins, CA2NQO1 and CA2NQO1 and then used those models to infer the structural basis of cosubstrate specificity for NADH.

CA2NQO2 model was constructed using extant human NQO2 structure as the template. The sequence alignment of CA2NQO2 and NQO2 does not have any insertions or deletions and has a sequence identity of 76%, so the global score, GMQE 0.90 and QMEANDisCo 0.86 ± 0.05 were high. Only two short loop regions in the CA2NQO2 model were modelled with low confidence of <50%. Homology modelling of CA2NQO1 was done with NQO1 structure as the template and sequence alignment of the two did not have any insertions or deletions, having a sequence identity of 73%. The predicted model was high quality, with a GMQE of 0.92 and

QMEANDisCo 0.86 ± 0.05. All regions of the CA2NQO1 model matched the template with high confidence, except for a few residues on loops.

We aligned the predicted structures of CA2NQO1 and CA2NQO2 to recognize residues in the protein that are different from each other and were conserved in their respective protein lineages. We found differences in residues all over the protein, as shown in the supplementary. Presumably, comparing the most recent common ancestor of NQO1 and the most recent common ancestor of NQO2 would allow for finding residues in the protein necessary for NADH cosubstrate specificity. We constructed chimeric proteins composed of various segments from both CA2NQO1 and CA2NQO2. The proteins were divided into four segments which were swapped between CA2NQO1 and CA1NQO2, segment 1 at the N-terminal (1- 100), segment 2 (101 – 154), segment 3 (155 – 219), segment 4 (220 – 231) for CA2NQO2 or (220 – 275) for CA2NQO1 (Figure 5A). There was a total of 14 chimeric constructs. They were purified as a fusion with the cellulose-binding protein CBM domain, which enabled protein isolation from the bacterial extract by binding to some filter paper. All the chimeric constructs were likely soluble, and isolation with the CBM strategy was very good as the resulting protein preparations were relatively pure, as visualized by SDS PAGE (Figure 5B). After protein isolation, the chimeric constructs were screened for enzymatic activity with both BNAH and NADH as the reducing cosubstrate and menadione as the quinone substrate. Eleven of the 14 chimeric constructs were enzymatically active.

All the chimeras had a higher turnover with BNAH than with NADH. Some chimeras had a higher turnover of BNAH than the parental protein, but activity with NADH did not show much

improvement. Some chimeras lost the ability to use NADH but retained the ability to use BNAH as a reducing cosubstrate. Although none of the chimeras showed increased activity to CA2QR2 with NADH as the cosubstrate, an overall conclusion can still be made – residues responsible for cofactor specificity are found all over the protein. Modelling showed that the residues in the third segment contributed to the active site charge.

Chimeras with the third segment from CA2QR1 were generally characterized by a more positive and open active site when compared with chimeras with the third segment from CA2QR2. The second segment of the protein seemed to influence activity in general but not specifically for NADH. The second and first segments also had the most influence on the pI of the protein. The closest similarity to CA2QR1 activity with BNAH was 2122 and 1211. These chimeras did not match CA2QR1 in action with NADH as a cofactor. However, 2122 was one of the two enzymes which did show slight movement with NADH, the other being 1122. The only commonality between all three of these chimeras is that they have a CA2QR1 second segment. Further supporting this, when analyzing which region showed the most improvement with a single change in the segment, segment two was found to be the most influential for change in activity

**Table 5. Kinetics of quinone reductase chimeras with NADH and BNAH**

| Proteins | Turnover, NADH ($s^{-1}$) [a] | | SE[b] | Proteins | Turnover, BNAH ($s^{-1}$)[a] | | SE[b] |
|---|---|---|---|---|---|---|---|
| NQO1 | 78 | ± | 3 | NQO1 | 348 | ± | 34 |
| NQO2 | 1.2 | ± | 0.2 | NQO2 | 101 | ± | 5 |
| CA2NQO1 | 116 | ± | 21 | CA2NQO1 | 712 | ± | 7 |
| CA2NQO2 | 5.4 | ± | 3.2 | CA2NQO2 | 21 | ± | 6 |
| 2122 | 6.4 | ± | 0.7 | 2122 | 574 | ± | 10 |
| 1121 | 4.7 | ± | 1.2 | 1121 | 513 | ± | 39 |
| 1122 | 4.5 | ± | 0.3 | 1211 | 263 | ± | 16 |
| 1211 | 3.2 | ± | 1 | 1122 | 182 | ± | 2 |
| 2211 | 3.1 | ± | 0.3 | 2211 | 136 | ± | 4 |
| 1221 | 1.3 | ± | 0.5 | 2111 | 35 | ± | 1 |
| 2221 | 1.2 | ± | 0.6 | 1221 | 22 | ± | 7 |
| 2111 | 0.8 | ± | 0.3 | 2221 | 20 | ± | 5 |
| 1222 | 0.73 | ± | 1.3 | 1222 | 20 | ± | 2 |
| 1112 | 0.56 | ± | 0.4 | 1112 | 17 | ± | 4 |
| 1212 | 0.42 | ± | 0.1 | 2121 | 6.4 | ± | 0.4 |
| 2212 | 0.21 | ± | 0.1 | 2212 | 0.87 | ± | 0.3 |
| 2121 | 0.11 | ± | 0.03 | 1212 | ND[c] | ± | ND[c] |
| 2112 | ND[c] | ± | ND[c] | 2112 | ND[c] | ± | ND[c] |

[a] Indicated are the turnover of 160 mM of NADH and 140 mM of BNAH given 50 mM of menadione.
[b] Errors indicated were calculated from three replicates
[c] These constructs did not show any catalytic activity at these concentrations of substrate even at high protein concentrations

**The molecular insights from the ancestral sequence reconstruction improved the ability of NQO2 to use NADH as a reducing cosubstrate.**

The evolutionary analysis revealed that following the gene duplication from a common ancestor, enzymes in the NQO2 lineage had mutations to decrease their catalytic efficiency with canonical reducing cosubstrates NADH. Although CA1NQO2 and CA have an 85% percentage identity, the changes reduced catalytic activity with NADH 1000-fold. Further changes down the NOQ2 lineage had less effect on the catalytic efficiency. NQO1 became progressively more efficient with NADH. We hypothesized then that, guided by the ancestral sequences, we could make amino acid substitutions on CA1NQO2 that would have a higher NADH turnover rate. Guided by the evolutionary conservation patterns, we made 25 substitutions (Figure 6A) on CA1NQO2, and they improved catalytic efficiency with NADH up to 400-fold. The 25 substitutions increased the turnover of NADH from 0.126 $s^{-1}$ (95% CI, 0.09615 to 0.1852 $s^{-1}$) to 157.9 $s^{-1}$ (95% CI, 142.1 to 178.3 $s^{-1}$). Interestingly, the mutations did not affect the $K_M$ for NADH much, as it only decreased to 972.4 μM (95% CI, 741.7 to 1290 μM) from 1964 μM (95% CI, 1255 to 3430 μM). The use of the smaller reducing cosubstrate, BNAH, did not change much as the catalytic efficiency remained high at $2.08 \cdot 10^6\ M^{-1} \cdot s^{-1}$ compared to $5.6 \cdot 10^6\ M^{-1} \cdot s^{-1}$ of CA1NQO2. The ability to use menadione did not change from CA1NQO2 as the $K_M$ for menadione was 9.04 μM (95% CI, 53.18 to 94.86 μM).

To test for the ability of the C-terminal of NQO1 to affect cosubstrate preference, we also made a chimeric CA1NQO2 mutant with the C-terminal 50 residue tail of CA2NQO1 along with

the 25 mutations (Figure 6B). We found that adding the tail did not affect cosubstrate preference or enzyme kinetics. We did homology modelling with the Swiss Model using NQO2 as a template of CA1NQO2 and the CA1NQO2 mutant to understand how these mutations affected the structure of NQO2. The 25 substitutions were on helices and loops near the FAD of the active site. The modifications changed the electrostatics of the protein's active site that houses the isoalloxazine ring by making the active site more positive, like extant human NQO1.

The active site of extant humans is more negative. Four of the 25 substitutions introduced positive charges, (P15R), (S31K), (N162H), and (Q183K). Further, five neutralized negative charges in NQO2, (E74L), (D128S), (E154S), (E194G), and (E198P). However, new negative charges were also introduced in (N24D), (S36E), and (N62D). The change in charge of the active site could have important implications for NADH binding. The phosphates in the ADP portion have a negative charge, so that the positive active site may contribute to favourable binding. Our results indicated that the increased catalytic efficiency of the mutants was not a result of improved NADH binding entirely, as the $K_M$ for NADH only decreased by a factor of 2 to 3. The bigger improvement was in the enzyme turnover with NADH as a reducing cosubstrate. Electrostatics is likely to be important for enzyme activity. Reduced FAD, $FADH_2$ has a negative charge and must be stabilized by the protein. It is currently thought that the chain of proton transfers achieves this goal. Y156 protonates the reduced flavin, and then Y156 is protonated by H162 in NQO1 and N162 in NQO2. It should be noted that the mutations only improved turnover when using NADH as a reducing cosubstrate and not the use of the smaller cosubstrate. We think that the change in the active site may have resolved some unfavorable charge interactions relating to the negative

charges of the active site, $FADH_2$, and NADH. The CA1NQO2 mutant does not have an intact metal binding site as we substituted the C223F, which coordinated the $Zn^+$ twice. We do not know the metal participating in the catalytic cycle of NQO2.

## Discussion

We found the ancestral proteins more thermostable than the extant protein sequences. This is a good indication that the protein is folding up correctly. Additionally, much literature found that ancestral proteins reconstructed by maximum likelihood tend to have higher thermostabilities. ML is preferred over Bayesian Inference (BI) reconstruction because it has fewer errors due to incorporated uncertainties; however, unlike Bayesian inference, ML is believed to be permissive to systematic bias for generating a more stable protein because of higher selectivity for common amino acids. ML is still commonly used for ancient state reconstruction and a previous study found that ML was the most effective method for predicting ancestral sequences correctly on a site-by-site basis (P. Williams et al., 2006). ML also provides a probabilistic method including branch lengths, substitution propensity and statistically confidence unlike the Maximum Parsimony (MP).

The novel cosubstrate preference provides evidence for neo-functionalization by NQO2. Many examples of neofunctionalization have been discovered because of ancestral sequence reconstruction (pyruvate dehydrogenase versus lactate dehydrogenase in the parasite from Revenant). I think there are many more too. There was an evolutionary selection for NADH use or use badly. Duplicated genes can end up resulting in several different outcomes depending on the acquisition of deleterious or alternative mutations. In order to generate inheritable altered genes, mutation events must take place in the germ line. Acquisition in the population of these changes will then come to a point of either fixation or extinction. Some fates of a gene duplication event include conservation where both copies are maintained, neofunctionalization where one

copy acquires a new function, sub-functionalization where both copies undergo complementary disabling mutations leading to both copies needing to be conserved to maintain ancestral function and non-functionalization where one loses function. Another interesting potential is for specialization, a fate where sub-functionalization is followed by neofunctionalization to result in two genes unique from each other and the common ancestor. In regard to the novel cosubstrate specificity of NQO2, we believe our results provide evidence for neo-functionalization. Unlike non-functionalization, where one duplicate completely loses function through deleterious mutations and often serves as an artifact known as a pseudogene, neofunctionalization is the fate that one copy maintains ancestral function, while the other is able to undertake a new function. While this is just one of several outcomes following gene duplication, it is seems to fit our predictions for NQO2 losing the ability to use NADH while undergoing selective pressure to acquire a novel function. The appearance of the metal binding site in the first ancestor following the duplication event and maintenance throughout the NQO2 lineage points to a structural aspect reflecting this potential new function and divergence from NQO1. As well, we find it unlikely sub functionalization or specialization occurred as ancient NQO1s still maintained the ability to use NADH like the common ancestor.

Sub-functionalization after gene duplication or neo-functionalization. Most genes, after undergoing duplication, may acquire deleterious mutations, which render them a pseudogene. NQO2 acquired mutations probably gave it a new function, which had something to do with not using NADH efficiently. It does look like it had become a pseudogene for many of the organisms which lived under water. Neo-functionalization should not be altogether too surprising. Quinone

reductase enzymes are very versatile. Bacteria have various quinone reductase enzymes (convergent evolution or divergent?). The bacterial quinone reductase, which is homologous to NQO1, seems to have many different functions.

Proteolytic cleavage was also seen for CA and CA1NQO1. We think it is likely that the C-terminal tail is getting proteolyzed by trace amounts of proteases. Trypsin digest supports that the tail is getting cleaved immediately on the CA indicating it is a flexible region. For the CA, some of the posterior probabilities are lower and are also ten residues shorter. As well, alternative predictions did not include the C-terminal. Inaccuracies within the C-terminal could lead to improper folding and great disorder makes this region more prone to cleavage. For CA1NQO1, it is clear again that the tail was subjected to some proteolysis; however, posterior probabilities were higher for this sequence and the lesser cleavage is probably due to the C-terminal being intrinsically more dynamic, like the C-terminal of NQO1 which has also been found to be dynamic and more likely preferential for cleavage (Medina-Carmona et al., 2016).

## References


Bantscheff, Marcus, Dirk Eberhard, Yann Abraham, Sonja Bastuck, Markus Boesche, Scott Hobson, Toby Mathieson, et al. 2007. "Quantitative Chemical Proteomics Reveals Mechanisms of Action of Clinical ABL Kinase Inhibitors." *Nat Biotechnol* 25 (9): 1035–44. https://doi.org/10.1038/nbt1328.

Benoit, C E, S Bastianetto, J Brouillette, Y Tse, J A Boutin, P Delagrange, T Wong, P Sarret, and R Quirion. 2010. "Loss of Quinone Reductase 2 Function Selectively Facilitates Learning Behaviors." *Journal of Neuroscience* 30 (38): 12690–700. https://doi.org/10.1523/JNEUROSCI.2808-10.2010.

Bieganowski, Pawel, and Charles Brenner. 2004. "Discoveries of Nicotinamide Riboside as a Nutrient and Conserved NRK Genes Establish a Preiss-Handler Independent Route to NAD+ in Fungi and Humans." *Cell* 117 (4): 495–502.

Brehmer, Dirk, Klaus Godl, Birgit Zech, Josef Wissing, and Henrik Daub. 2004. "Proteome-Wide Identification of Cellular Targets Affected by Bisindolylmaleimide-Type Protein Kinase C Inhibitors." *Molecular & Cellular Proteomics* 3 (5): 490–500. https://doi.org/10.1074/mcp.M300139-MCP200.

Brouillette, Jonathan, Deborah Young, Matthew J. During, and Rémi Quirion. 2007. "Hippocampal Gene Expression Profiling Reveals the Possible Involvement of *Homer1* and *GABA* $_B$ Receptors in Scopolamine-Induced Amnesia: Gene Expression Profiling in Scopolamine-Treated Rats." *Journal of Neurochemistry* 102 (6): 1978–89. https://doi.org/10.1111/j.1471-4159.2007.04666.x.

Celli, Claudia M, Namphuong Tran, Richard Knox, and Anil K Jaiswal. 2006. "NRH:Quinone Oxidoreductase 2 (NQO2) Catalyzes Metabolic Activation of Quinones and Anti-Tumor Drugs." *Biochemical Pharmacology* 72 (3): 366–76. https://doi.org/10.1016/j.bcp.2006.04.029.

Chen, Shiuan, Kebin Wu, and Richard Knox. 2000. "Structure-Function Studies of DT-Diaphorase (NQO1) and NRH:Quinone Oxidoreductase (NQO2)11This Article Is Dedicated to the Memory of Dr. Lars Ernster, Who Recently Passed Away." *Free Radical Biology & Medicine* 29 (3–4): 276–84. https://doi.org/10.1016/S0891-5849(00)00308-7.

Deller, S, P Macheroux, and S Sollner. 2008. "Flavin-Dependent Quinone Reductases." *Cellular and Molecular Life Sciences : CMLS* 65 (1): 141–60. https://doi.org/10.1007/s00018-007-7300-y.

Duncan, James S, John Lenehan, Lee M Graves, Timothy A Haystead, and David W Litchfield. 2008. "An Unbiased Evaluation of CK2 Inhibitors by Chemoproteomics: Characterization of

Inhibitor Effects on CK2 and Identification of Novel Inhibitor Targets." *Molecular & Cellular Proteomics : MCP* 7 (6): 1077–88. https://doi.org/10.1074/mcp.M700559-MCP200.

Giroud-Gerbetant, Judith, Magali Joffraud, Maria Pilar Giner, Angelique Cercillieux, Simona Bartova, Mikhail V. Makarov, Rubén Zapata-Pérez, et al. 2019. "A Reduced Form of Nicotinamide Riboside Defines a New Path for NAD+ Biosynthesis and Acts as an Orally Bioavailable NAD+ Precursor." *Molecular Metabolism* 30 (December): 192–202. https://doi.org/10.1016/j.molmet.2019.09.013.

Gould, Nathaniel L., Sailendrakumar Kolatt Chandran, Haneen Kayyal, Efrat Edry, and Kobi Rosenblum. 2021. "Somatostatin Interneurons of the Insula Mediate QR2-Dependent Novel Taste Memory Enhancement." *ENeuro* 8 (5): ENEURO.0152-21.2021. https://doi.org/10.1523/ENEURO.0152-21.2021.

Gould, Nathaniel L., Vijendra Sharma, Mohammad Hleihil, Sailendrakumar Kolatt Chandran, Orit David, Efrat Edry, and Kobi Rosenblum. 2020. "Dopamine-Dependent QR2 Pathway Activation in CA1 Interneurons Enhances Novel Memory Formation." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 40 (45): 8698–8714. https://doi.org/10.1523/JNEUROSCI.1243-20.2020.

Graves, Paul R, Jesse J Kwiek, Patrick Fadden, Rupa Ray, Klaas Hardeman, Andrew M Coley, Michael Foley, and Timothy A J Haystead. 2002. "Discovery of Novel Targets of Quinoline Drugs in the Human Purine Binding Proteome." *Molecular Pharmacology* 62 (6): 1364–72.

Harris, Kurt L., Raine E.S. Thomson, Yosephine Gumulya, Gabriel Foley, Saskya E. Carrera-Pacheco, Parnayan Syed, Tomasz Janosik, et al. 2022. "Ancestral Sequence Reconstruction of a Cytochrome P450 Family Involved in Chemical Defense Reveals the Functional Evolution of a Promiscuous, Xenobiotic-Metabolizing Enzyme in Vertebrates." Edited by Meredith Yeager. *Molecular Biology and Evolution* 39 (6): msac116. https://doi.org/10.1093/molbev/msac116.

Hedges, S. Blair. 2009. "The Timetree of Life, Vertebrates." In *The Timetree of Life*, edited by S. Blair Hedges and Sudhir Kumar, 309–14. Oxford Biology. Oxford ; New York: Oxford University Press.

Hobbs, J. K., C. Shepherd, D. J. Saul, N. J. Demetras, S. Haaning, C. R. Monk, R. M. Daniel, and V. L. Arcus. 2012. "On the Origin and Evolution of Thermophily: Reconstruction of Functional Precambrian Enzymes from Ancestors of Bacillus." *Molecular Biology and Evolution* 29 (2): 825–35. https://doi.org/10.1093/molbev/msr253.

Islam, Faiza, Kevin K. Leung, Matthew D. Walker, Shahed Al Massri, and Brian H. Shilton. 2022. "The Unusual Cosubstrate Specificity of NQO2: Conservation Throughout the Amniotes and

Implications for Cellular Function." *Frontiers in Pharmacology* 13 (April): 838500. https://doi.org/10.3389/fphar.2022.838500.

Jaiswal, A K, P Burnett, M Adesnik, and O W McBride. 1990. "Nucleotide and Deduced Amino Acid Sequence of a Human CDNA (NQO2) Corresponding to a Second Member of the NAD(P)H:Quinone Oxidoreductase Gene Family. Extensive Polymorphism at the NQO2 Gene Locus on Chromosome 6." *Biochemistry* 29 (7): 1899–1906.

Klaeger, Susan, Stephanie Heinzlmeir, Mathias Wilhelm, Harald Polzer, Binje Vick, Paul-Albert Koenig, Maria Reinecke, et al. 2017. "The Target Landscape of Clinical Kinase Drugs." *Science* 358 (6367): eaan4368. https://doi.org/10.1126/science.aan4368.

Knox, Richard J, Terence C Jenkins, Stephen M Hobbs, Shiuan Chen, Roger G Melton, and Philip J Burke. 2000. "Bioactivation of 5-(Aziridin-1-Yl)-2,4-Dinitrobenzamide (CB 1954) by Human NAD(P)H Quinone Oxidoreductase 2: A Novel Co-Substrate-Mediated Antitumor Prodrug Therapy." *Cancer Research* 60: 4179–86.

Kwiek, Jesse J, Timothy A J Haystead, and Johannes Rudolph. 2004. "Kinetic Mechanism of Quinone Oxidoreductase 2 and Its Inhibition by the Antimalarial Quinolines." *Biochemistry* 43 (15): 4538–47. https://doi.org/10.1021/bi035923w.

Leung, Kevin K. K., and Brian H. Shilton. 2015. "Quinone Reductase 2 Is an Adventitious Target of Protein Kinase CK2 Inhibitors TBBz (TBI) and DMAT." *Biochemistry* 54 (1): 47–59. https://doi.org/10.1021/bi500959t.

Leung, Kevin Ka Ki, David W Litchfield, and Brian H Shilton. 2012. "Flavin Adenine Dinucleotide Content of Quinone Reductase 2: Analysis and Optimization for Structure-Function Studies." *Analytical Biochemistry* 420 (1): 84–89. https://doi.org/10.1016/j.ab.2011.09.011.

Liao, S, J T Dulaney, and H G Williams-Ashman. 1962. "Purification and Properties of a Flavoprotein Catalyzing the Oxidation of Reduced Ribosyl Nicotinamide." *The Journal of Biological Chemistry* 237 (September): 2981–87.

Liao, Shutsung, and H.G. Williams-Ashman. 1961. "Enzymatic Oxidation of Some Non-Phosphorylated Derivatives of Dihydronicotinamide." *Biochemical and Biophysical Research Communications* 4 (3): 208–13. https://doi.org/10.1016/0006-291X(61)90272-8.

Lind, Christina, Paul Hochstein, and Lars Ernster. 1982. "DT-Diaphorase as a Quinone Reductase: A Cellular Control Device against Semiquinone and Superoxide Radical Formation." *Archives of Biochemistry and Biophysics* 216 (1): 178–85. https://doi.org/10.1016/0003-9861(82)90202-8.

Lynch, Michael, and John S. Conery. 2000. "The Evolutionary Fate and Consequences of Duplicate Genes." *Science* 290 (5494): 1151–55. https://doi.org/10.1126/science.290.5494.1151.

McGrath, Casey L., Jean-Francois Gout, Parul Johri, Thomas G. Doak, and Michael Lynch. 2014. "Differential Retention and Divergent Resolution of Duplicate Genes Following Whole-Genome Duplication." *Genome Research* 24 (10): 1665–75. https://doi.org/10.1101/gr.173740.114.

Merkl, Rainer, and Reinhard Sterner. 2016. "Reconstruction of Ancestral Enzymes." *Perspectives in Science* 9 (December): 17–23. https://doi.org/10.1016/j.pisc.2016.08.002.

Ohno, Susumu. 1970. "Introduction." In *Evolution by Gene Duplication*, by Susumu Ohno, 1–2. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-86659-3_1.

Rappaport, Akiva N., Eyal Jacob, Vijendra Sharma, Sharon Inberg, Alina Elkobi, Hadile Ounallah-Saad, Metsada Pasmanik-Chor, Efrat Edry, and Kobi Rosenblum. 2015. "Expression of Quinone Reductase-2 in the Cortex Is a Muscarinic Acetylcholine Receptor-Dependent Memory Consolidation Constraint." *The Journal of Neuroscience* 35 (47): 15568–81. https://doi.org/10.1523/JNEUROSCI.1170-15.2015.

Rix, Uwe, Oliver Hantschel, Gerhard Dürnberger, Lily L Remsing Rix, Melanie Planyavsky, Nora V Fernbach, Ines Kaupe, et al. 2007. "Chemical Proteomic Profiles of the BCR-ABL Inhibitors Imatinib, Nilotinib, and Dasatinib Reveal Novel Kinase and Nonkinase Targets." *Blood* 110 (12): 4055–63. https://doi.org/10.1182/blood-2007-07-102061.

Ross, David, and David Siegel. 2021. "The Diverse Functionality of NQO1 and Its Roles in Redox Control." *Redox Biology* 41 (May): 101950. https://doi.org/10.1016/j.redox.2021.101950.

Sihvola, Virve, and Anna-Liisa Levonen. 2017. "Keap1 as the Redox Sensor of the Antioxidant Response." *Archives of Biochemistry and Biophysics* 617 (March): 94–100. https://doi.org/10.1016/j.abb.2016.10.010.

Vasiliou, Vasilis, David Ross, and Daniel W Nebert. 2006. "Update of the NAD(P)H:Quinone Oxidoreductase (NQO) Gene Family." *Human Genomics* 2 (5): 329–35.

Vella, Fanny, Gilles Ferry, Philippe Delagrange, and Jean A Boutin. 2005. "NRH:Quinone Reductase 2: An Enzyme of Surprises and Mysteries." *Biochemical Pharmacology* 71 (1–2): 1–12. https://doi.org/10.1016/j.bcp.2005.09.019.

Walsh, Bruce. 2003. "Population-Genetic Models of the Fates of Duplicate Genes." In *Origin and Evolution of New Gene Functions*, edited by M. Long, 10:279–94. Contemporary Issues in Genetics and Evolution. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-0229-5_16.

Wilson, C, R V Agafonov, M Hoemberger, S Kutter, A Zorba, J Halpin, V Buosi, et al. 2015. "Kinase Dynamics. Using Ancient Protein Kinases to Unravel a Modern Cancer Drug's

Mechanism." *Science (New York, NY)* 347 (6224): 882–86. https://doi.org/10.1126/science.aaa1823.

Winger, Jonathan A, Oliver Hantschel, Giulio Superti-Furga, and John Kuriyan. 2009. "The Structure of the Leukemia Drug Imatinib Bound to Human Quinone Reductase 2 (NQO2)." *BMC Structural Biology* 9 (1): 7. https://doi.org/10.1186/1472-6807-9-7.

Yang, Yue, Farheen Sultana Mohammed, Ning Zhang, and Anthony A. Sauve. 2019. "Dihydronicotinamide Riboside Is a Potent NAD+ Concentration Enhancer in Vitro and in Vivo." *Journal of Biological Chemistry* 294 (23): 9295–9307. https://doi.org/10.1074/jbc.RA118.005772.

Yang, Yue, Ning Zhang, Guoan Zhang, and Anthony A. Sauve. 2020. "NRH Salvage and Conversion to NAD+ Requires NRH Kinase Activity by Adenosine Kinase." *Nature Metabolism* 2 (4): 364–79. https://doi.org/10.1038/s42255-020-0194-9.

Zarei, Amin, Leila Khazdooz, Mojtaba Enayati, Sara Madarshahian, Timothy J. Wooster, Gerhard Ufheil, and Alireza Abbaspourrad. 2021. "Dihydronicotinamide Riboside: Synthesis from Nicotinamide Riboside Chloride, Purification and Stability Studies." *RSC Advances* 11 (34): 21036–47. https://doi.org/10.1039/D1RA02062E.

Zhang, Jianzhi. 2003. "Evolution by Gene Duplication: An Update." *Trends in Ecology & Evolution* 18 (6): 292–98. https://doi.org/10.1016/S0169-5347(03)00033-8.

Zhao, Q., X. L. Yang, W. D. Holtzclaw, and P. Talalay. 1997. "Unexpected Genetic and Structural Relationships of a Long-Forgotten Flavoenzyme to NAD(P)H:Quinone Reductase (DT-Diaphorase)." *Proceedings of the National Academy of Sciences* 94 (5): 1669–74. https://doi.org/10.1073/pnas.94.5.1669.

# Figure Legends

## Figure 1. Reconstruction of NQO1 and NQO2 Ancient Sequences

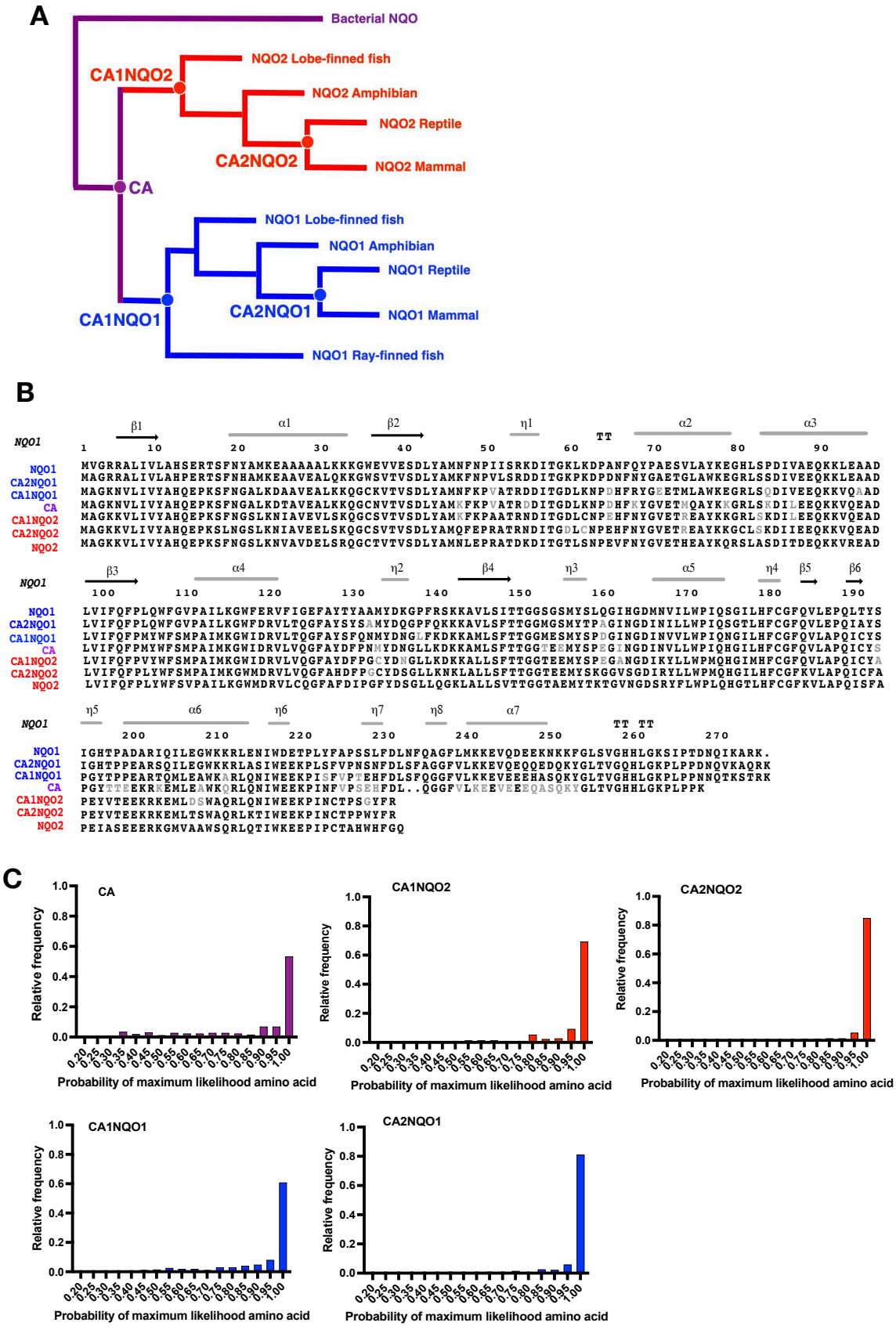
(A) Schematic of the phylogenetic tree of NQO1 and NQO2 sequences highlighting critical ancestral proteins that were resurrected in this work. CA or common ancestor was the ancestral gene that duplicated to NQO1 and NQO2. (B) Multiple sequence alignment of NQO1 and NQO2 along with the maximum likelihood sequence of the resurrected ancestral proteins. The residues in grey were predicted with probability below 0.6. NQO1 and NQO2 have the same overall structure and the secondary structure schematic shown at the top of the figure is of NQO1. (C) To assess the reliability of the ancient sequence reconstructions, the maximum likelihood probability for predicted residues in the ancestral proteins are shown against the relative frequency or fraction of all residues for a given protein.

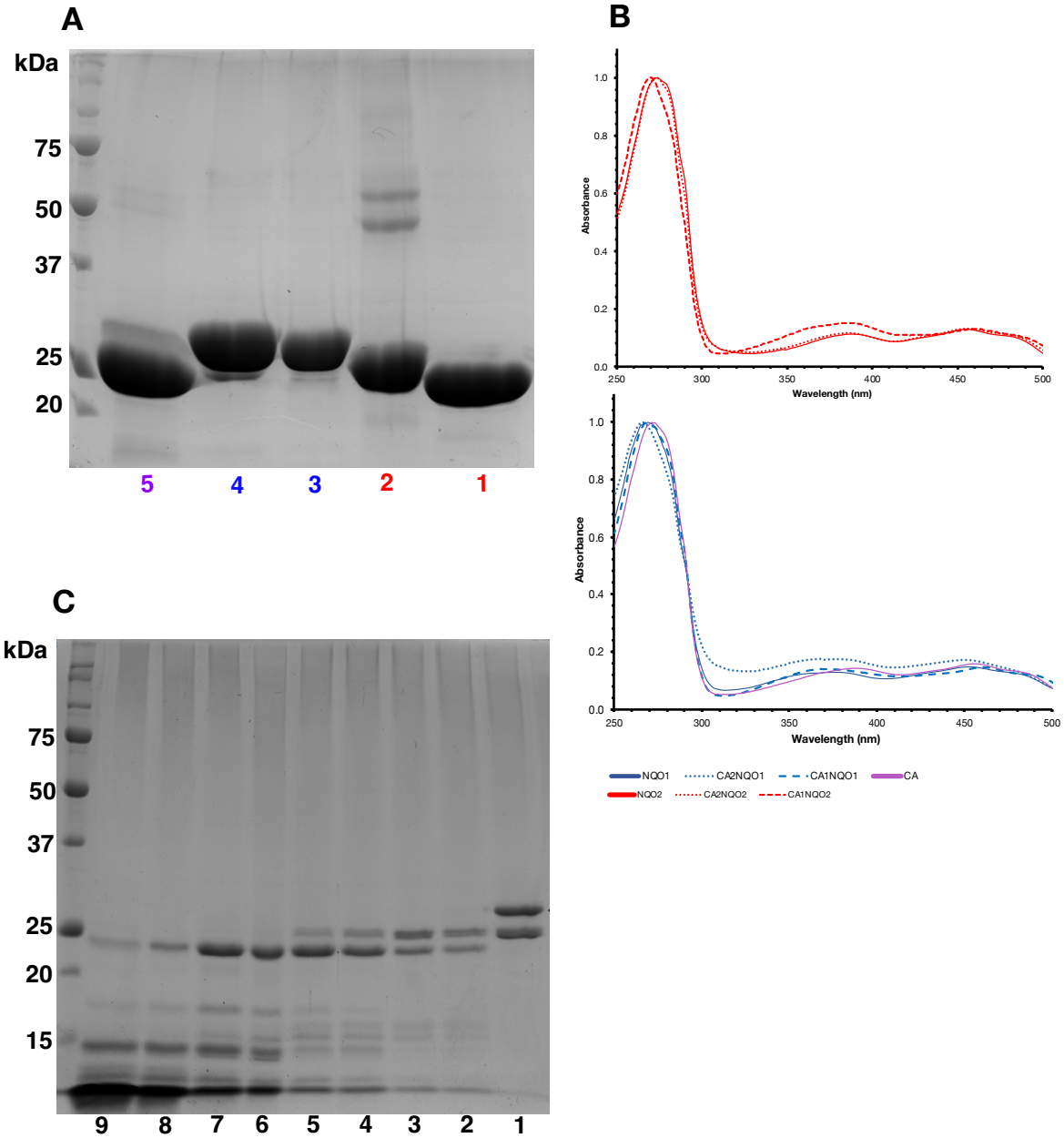
## Figure 2. Purification and Characterization of Ancient Quinone Reductases

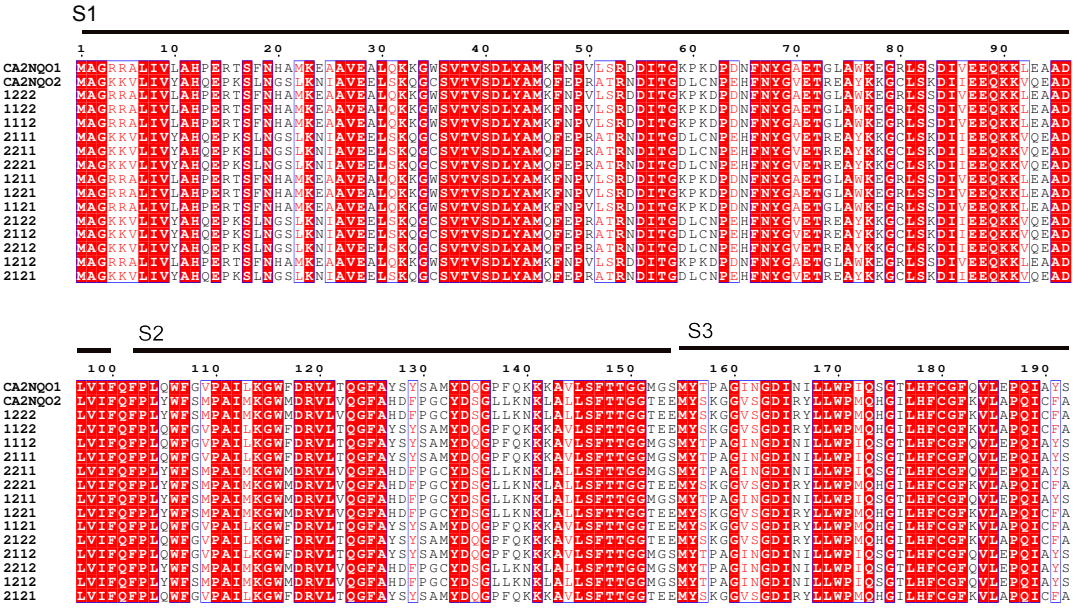
(A) SDS-PAGE analysis of the final ancient protein preparations: lanes 1 and 2 are the ancestors of NQO2 (CA2NQO2 and CA1NQO2, respectively), lanes 3 and 4 are the ancestors of NQO1 (CA2NQO1 and CA1NQO1, respectively) and lane 5 is the common ancestor (CA) of the NQO1 and NQO2 lineages. (B) UV-Visible absorbance spectra of the purified quinone reductases. (C) Partial proteolysis of the ancient CA protein, demonstrating that the predicted C-terminal tail for this protein is likely unstructured and subject to proteolysis, as was observed during its purification.

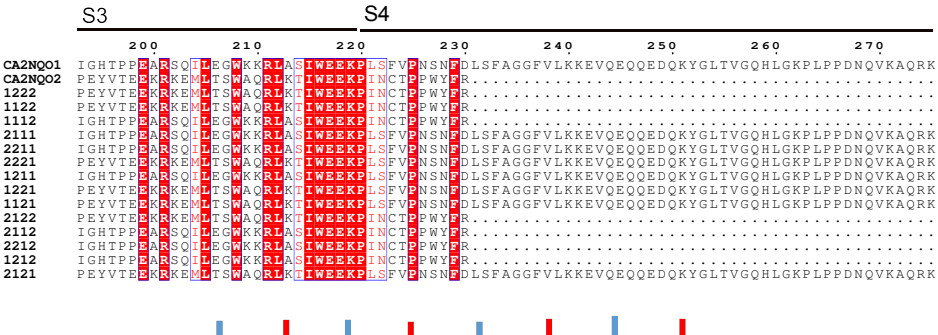
**Figure 3. Sequences and purification of chimeric quinone reductase sequences**

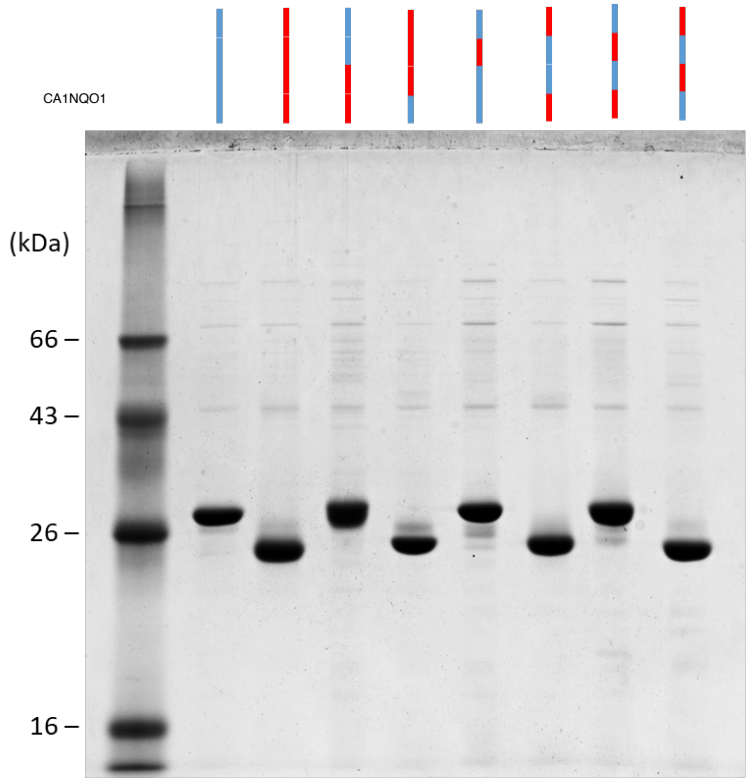
The sequences of the common ancestors for NQO1 and NQO2, at the top of the phylogenetic tree for the amniotes (i.e. CA2NQO1 and CA2NQO2) were divided into four regions that were then mixed together to produce 14 chimeric proteins that were purified as fusions with a cellulose binding domain. (A) The sequences of the chimeric proteins are provided along with the two parental sequences, CA2NQO1 and CA2NQO2. The lines above the sequence alignment, S1, S2, S3, and S4 indicate the 4 regions of the parental sequences that were shuffled to produce the chimeras. (B) SDS-PAGE analysis of representative purifications of 8 of the chimeras. The sequence composition of the chimeras is indicated by the schematics above the lanes, with blue representing sequence from CA2NQO1, and red representing sequence from CA2NQO2.


A
Bacterial NQO
CA1NQO2
NQO2 Lobe-finned fish
NQO2 Amphibian
NQO2 Reptile
CA2NQO2
NQO2 Mammal
CA
NQO1 Lobe-finned fish
NQO1 Amphibian
NQO1 Reptile
CA1NQO1
CA2NQO1
NQO1 Mammal
NQO1 Ray-finned fish
B
NQO1
NQO1 MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKLKDPANFQYPAESVLAYKEGHLSPDIVAEQKKLEAAD
CA2NQO1 MAGRRALIVLAHPERTSFNHAMKEAAVEALQKKGWSVTVSDLYAMKFNPVLSRDDITGKPKDPDNFNYGAETGLAWKEGRLSSDIVEEQKKLEAAD
CA1NQO1 MAGKNVLIVYAHQEPKSFNGALKDAAVEALKKQGCKVTVSDLYAMNFKPVATRDDITGDLKNPDHFRYGEETMLAWKEGRLSQDIVEEQKKVQAAD
CA MAGKNVLIVYAHQEPKSFNGALKDTAVEALKKQGCNVTVSDLYAMKFKPVATRDDITGDLKNPDHFKYGVETMQAYKKGRLSKDILEEQKKVQEAD
CA1NQO2 MAGKKVLIVYAHQEPKSFNGSLKNIAVEVLSKQGCSVTVSDLYAMKFKPAATRNDITGDLCNPEHFNYGVETREAYKKGRLSKDILEEQKKVQEAD
CA2NQO2 MAGKKVLIVYAHQEPKSLNGSLKNIAVEELSKQGCSVTVSDLYAMQFEPRATRNDITGDLCNPEHFNYGVETREAYKKGCLSKDIIEEQKKVQEAD
NQO2 MAGKKVLIVYAHQEPKSFNGSLKNVAVDELSRQGCTVTVSDLYAMNLEPRATDKDITGTLSNPEVFNYGVETHEAYKQRSLASDITDEQKKVREAD
NQO1 LVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRSKKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYS
CA2NQO1 LVIFQFPLQWFGVPAILKGWFDRVLTQGFAYSYSAMYDQGPFQKKKAVLSFTTGGMGSMYTPAGINGDINILLWPIQSGTLHFCGFQVLEPQIAYS
CA1NQO1 LVIFQFPMYWFSMPAIMKGWIDRVLTQGFAYSFQNMYDNGLFKDKKAMLSFTTGGMESMYSPDGINGDINVVLWPIQNGILHFCGFQVLAPQICYS
CA LVIFQFPMYWFSMPAIMKGWIDRVLVQGFAYDFPNMYDNGLLKDKKAMLSFTTGGTEEMYSPEGINGDINVLLWPIQHGILHFCGFQVLAPQICYS
CA1NQO2 LVIFQFPVYWFSMPAIMKGWIDRVLVQGFAYDFPGCYDNGLLKDKKALLSFTTGGTEEMYSPEGANGDIKYLLWPMQHGIMHFCGFQVLAPQICYA
CA2NQO2 LVIFQFPLYWFSMPAIMKGWMDRVLVQGFAHDFPGCYDSGLLKNKLALLSFTTGGTEEMYSKGGVSGDIRYLLWPMQHGILHFCGFKVLAPQICFA
NQO2 LVIFQFPLYWFSVPAILKGWMDRVLCQGFAFDIPGFYDSGLLQGKLALLSVTTGGTAEMYTKTGVNGDSRYFLWPLQHGTLHFCGFKVLAPQISFA
NQO1 IGHTPADARIQILEGWKKRLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK.
CA2NQO1 IGHTPPEARSQILEGWKKRLASIWEEKPLSFVPNSNFDLSFAGGFVLKKEVQEQQEDQKYGLTVGQHLGKPLPPDNQVKAQRK
CA1NQO1 PGYTPPEARTQMLEAWKARLQNIWEEKPISFVPTEHFDLSFQGGFVLKKEVEEEHASQKYGLTVGHHLGKPLPPNNQTKSTRK
CA PGYTTEEKRKEMLEAWKQRLQNIWEEKPINFVPSEHFDL..QGGFVLKEEVEEEQASQKYGLTVGHHLGKPLPPK
CA1NQO2 PEYVTEEKRKEMLDSWAQRLQNIWEEKPINCTPSGYFR
CA2NQO2 PEYVTEEKRKEMLTSWAQRLKTIWEEKPINCTPPWYFR
NQO2 PEIASEEERKGMVAAWSQRLQTIWKEEPIPCTAHWHFGQ
C
CA
CA1NQO2
CA2NQO2
CA1NQO1
CA2NQO1
Relative frequency
Probability of maximum likelihood amino acid


A
kDa
75
50
37
25
20
5
4
3
2
1
B
Absorbance
Wavelength (nm)
NQO1
CA2NQO1
CA1NQO1
CA
NQO2
CA2NQO2
CA1NQO2
C
kDa
75
50
37
25
20
15
9
8
7
6
5
4
3
2
1


S1
S2
S3


S3
S4


CA1NQO1
(kDa)
66 –
43 –
26 –
16 –