Supporting Information

Lohman et al. 10.1073/pnas.1413128111

SI Materials and Methods

Proteomics Analysis of Mouse Tissue by LC-MS/MS.

Sample preparation. Mouse tissue (heart, cerebellum, and kidney) from C57BL/6 mice (three WT, three *Coq9*^{R239X}; females) was harvested at 3 mo as described by García-Corzo et al. (1). Tissue homogenization was performed at 4 °C with proteomics lysis buffer [urea (8 M), Tris, pH 8 (40 mM), NaCl (30 mM), $CaCl_2$ (1 mM), phosphatase inhibitors (prepared in-house), protease inhibitors (cOmplete, Mini; Roche)] in a glass homogenizer with a PTFE pestle (4 mL; Kontes) (1,000 rpm, 15 strokes). The homogenate was flash-frozen in liquid nitrogen and stored at −80 °C.

Homogenate was thawed at 4 °C and lysed on ice using a probe sonicator. Protein content was evaluated using a bicinchoninic acid assay (Thermo Fisher Scientific). Proteins were reduced with 5 mM DTT (incubation at 58 °C for 30 min) and alkylated with 15 mM iodoacetamide (incubation in the dark, at ambient temperature, for 30 min). Alkylation was quenched by adding an additional 5 mM DTT (incubation at ambient temperature for 15 min). Samples were diluted to a final concentration of 1.5 M urea (pH 8) with a solution of 50 mM Tris and 5 mM $CaCl_2$ before undergoing enzymatic digestion with sequencing-grade trypsin (Promega) at a ratio of 1:50 (enzyme/protein, mass/mass). The resulting mixtures were incubated at ambient temperature overnight. The following morning, each sample was incubated with an additional aliquot of sequencing-grade trypsin at ratio of 1:100 (enzyme/protein, mass/mass) for ∼1.5 h. Digests were quenched by bringing the pH ∼2 with trifluoroacetic acid and immediately desalted using C18 solid-phase extraction columns (SepPak; Waters).

Desalted material was labeled with TMT six-plex isobaric labels (Thermo-Pierce). Before quenching the TMT reactions, ∼5 μg of material from each TMT channel was combined into a test mix and analyzed by LC-MS/MS to evaluate labeling efficiency and obtain optimal ratios for sample recombination. Following quenching, tagged peptides were combined in equal amounts by mass (∼500 μg per channel) and desalted. All experiments had ≥97% labeling efficiency, calculated by the number of N-terminal–labeled peptides divided by the total number of peptide identifications.

Sample fractionation. Labeled peptides were fractionated by strong cation exchange using a polysulfoethylaspartamide column (9.4 × 200 mm; PolyLC) on a Surveyor LC quaternary pump (Thermo Scientific). Each dried and mixed TMT sample was resuspended in buffer A, injected onto the column, and subjected to the following gradient for separation: 100% buffer A from 0 to 2 min, 0–15% buffer from 2 to 5 min, and 15–100% buffer B from 5 to 35 min. Buffer B was held at 100% for 10 min, and then the column was washed extensively with buffer C and water before recalibration. Flow rate was held at 3.0 mL/min throughout the separation. Buffer compositions were as follows: buffer A [5 mM KH_2PO_4, 30% (vol/vol) acetonitrile (ACN) (pH 2.65)], buffer B [5 mM KH_2PO_4, 350 mM KCl, 30% (vol/vol) ACN (pH 2.65)], and buffer C [50 mM KH_2PO_4, 500 mM KCl (pH 7.5)]. Eleven fractions were collected over the first 50-min elution period and were immediately frozen, lyophilized, and desalted.

LC-MS/MS analysis. All experiments were performed using a NanoAcquity UPLC system (Waters) coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). Reverse-phase columns were made in-house by packing a fused silica capillary (75-μm i.d., 360-μm o.d., with a laser-pulled electrospray tip) with 1.7-μm diameter, 130-Å pore size Bridged Ethylene Hybrid C18 particles (Waters) to a final length of 30 cm. The column was heated to 60 °C for all experiments. Samples were loaded onto the column for 12 min in 95:5 buffer A [water, 0.2% formic acid, and 5% DMSO (vol/vol/vol)]/buffer B (ACN, 0.2% formic acid, and 5% DMSO) at a flow rate of 0.30 μL/min. Peptides were eluted using the following gradient: an increase to 7% B over 1 min, followed by a 42-min linear gradient from 7% to 18% B, followed by a 28-min linear gradient from 18% to 27% B, followed by a final 1-min ramp to 75% B, which was held for 3 min. The column was equilibrated with 5% buffer B for an additional 25 min. Precursor peptide cations were generated from the eluent through the utilization of a nano-electrospray ionization (ESI) source.

Mass spectrometry instrument methods consisted of MS^1 survey scans [1e6 target value; 60,000 resolution; 300–1,500 thomsons (Th)] that were used to guide 15 subsequent data-dependent MS/MS scans [3-Th isolation window, higher-energy collisional dissociation (HCD) fragmentation, normalized collision energy of 35; 5e4 target value, 30,000 resolution]. Dynamic exclusion duration was set to 30 s, with a maximum exclusion list of 500 and an exclusion width of 0.55 Th below and 2.55 Th above the selected average mass. Maximum injection times were set to 50 ms for all MS^1 scans and 150 ms for MS/MS scans.

Data analysis. Data were processed using the in-house software suite COMPASS (2). OMSSA (3) (version 2.1.8) searches were performed against a target-decoy database [Uniprot (mouse); www.uniprot.org; August 7, 2013]. Searches were conducted using a 150-ppm precursor mass tolerance and a 0.015-Da product mass tolerance. A maximum of three missed tryptic cleavages were allowed. The fixed modifications specified were carbamidomethylation of cysteine residues, TMT six-plex on peptide N termini, and TMT six-plex on lysine residues. The variable modifications specified were oxidation of methionine and TMT six-plex on tyrosine residues. TMT quantification of identified peptides was performed within COMPASS as described previously (4). Peptides identified within each of 11 fractions were grouped into proteins according to previously reported rules (5) using COMPASS. Protein quantification was performed by summing all of the reporter ion intensities within each channel for all peptides uniquely mapping back to a given protein. All quantitative data were normalized at the protein level, $\log_2$ transformed, and mean normalized. Fold changes were calculated by averaging normalized protein values for each condition and calculating the difference of averages. For each comparison, a P value was calculated using Student t test (assuming equal variance).

Lipidomics Analysis of Mouse Tissue by LC-MS/MS.

Sample preparation. Mouse cerebellums were prepared as described in proteomic analysis of mouse tissue by LC-MS/MS with the lipid homogenization buffer [phosphate (11.8 mM), NaCl (137 mM), KCl (2.7 mM), pH 7.2, phosphatase inhibitors (prepared in-house), protease inhibitors (prepared in-house)]. Lipids were extracted by the following procedure at 4 °C. All lipid standards were purchased from Avanti and prepared in CH_3Cl/MeOH (1:1, vol/vol) or ACN/isopropanol (IPA)/H_2O (65:30:5, vol/vol/vol): PE(15:0/15:0) 850704P; PC(11:011:0) 850330C; PG(15:0/15:0) 840446X; CL(14:0)$_4$ 710332C; PA(17:0/17:0) 830856P; PS (17:0/17:0) 840028P; DG(12:012:0) 800812P; Q_6 901150; DHE 810253P; LPC(20:0) 855777P; PG(17:0/17:0) 830456P. Mouse cerebellum homogenate (1 mg of protein, 375 μL) was mixed with internal standard (25 μL of 0.2 mM each lipid) in a glass tube. $CHCl_3$/MeOH (1:1, vol/vol, 5 mL) (one phase) was added and vortexed (60 s). HCl (1 M, 100 μL) was added to acidify the extract followed by addition of 0.5 NaCl, 0.5 M HCl (600 μL) to induce phase separation. Phase separation was completed by

centrifugation (1,000 × *g*, 5 min, 4 °C). The upper aqueous phase was removed by aspiration. A volume of 3 mL of the lower organic phase was isolated and dried under $N_{2(g)}$. The organic residue was reconstituted in ACN/IPA/H_2O (65:30:5, vol/vol/vol, 100 μL) containing PG(17:0/17:0) (50 μM) by vortexing (2 min) and sonicating (1 s).

LC-MS/MS analysis. Discovery-type lipidomics analyses were conducted using the instrumentation and column described in Yeast CoQ quantitation by LC-MS. Mobile phase A consisted of 10 mM ammonium formate in H_2O/ACN (40:60, vol/vol) containing 0.1% formic acid. Mobile phase B consisted of 10 mM ammonium formate in IPA/ACN (90:10, vol/vol) containing 0.1% formic acid. Ten microliters of lipid extract were separated by a nonlinear gradient from 32% to 97% B over 35 min at 260 μL/min and 45 °C by adapting previously published methods (6, 7). The MS was operated in polarity switching mode acquiring positive and negative mode MS^1 and MS^2 spectra (Top1) during the same separation with a 1.29-s total duty cycle. MS acquisition parameters were 17,500 resolving power, 1×10^6 automatic gain control (AGC) target for MS^1 and 2×10^5 AGC target for MS^2 scans, 35 units of sheath gas and 10 units of auxiliary gas, 300 °C HESI II and inlet capillary temperature, 100-ms MS^1 and 50-ms MS^2 ion accumulation time, 200- to 1,600-Th MS^1 scan range, 2-Th isolation width for fragmentation, stepped HCD collision energy (15, 30, 60 units), 0.5% under fill ratio, and 8-s dynamic exclusion. The separation was monitored for irregularities by acquisition of UV/Vis spectra (200–600 nm).

Data analysis. LC-MS/MS data were analyzed using SIEVE (Thermo; version 2.0). The output from this analysis was a list of 5,000 frames with corresponding MS^1 masses, RT values, and peak intensities for each sample. Individual lipid species were identified by searching the Metlin database (metlin.scripps.edu), human metabolome database (HMDB; www.hbmd.ca), or LIPID MAPS database (www.lipidmaps.org) for species that matched the MS^1 masses. To distinguish lipid species that have the same MS^1 masses, MS^2 spectra were manually inspected to identify characteristic fragments. Lipid quantitation was performed by normalizing the MS^1 signal integration to the corresponding lipid species of internal standard added during extraction.

Cloning of $COQ9^{N\Delta79}$ Constructs and Mutants for *Escherichia coli* Expression. The COQ9 (COQ9_HUMAN) protein used in this study was cloned, expressed, and purified using a modified NESG standard procedure (8–10). Briefly, residues 79–318 of WT COQ9 (HR5043) were cloned into the expression vector pET15_8HMbpTEV_NESG, yielding plasmid HR5043-79–318-MBP3.11. The resulting fused ORF consists of an N-terminal eight-residue His-tag followed by a derivative of the maltose binding protein (MBP), a 3C protease recognition site, a TEV protease recognition site, and finally residues 79–318 of COQ9 (8His-MBP-[TEV]-$COQ9^{N\Delta79}$). Following purification and TEV protease cleavage, three nonnative residues (SHM) remain at the N terminus of the COQ9 (79–318) protein variant. The crystallized $COQ9^{N\Delta79}$ construct contained the E144G mutation, but all other constructs were WT COQ9. $COQ9^{N\Delta79}$ mutants were generated via standard site-directed mutagenesis [50-ng template, 0.2 mM dNTPs, 0.2 μM each primer, 3% (vol/vol) DMSO, 0.625 U of Pfu Ultra AD]. After an initial 2-min denaturation at 95 °C, the reaction temperature was cycled 26 times: 95 °C for 15 s, 60 °C for 30 s, and 68 °C for 7 min. The reactions were incubated with 10 U of DpnI to destroy template DNA, and then transformed into DH5α *E. coli* cells. Mutants were confirmed by DNA sequencing.

Cloning of $COQ9^{N\Delta45}$ and $COQ7^{N\Delta38}$ for Cell-Free Expression. The human COQ9 gene (Uniprot O75208) encoding amino acids 46–318 was cloned into the SgfI and PmeI sites of the pEU wheat germ cell-free expression vector (11) with a N-terminal Strep(II) affinity tag (MGWSHPQFEKAIAENLYFQS). The human COQ7 isoform 2 (Uniprot Q99807-2) was similarly cloned, but with an N-terminal 6×His tag (HHHHHHAIAENLYFQ). The full-length human COQ7 and COQ9 ORFs were PCR amplified while appending an N-terminal TEV cleavage site using a two-step PCR protocol and the PCR product inserted into the pEU-His-Flexi vector (12). PIPE cloning was used to make the COQ9 and COQ7 truncations as well as exchanging the His6-tag for the Strep(II)-tag in the pEU-Strep2-$COQ9^{N\Delta45}$ construct (13). All PCR steps were done using PfuUltra II polymerase following manufacturer's protocols. The coding regions for all constructs were verified by DNA sequencing. Point mutations in the COQ9 ORF were made in situ using standard site-directed mutagenesis as described in the cloning of $COQ9^{N\Delta79}$ constructs and mutants for *E. coli* expression.

Cloning of *Coq9* Constructs and Mutants. The *Saccharomyces cerevisiae* gene *Coq9* was amplified by Accuprime Pfu polymerase (Invitrogen) with primers generating MluI (forward) and KpnI (reverse) restriction sites. Both the *Coq9* amplicon and the yeast expression vector p426 GPD were digested with MluI and KpnI and subsequently ligated and transformed into DH5α *E. coli*. Plasmid minipreps were performed, and recombinants were confirmed by sequencing. *Coq9* mutants were generated via standard site-directed mutagenesis as described in the cloning of $COQ9^{N\Delta79}$ constructs and mutants for *E. coli* expression.

Isolation and Purification of $COQ9^{N\Delta79}$ for Crystallization. *Escherichia coli* BL21 (DE3) pMGK cells, a rare-codon enhanced strain (14), were transformed with the DNA sequence-verified HR5043-79–318-MBP3.11 plasmid. A single isolate was cultured in MJ9 minimal media (9) supplemented with selenomethionine, lysine, phenylalanine, threonine, isoleucine, leucine, and valine for the production of selenomethionine-labeled $COQ9^{N\Delta79}$ (8). Initial growth was carried out at 37 °C until the OD_{600} of the culture reached ∼0.8 units. The incubation temperature was then decreased to 17 °C, and protein expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 1 mM. Following overnight incubation at 17 °C, the cells were harvested by centrifugation and resuspended in Lysis Buffer [50 mM Tris, pH 7.5, 500 mM NaCl, 1 mM tris (2-carboxyethyl)phosphine, 40 mM imidazole]. After sonication, the supernatant was collected by centrifugation for 40 min at 30,000 × *g*. The supernatant was loaded first onto a Ni affinity column (HisTrap HP; GE Healthcare) and the eluate loaded into a gel filtration column (Superdex 75 26/60; GE Healthcare). The fractions containing 8His-MBP-[TEV]-$COQ9^{N\Delta79}$ were collected, and TEV protease was added at a ratio of 1:50 (wt/wt). After incubating the reaction overnight at 4 °C, the uncleaved MBP-COQ9, MBP tag, and TEV protease were removed by passing through a second Ni affinity column. The $COQ9^{N\Delta79}$ protein was then concentrated and further purified on a Superdex75 26/60 gel filtration column. The purified $COQ9^{N\Delta79}$ protein was in a buffer containing 10 mM Tris·HCl, 100 mM NaCl, 5 mM DTT, pH 7.5, and concentrated to ∼5.5 mg/mL Sample was flash-frozen in 50-μL aliquots using liquid nitrogen and stored at −80 °C before crystallization trials. The sample purity (>98%), molecular weight, and oligomerization state were verified by SDS/PAGE, MALDI-TOF mass spectrometry, and analytic gel filtration followed by static light scattering, respectively. For static light scattering, selenomethionine-labeled $COQ9^{N\Delta79}$ (30 μL at 10 mM in Tris·HCl, pH 7.5, 100 mM NaCl, 5 mM DTT) was injected onto an analytical gel filtration column (Shodex KW-802.5; Shodex) with the effluent monitored by refractive index (Optilab rEX) and 90° static light-scattering (miniDAWN TREOS; Wyatt Technology) detectors.

Isolation and Purification of $COQ9^{N\Delta79}$ for Differential Scanning Fluorimetry Experiments. *E. coli* cell pellets with overexpressed protein (15) (8His-MBP-[TEV]-$COQ9^{N\Delta79}$) were thawed on ice. The pellets were resuspended in Lysis Buffer [50 mM KH_2PO_4, 20 mM Tris·HCl (pH 7.2), 300 mM NaCl, 5 mM BME (β-mercaptoethanol), 0.25 mM PMSF, 1 mg/mL lysozyme, peptide protease inhibitors, pH 7.2] and incubated (1 h, 4 °C, with rocking). The cells were lysed by sonication (4 °C, 6 V, 60 s × 5, with ≥60-s rests between sonications to cool lysate). The lysate was clarified by centrifugation (15,000 × *g*, 30 min, 4 °C). The cleared lysate was mixed with cobalt immobilized metal affinity chromatography (IMAC) resin (Talon resin) and incubated (4 °C, with gentle agitation, 1 h). The resin was pelleted by centrifugation (700 × *g*, 5 min, 4 °C) and washed twice with Wash Buffer (50 mM KH_2PO_4, 20 mM Tris·HCl [pH 7.2], 300 mM NaCl, 5 mM BME, 0.25 mM PMSF, 10 mM imidazole, peptide protease inhibitors [500 ng/mL leupeptin hemisulfate, pepstatin A, chymostatin, aprotinin, and antipain dihydrochloride (Sigma)], pH 7.2) (10 bed volumes). The resin was resuspended in Wash Buffer (1 bed volume), transferred to a polypropylene chromatography column, and allowed to settle by gravity. The resin was washed again with Wash Buffer (5 bed volumes). His-tagged protein was eluted with Elution Buffer [50 mM KH_2PO_4, 20 mM Tris·HCl (pH 7.2), 300 mM NaCl, 5 mM BME, 0.25 mM PMSF, 100 mM imidazole, pH 7.2] (8 mL). The eluted protein was concentrated with a M_r-cutoff spin filter (50-kDa M_r cutoff) and exchanged into Storage Buffer [50 mM KH_2PO_4, 20 mM Tris·HCl (pH 7.2), 300 mM NaCl, 5 mM BME, 0.25 mM PMSF, pH 7.2]. The concentration of 8His-MBP-[TEV]-$COQ9^{N\Delta79}$ was determined by its absorbance at 280 nm ($\varepsilon = 102{,}130\ M^{-1}\cdot cm^{-1}$) ($M_r = 71.3$ kDa). The protein was incubated with TEV protease (1:50, TEV/fusion protein, mass/mass) (2 h, 20 °C, with gentle agitation). The TEV protease reaction mixture was mixed with cobalt IMAC resin (Talon resin), incubated (4 °C, with gentle agitation, 1 h), transferred to a polypropylene chromatography column, and allowed to settle by gravity. The unbound $COQ9^{N\Delta79}$ was allowed to flow through the column and was isolated. The flow through was concentrated with a M_r-cutoff spin filter (10-kDa M_r cutoff) and exchanged into Storage Buffer [50 mM KH_2PO_4, 20 mM Tris·HCl (pH 7.2), 300 mM NaCl, 5 mM BME, 0.25 mM PMSF, pH 7.2]. The concentration of $COQ9^{N\Delta79}$ was determined by its absorbance at 280 nm ($\varepsilon = 36{,}130\ M^{-1}\cdot cm^{-1}$) ($M_r = 27.4$ kDa). The protein was aliquoted, frozen in $N_{2(l)}$, and stored at −80 °C. Fractions from the protein preparation were analyzed by lithium dodecyl sulfate/PAGE.

$COQ9^{N\Delta45}$ and $COQ7^{N\Delta38}$ Cell-Free Expression and Purification. $COQ9^{N\Delta45}$ and $COQ7^{N\Delta38}$ plasmid DNA was treated with 0.05 μg/μL proteinase K to remove trace amounts of RNase, purified, and used as individual transcription templates with SP6 RNase polymerase. Transcription and translation methods are as previously described (16). Briefly, transcription reactions included 0.2 mg/mL DNA, 20 mM magnesium acetate, 2 mM spermidine trihydrochloride, 10 mM DTT, 80 mM Hepes-KOH, pH 7.8, 4 mM each NTP, pH 7.0, 1.6 U/μL SP5 RNA polymerase (Promega), and 1 U/μL RNasin (Promega). Transcriptions were incubated for 4 h at 37 °C. Nonpurified transcription reactions were used singly or mixed 1:1 $COQ9^{N\Delta45}$:$COQ7^{N\Delta38}$ and added to wheat germ extract (WEPRO2240; CellFree Sciences) in a standard dialysis cup reaction (16). Each 25-μL reaction contained 60 O.D. wheat germ extract (6.25 μL), 24 mM Hepes-KOH, pH 7.8, 100 mM potassium acetate, 6.25 mM magnesium acetate, 0.4 mM spermidine trihydrochloride, 4 mM DTT, 1.2 mM ATP, 0.25 mM GTP, 16 mM creatine phosphate, 0.0005% sodium azide, 0.04 mg/mL creatine kinase, 0.3 mM each amino acid (amino acid mixture adjusted to pH 7.0 with KOH), and 5 μL of nonpurified RNA. Thirty-two–fold excess dialysis buffer was used for each reaction, containing all reaction components except for wheat germ extract, creatine kinase, and RNA. After assembly of dialysis cups and 18-h incubation at 22 °C, duplicate translations (50 μL) of each combination were pooled and centrifuged for 5 min at 20,000 × *g* at 10 °C. The soluble fraction was removed and added to 20 μL of StrepTactin resin equilibrated in 25 mM Hepes, pH 7.8, 150 mM NaCl, and 1 mM DTT in a 96-well filter plate (HTS Multiscreen; Millipore). The resin with bound protein was washed three times with 150 μL of binding buffer. Strep(II)-tagged protein and protein complexes were eluted in 2.5 mM desthiobiotin in the same buffer. After sampling for SDS/PAGE, the eluate was applied to Ni-Sepharose (GE Healthcare) in 50 mM Hepes, pH 8.0, 300 mM NaCl, 50 mM imidazole, and 1 mM DTT and subjected to IMAC. Bound samples were eluted by increasing the imidazole concentration to 500 mM. Samples of both the StrepTactin and IMAC elutions were loaded without heating on 4–20% Stain-free TGX Criterion gels (Bio-Rad) and subjected to denaturing SDS/PAGE. Gels were imaged by tryptophan fluorescence using a Bio-Rad Stain Free Imager, followed by staining in Coomassie Brilliant Blue R-250.

Crystallization. The human $COQ9^{N\Delta79}$ was crystallized by microbatch method at 18 °C. Two microliters of protein solution containing $COQ9^{N\Delta79}$ protein (5.5 mg/mL), 5 mM Tris (pH 7.5), 100 mM NaCl, 5 mM DTT, and 0.02% NaN_3 were mixed with 2 μL of the precipitating solution consisting of 100 mM Tris (pH 8.5) and 20% (vol/wt) PEG 3350, and 200 mM $MgCl_2$. Crystals of $COQ9^{N\Delta79}$ appeared in 2 d and were grown to full-length after several days, which were subsequently cryoprotected by 25% (vol/vol) ethylene glycol, and then flash-frozen in liquid N_2 for data collection at 100 K.

X-Ray Data Collection and Structure Determination. The crystals of human $COQ9^{N\Delta79}$ belong to space group $P2_1$ with cell parameters of $a = 38.3$ Å, $b = 97.6$ Å, and $c = 64.1$ Å, and $\beta = 95.8°$. There are two molecules of $COQ9^{N\Delta79}$ in the crystallographic asymmetric unit of the protein. A single-wavelength anomalous diffraction dataset to 2.4-Å resolution was collected at the peak absorption wavelength of selenium at the X4A beamline of the National Synchrotron Light Source. The diffraction images were processed with the HKL package (17), and 17 of 22 possible selenium sites were located with the program Shelx (18). SOLVE/RESOLVE (19) was used for phasing the reflections and automated model building, which correctly placed 49% of the residues with side chains in each protomer of ASU. The entire model was built with the program XtalView (20) and refined by CNS (21). Noncrystallographic symmetry restraint was applied for most stages of the refinement of the structure, but it was released at the final refinement stage. The data processing and refinement statistics are summarized in Table 1.

Dali Server Structure Analysis. The COQ9 structure was used as a query protein structure to search the Protein Structure Database with DaliLite, version 3, through the Dali server (22).

Computational Modeling of DNA-Binding Residues. DNA binding site identifier (DBSI) (23) is a structure-based model that accurately predicts binding sites for DNA on the 3D surface of a protein. This model uses a machine learning strategy for the classification of surface residues as binding or nonbinding with respect to DNA. DBSI uses carefully designed and selected features that encapsulate a range of physical, chemical, geometric, and evolutionary properties of the protein surface that enable DNA recognition. These features include various residue characteristics, secondary structure, solvent accessibility, electrostatic potential, atomic density, and surface geometry. Microenvironment features at near scale average out the effects

of small-scale structural perturbation. At the larger scale, microenvironment features analyzes nonlocal cooperative effects.

DBSI has been trained, validated, and tested on several datasets consisting of high-resolution structures of protein–DNA complexes, and is among the most accurate and robust methods for identification of binding sites on protein structures. Structures tested for this study are as follows: COQ9, monomer and dimer; FadR, monomer and dimer (3ANP); TetR, 1QPI and 4ABZ; DesT, 3LSP. Structures were stripped of ligands, ions, water molecules, and other hetero-atoms. MSEs were replaced by MET residues. To generate electrostatic features within DBSI, solutions to the Poisson–Boltzmann equation were generated using PBEQ (24) via the CHARMM-GUI (25) website (www.charmm-gui.org/). The CHARMM-GUI parameters that differ from default (23) are as follows: dielectric constant for the protein interior (2.0); coarse finite-difference grid spacing (1.0 Å); fine finite-difference grid spacing (0.5 Å). The calculations were repeated independently for these structures in both monomeric and dimeric forms. The DBSI program was then run with these structures and their respective electrostatic potential maps as inputs. The resulting predictions for each surface residue were recorded, and each residue was colored by its DBSI prediction score (red indicating strong binding prediction and blue indicating a surface that is predicted to bind DNA very poorly).

Homology Modeling. Models of mouse and yeast Coq9/Coq9p were manually made and refined by program XtalView (20), and CNS (21) was used for minimization of the steric clashes in the yeast model.

Residue Conservation Analysis. The COQ9 human sequence was used as a query protein using the Consurf server and default constraints (26–28).

Copurifying Lipid Identification by LC-MS/MS. Lipids were extracted using the procedure described in *Lipidomics Analysis of Mouse Tissue by LC-MS/MS* with the following adaptations. Three internal standard mixtures were prepared (0.1×, 0.5×, or 0.01× internal standard described above) and one internal standard mixture was added to each COQ9 sample. Recombinant human COQ9 samples purified for crystallization ($COQ9^{N\Delta 95}$ NESG HR5043.004 or $COQ9^{N\Delta 79}$ HR5043-79–318-MBP3.11) or COQ9 purified for cell-free expression ($COQ9^{N\Delta 45}$ CESG JR.28840) were prepared as follows: 750 μg of protein was mixed with CH_3Cl/MeOH (1:1, vol/vol, 5 mL). $HCl_{(aq)}$ (1 M to 400 μL) was added to acidify the mixture. Lipid LC-MS/MS analysis was performed as described in *Lipidomics Analysis of Mouse Tissue by LC-MS/MS*. LC-MS/MS data were analyzed as described in *Lipidomics Analysis of Mouse Tissue by LC-MS/MS* with the following adaptations: only the most abundant lipids were identified and lipid species were quantified (in picomoles) by comparison with the internal standard curve.

Differential Scanning Fluorimetry (29). Mixtures (20 μL, total volume) of $COQ9^{N\Delta 79}$ (2 μM), SYPRO Orange dye (Invitrogen) (5×), NaCl (150 mM), Hepes (100 mM, pH 7.5) were prepared in Micro Amp Optical 96 (0.2 mL) Well Reaction Plates (Applied Biosystems). Plates were sealed with Optical Adhesive Covers (Applied Biosystems). The fluorescence of the mixture was measured as the temperature was increased from 25 °C to 95 °C using an Applied Biosystems ViiA7 Real-Time PCR System. The samples were heated initially at a rate of 2 °C/s until the temperature reached 15 °C. Then, the samples were heated at a rate of 0.017 °C/s until the temperature reached 99.9 °C. Fluorescence measurements were collected continuously from 25 °C to 95 °C using excitation filter X1 (470 ± 15 nm) and emission filter m4 (623 ± 14 nm). The fluorescence data were analyzed with Protein Thermal Shift software (Applied Biosystems) to determine T_m values. ΔT_m values were determined by subtracting the T_m of the mutant protein from the T_m of WT $COQ9^{N\Delta 79}$ ($\Delta T_m = T_{m,\ Mutant} - \Delta T_{m,\ WT}$).

$COQ9^{N\Delta 45}$-$COQ7^{N\Delta 38}$ Complex Quantitation by Tryptophan Fluorescence. The intensity of the bands, determined using Image Lab software, was calibrated against a fitted intensity curve generated using the known molecular weights, number of tryptophans, and nanogram load of bands in the protein marker lanes. The intensity was integrated with the molecular weight and tryptophan count in each band to determine the relative molar yield of each protein. Strength of the $COQ9^{N\Delta 45}$–$COQ7^{N\Delta 38}$ interaction was quantified by determining the molar ratio of $COQ7^{N\Delta 38}$ with each mutant $COQ9^{N\Delta 45}$-containing complex after two rounds of affinity purification to that containing the WT $COQ9^{N\Delta 45}$. Only samples on the same gel were compared with determine relative molar yield; each gel was calibrated to its intrinsic marker lanes.

Yeast Transformation. Yeast transformations were carried out as previously described (30). *Δcoq9* yeast transformed with p426 plasmids encoding for Coq9p variants were grown on uracil dropout (Ura^-) synthetic media plates with glucose [2% (wt/vol)]. Individual colonies of yeast were used to inoculate Ura^- media (10 g/L glucose) starter cultures (3 mL), which were incubated (30 °C, ∼12 h, 230 rpm).

Yeast Drop Assay. Serial dilutions of yeast from a starter culture were prepared in Ura^- media without glucose. 10^4, 10^3, 10^2, or 10 yeast cells were dropped onto Ura^- media solid media plates containing either glucose [2% (wt/vol)] or glycerol [3% (vol/vol)] and incubated (30 °C, 2–4 d).

Yeast CoQ Quantitation by LC-MS. A total of 2.5×10^6 yeast cells (as determined by OD_{600} of a starter culture) was used to inoculate a 25-mL culture of Ura^- media (10 g/L glucose), which was incubated (30 °C, 230 rpm) for 23 h. After 23 h, the yeast cultures were ∼4 h past the diauxic shift and the medium was depleted of glucose. The OD_{600} of the culture was measured and used to determine the volume of culture needed to isolate 1×10^9 yeast cells. A total of 1×10^9 yeast cells was pelleted by centrifugation (3,000 × *g*, 3 min, 4 °C), the supernatant was discarded, and the yeast pellet was frozen at –20 °C.

Frozen pellets of yeast (2×10^9 cells total) were thawed on ice and mixed with PBS (400 μL), glass beads (0.5 mm diameter, 200 μL), and the yeast were lysed by vortexing with the glass beads (30 s). Q_{10} was added as an internal standard (10 μM, 10 μL). The yeast lysate was vortexed again (30 s). Hexanes/2-propanol (10:1, vol/vol) (1 mL) was added and vortexed (2 × 30 s). Brine (600 μL) was added and vortexed (2 × 30 s). The samples were centrifuged (3,000 × *g*, 5 min, 4 °C) to complete phase separation. A volume of 700 μL of the organic phase was transferred to a clean tube and dried under $N_{2(g)}$. The organic residue was reconstituted in ACN/IPA/H_2O (65:30:5, vol/vol/vol) (100 μL) by vortexing (30 s) and transferred to a glass vial for LC-MS analysis.

LC-MS analysis was performed on an Ascentis Express C18 column held at 35 °C (150 mm × 2.1 mm × 2.7 μm particle size; Supelco) using an Accela LC Pump (500 μL/min flow rate; Thermo Scientific). Mobile phase A consisted of 10 mM ammonium acetate in ACN/H_2O (70:30, vol/vol) containing 250 μL/L acetic acid. Mobile phase B consisted of 10 mM ammonium acetate in IPA/ACN (90:10, vol/vol) with the same additives. Initially, mobile phase B was held at 50% for 1.5 min and then increased to 95% over 6.5 min where it was held for 2 min. The column was then reequilibrated for 3.5 min before the next injection. Ten microliters of sample were injected by an HTC PAL autosampler (Thermo Scientific). The LC system was coupled to a Q Exactive mass spectrometer (Build 2.3 SP2) by a HESI II

heated ESI source kept at 325 °C (Thermo Scientific). The inlet capillary was kept at 350 °C, sheath gas was set to 60 units, and auxiliary gas to 20 units, and the spray voltage was set to 3,000 V. The MS was operated in scheduled targeted MS^2 (tMS^2) mode to quantify DMQ_6 and CoQ_6. From 4.5 to 6.25 min, the $[M+H]^+$ ions of DMQ_6 and CoQ_6 at 561.43 and 591.44 Th were isolated and fragmented, and from 7.5 to 9.0 min, the $[M+H]^+$ ion of the internal standard CoQ_{10} at 880.72 Th was isolated and fragmented. Resolving power was set to 17,500, AGC target to 2×10^5, maximum injection time to 500 ms, isolation window to 1 Th, and collision energy to 27 units. Quantitation proceeded by integrating the peak areas of the characteristic product ions at 167.07 Th (DMQ_6) and 197.08 Th (CoQ species) using the Xcalibur software suite (2.2 SP1.48; Thermo Scientific) and normalizing to the internal standard. The separation was monitored for unusual peaks by UV detection at 265 nm (5 Hz, 49-nm filter bandwidth, 200-ms filter rise time) using an Accela PDA Detector (Thermo Scientific) set up in line with the MS.

1. García-Corzo L, et al. (2013) Dysfunctional Coq9 protein causes predominant encephalomyopathy associated with CoQ deficiency. *Hum Mol Genet* 22(6):1233–1248.
2. Wenger CD, Phanstiel DH, Lee MV, Bailey DJ, Coon JJ (2011) COMPASS: A suite of pre- and post-search proteomics software tools for OMSSA. *Proteomics* 11(6):1064–1074.
3. Geer LY, et al. (2004) Open mass spectrometry search algorithm. *J Proteome Res* 3(5): 958–964.
4. Phanstiel DH, et al. (2011) Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nat Methods* 8(10):821–827.
5. Nesvizhskii AI, Aebersold R (2005) Interpretation of shotgun proteomic data: The protein inference problem. *Mol Cell Proteomics* 4(10):1419–1440.
6. Bird SS, Marur VR, Sniatynski MJ, Greenberg HK, Kristal BS (2011) Lipidomics profiling by high-resolution LC-MS and high-energy collisional dissociation fragmentation: Focus on characterization of mitochondrial cardiolipins and monolysocardiolipins. *Anal Chem* 83(3):940–949.
7. Hu C, et al. (2008) RPLC-ion-trap-FTMS method for lipid profiling of plasma: Method validation and application to p53 mutant mouse model. *J Proteome Res* 7(11):4982–4991.
8. Doublié S, et al. (1996) Crystallization and preliminary X-ray analysis of the 9 kDa protein of the mouse signal recognition particle and the selenomethionyl-SRP9. *FEBS Lett* 384(3):219–221.
9. Jansson M, et al. (1996) High-level production of uniformly ^{15}N- and ^{13}C-enriched fusion proteins in *Escherichia coli*. *J Biomol NMR* 7(2):131–141.
10. Xiao R, et al. (2010) The high-throughput protein sample production platform of the Northeast Structural Genomics Consortium. *J Struct Biol* 172(1):21–33.
11. Sawasaki T, Hasegawa Y, Tsuchimochi M, Kasahara Y, Endo Y (2000) Construction of an efficient expression vector for coupled transcription/translation in a wheat germ cell-free system. *Nucleic Acids Symp Ser* 2000(44):9–10.
12. Blommel PG, Martin PA, Wrobel RL, Steffen E, Fox BG (2006) High efficiency single step production of expression plasmids from cDNA clones using the Flexi Vector cloning system. *Protein Expr Purif* 47(2):562–570.
13. Klock HE, Lesley SA (2009) The polymerase incomplete primer extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. *Methods Mol Biol* 498:91–103.
14. Burgess-Brown NA, et al. (2008) Codon optimization can improve expression of human genes in *Escherichia coli*: A multi-gene study. *Protein Expr Purif* 59(1):94–102.
15. Fox BG, Blommel PG (2009) Autoinduction of protein expression. *Curr Protoc Protein Sci* Chap 5:Unit 5.23.
16. Makino S, Beebe ET, Markley JL, Fox BG (2014) Cell-free protein synthesis for functional and structural studies. *Methods Mol Biol* 1091:161–178.
17. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
18. Sheldrick GM (1990) Phase annealing in Shelx-90—direct methods for larger structures. *Acta Crystallogr A* 46:467–473.
19. Terwilliger TC (2003) SOLVE and RESOLVE: Automated structure solution and density modification. *Methods Enzymol* 374:22–37.
20. McRee DE (1999) XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* 125(2-3):156–165.
21. Brünger AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54(Pt 5): 905–921.
22. Holm L, Rosenström P (2010) Dali server: Conservation mapping in 3D. *Nucleic Acids Res* 38(web server issue):W545–W549.
23. Zhu X, Ericksen SS, Mitchell JC (2013) DBSI: DNA-binding site identifier. *Nucleic Acids Res* 41(16):e160.
24. Jo S, Vargyas M, Vasko-Szedlar J, Roux B, Im W (2008) PBEQ-Solver for online visualization of electrostatic potential of biomolecules. *Nucleic Acids Res* 36(web server issue):W270–W275.
25. Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *J Comput Chem* 29(11):1859–1865.
26. Berezin C, et al. (2004) ConSeq: The identification of functionally and structurally important residues in protein sequences. *Bioinformatics* 20(8):1322–1324.
27. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 38(web server issue):W529–W533.
28. Celniker G, et al. (2013) ConSurf: Using evolutionary data to raise testable hypotheses about protein function. *Isr J Chem* 53(3-4):199–206.
29. Niesen FH, Berglund H, Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* 2(9): 2212–2221.
30. Gietz RD, Woods RA (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 350:87–96.

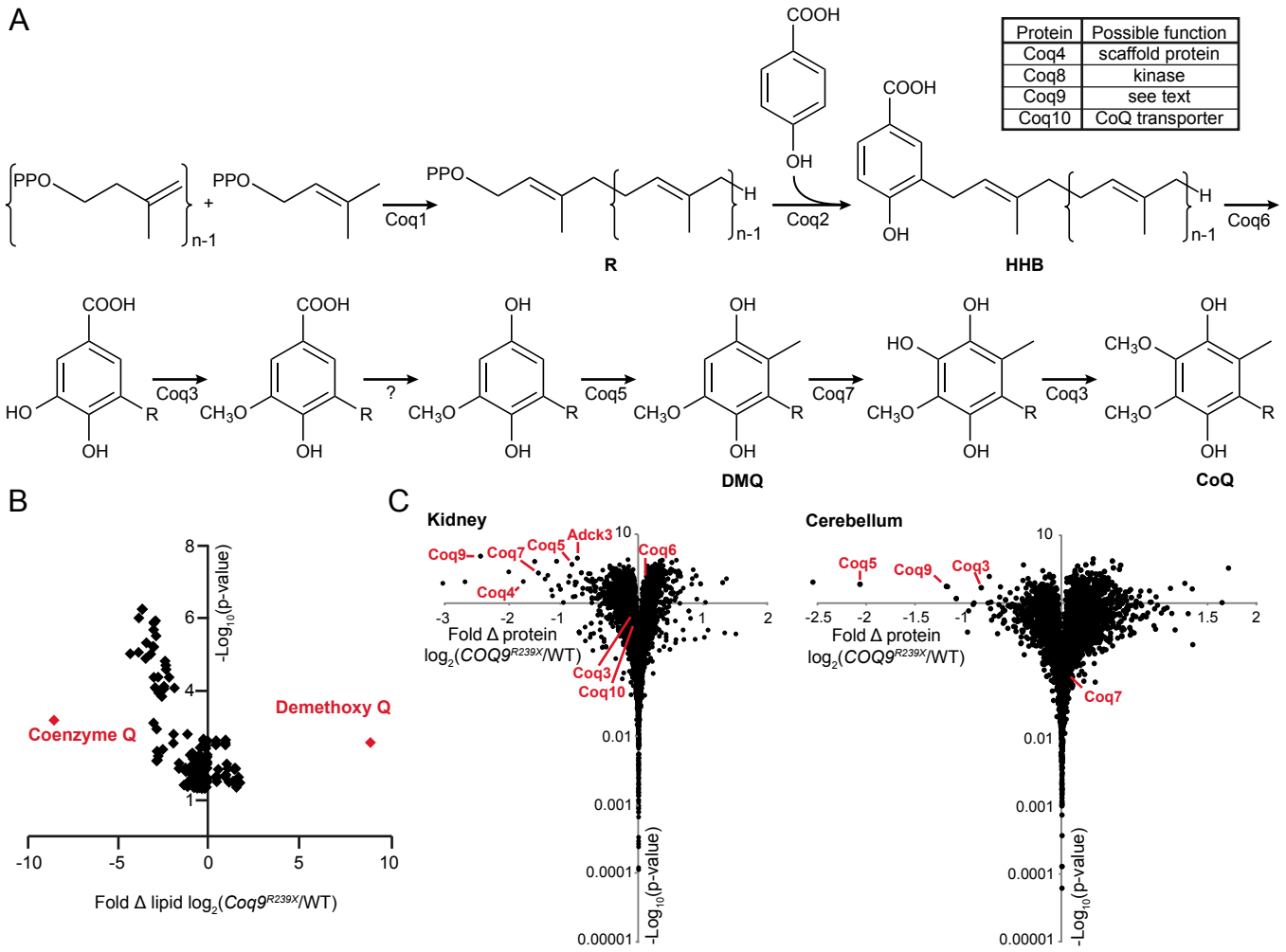
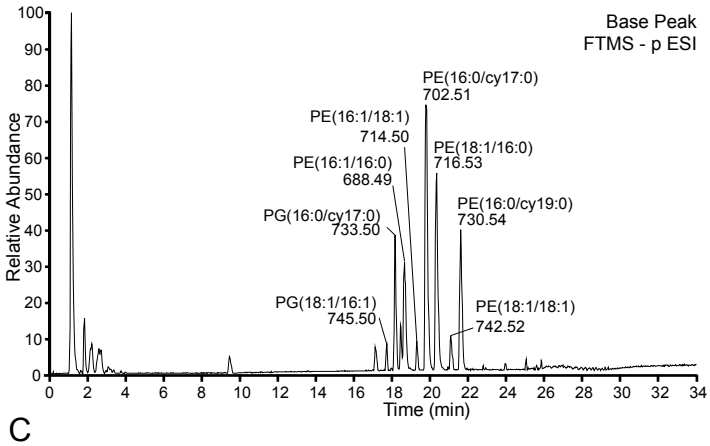


Fig. S1. Coenzyme Q (CoQ) biosynthesis in WT and *Coq9*^{R239X} mice (related to Fig. 1). (*A*) CoQ biosynthesis in eukaryotes, with proteins of unknown function tabulated and CoQ intermediates 3-hexaprenyl-4-hydroxybenzoic acid (HHB) and demethoxy Q (DMQ) labeled (1). (*B*) Lipids (◆) identified by LC-MS/MS in positive ionization mode (*P* value < 0.05 from mice cerebellum; $n = 3$) plotted as Fold Δ $\log_2$(*Coq9*^{R239X}/WT) vs. $-\text{Log}_{10}$(*P* value). The most changing lipid species are labeled in red. (*C*) Proteins (●) identified by LC-MS/MS in mice kidney and cerebellum plotted as $-\text{Log}_{10}$(*P* value) vs. Fold Δ $\log_2$(*Coq9*^{R239X}/WT) with COQ proteins labeled ($n = 3$). See also Dataset S1.

1. He CH, Xie LX, Allan CM, Tran UC, Clarke CF (2014) Coenzyme Q supplementation or over-expression of the yeast Coq8 putative kinase stabilizes multi-subunit Coq polypeptide complexes in yeast coq null mutants. *Biochim Biophys Acta* 1841(4):630–644.

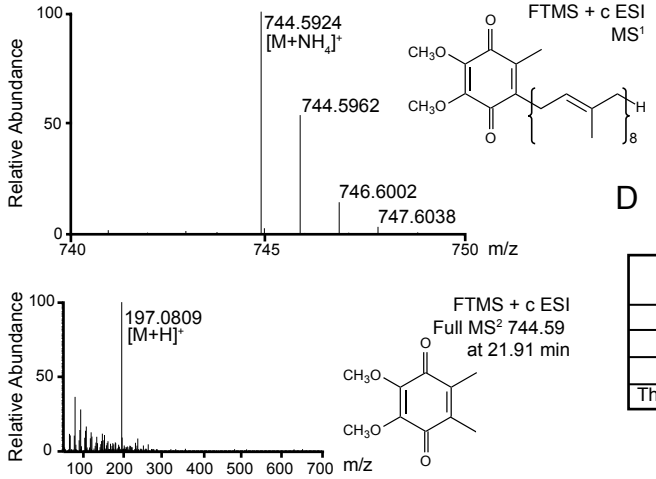
A



B

E. coli		Wheat germ	
[M-H]⁻	**Lipid Species**	**[M-H]⁻**	**Lipid Species**
702	PE(16:0, cy17:0)	738	PE(18:2/18:2)
716	PE(18:1/16:0)	714	PE(18:2/16:0)
730	PE(16:0/cy19:0)	736	PE(18:2/18:3)
688	PE(16:1/16:0)	740	PE(18:1/18:2)
728	PE(cy17:0/18:1)	738	PE(18:2/18:2)
714	PE(16:1/18:1) and PE(cy17:0/cy17:0)	792	PC(18:2/16:0)
756	Pe(18:1/cy19:0)	802	PC(18:2/16:0)
690	PE(16:0/16:0)	824	PC(18:2/18:3)
700	PE(16:1/cy17:0)	816	PC(18:2/18:2)
733	PG(16:0/cy17:0)	745	PG(18:2/16:0)
747	PG(16:0/18:1)		
761	PG(16:0/cy19:0)		
773	PG(cy19:0/cy17:0)		
759	PG(18:1/cy17:0)		
745	PG(18:1/16:1)		
719	PG(16:1/16:0)		
742	PE(18:1/18:1) and PE(cy17:0/cy19:0)		
[M+H]⁺		**[M+NH₄]⁺**	
744	Coenzyme Q_8	812	Coenzyme Q_9

C



D

Lipid	Molar ratio (COQ9/Lipid)
PE	4.8
PG	12.7
CoQ_8	152.2
Theoretical	4

E

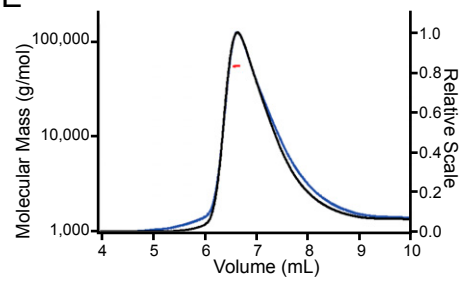


Fig. S2. Identification of copurifying lipids with recombinant COQ9 (related to Fig. 3). (*A*) Base peak chromatogram of copurifying lipids with recombinant COQ9 from *E. coli* identified by LC-MS/MS in negative ionization mode. (*B*) Table of lipid species copurifying with recombinant COQ9 from *E. coli* or wheat germ extract. PE, phosphatidylethanolamine; PG, phosphatidylglycerol. (*C*) Representative LC-MS/MS MS^1 and MS^2 chromatograms for lipid identification with the identification of coenzyme Q_8 in positive mode shown. (*D*) Table of experimental molar ratios between recombinant COQ9 and copurifying lipids from *E. coli* along with the theoretical ratio assuming one lipid molecule per COQ9 dimer at 50% occupancy, as in crystal structure. (*E*) Static light scattering shows that molecular weight of recombinant $COQ9^{N\Delta 79}$ is 55.12 kDa (red), consistent with the expected dimer (55.89 kDa), as determined by gel filtration analysis monitoring the refractive index (blue line) and 90° light scattering (black line).

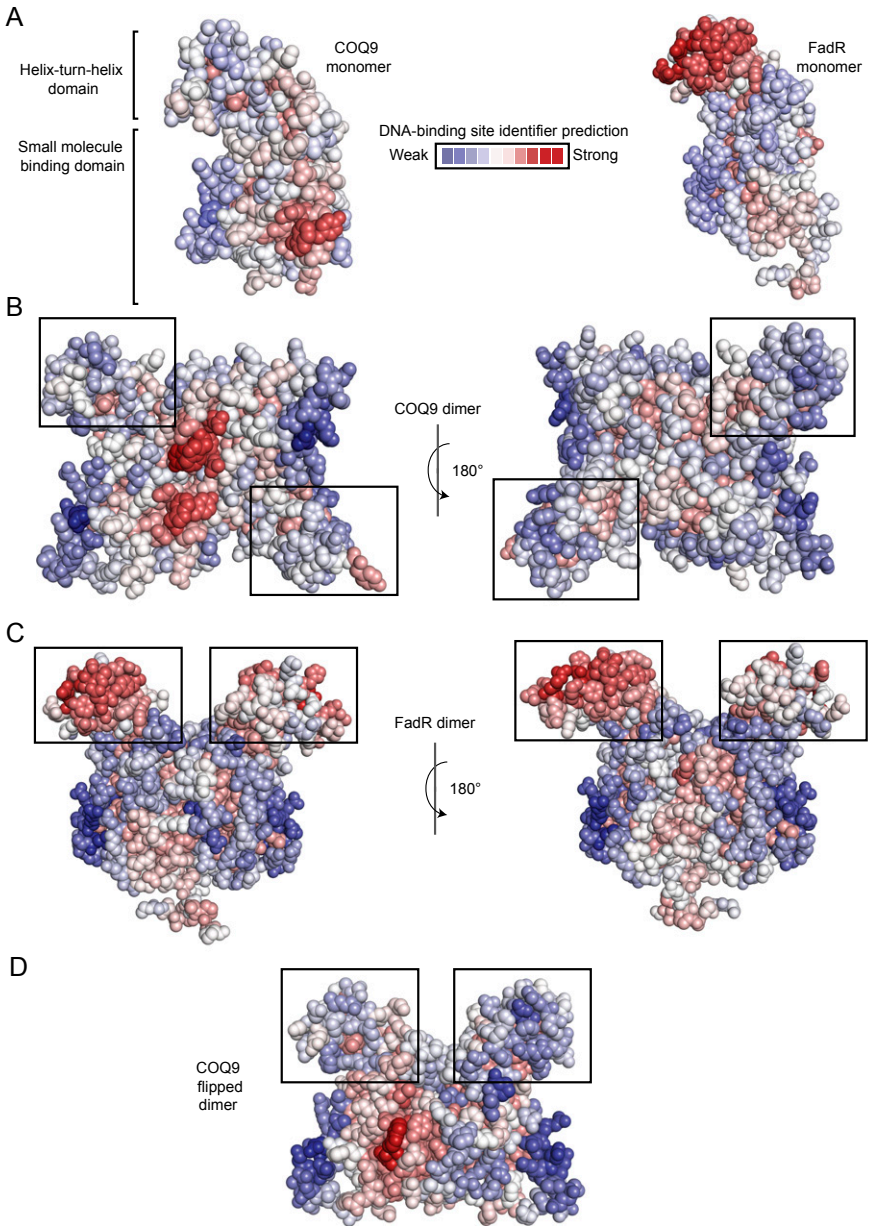


Fig. S3. COQ9 helix-turn-helix domain is not predicted to bind DNA (related to Fig. 4). Residue predictions made by DBSI (1) for DNA binding are shown and colored on a scale from a weak (blue) to strong (red) likelihood of interacting with DNA. (*A*) COQ9 and FadR monomers. (*B*) COQ9 dimer with helix-turn-helix motifs boxed. (*C*) FadR dimer (PDB: 3ANP) with helix-turn-helix motifs boxed. (*D*) The flipped COQ9 dimer by threading sequence through the FadR structure with helix-turn-helix motifs boxed.

1. Zhu X, Ericksen SS, Mitchell JC (2013) DBSI: DNA-binding site identifier. *Nucleic Acids Res* 41(16):e160.

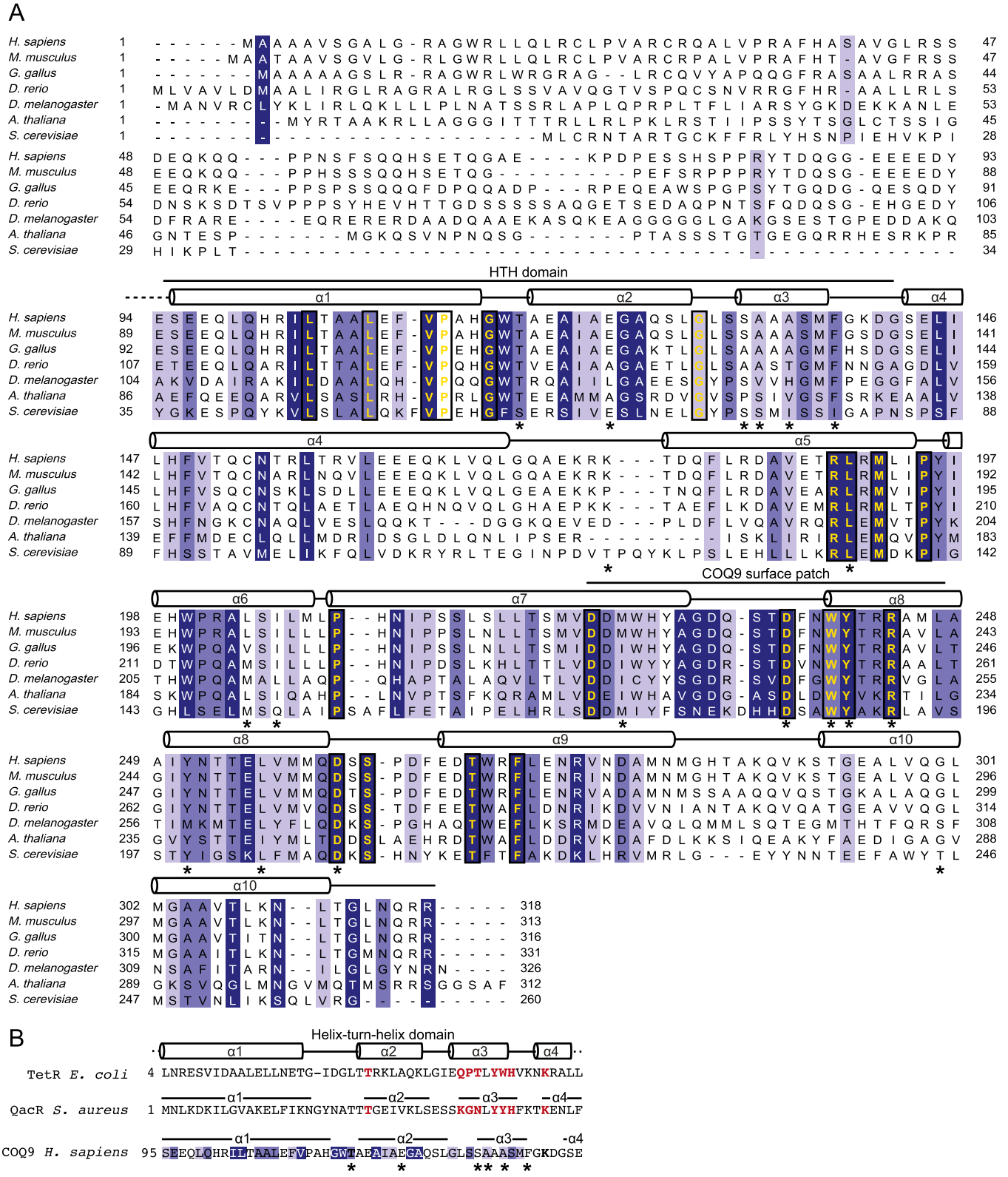


Fig. S4. COQ9 has distinct features from TFRs (related to Fig. 5). (*A*) ClustalW sequence alignment with highly conserved residues identified using the Consurf server (104 homologs) (1–3) and highlighted in increasing blue shades. Invariant residues are boxed and yellow, residues mutated in the yeast genetic screen are annotated (*), and the human COQ9 secondary structure is annotated. (*B*) Structure-based alignment of the helix-turn-helix regions of *Staphylococcus aureus* QacR and *E. coli* TetR with *Homo sapiens* COQ9 (4). Residues that interact with DNA are in red and the conservation coloring scheme for COQ9 is as described in *A*, with residues mutated in the yeast genetic screen annotated (*).

1. Berezin C, et al. (2004) ConSeq: The identification of functionally and structurally important residues in protein sequences. *Bioinformatics* 20(8):1322–1324.

2. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 38(web server issue):W529–W533.
3. Celniker G, et al. (2013) ConSurf: Using evolutionary data to raise testable hypotheses about protein function. *Isr J Chem* 53(3-4):199–206.
4. Ramos JL, et al. (2005) The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* 69(2):326–356.

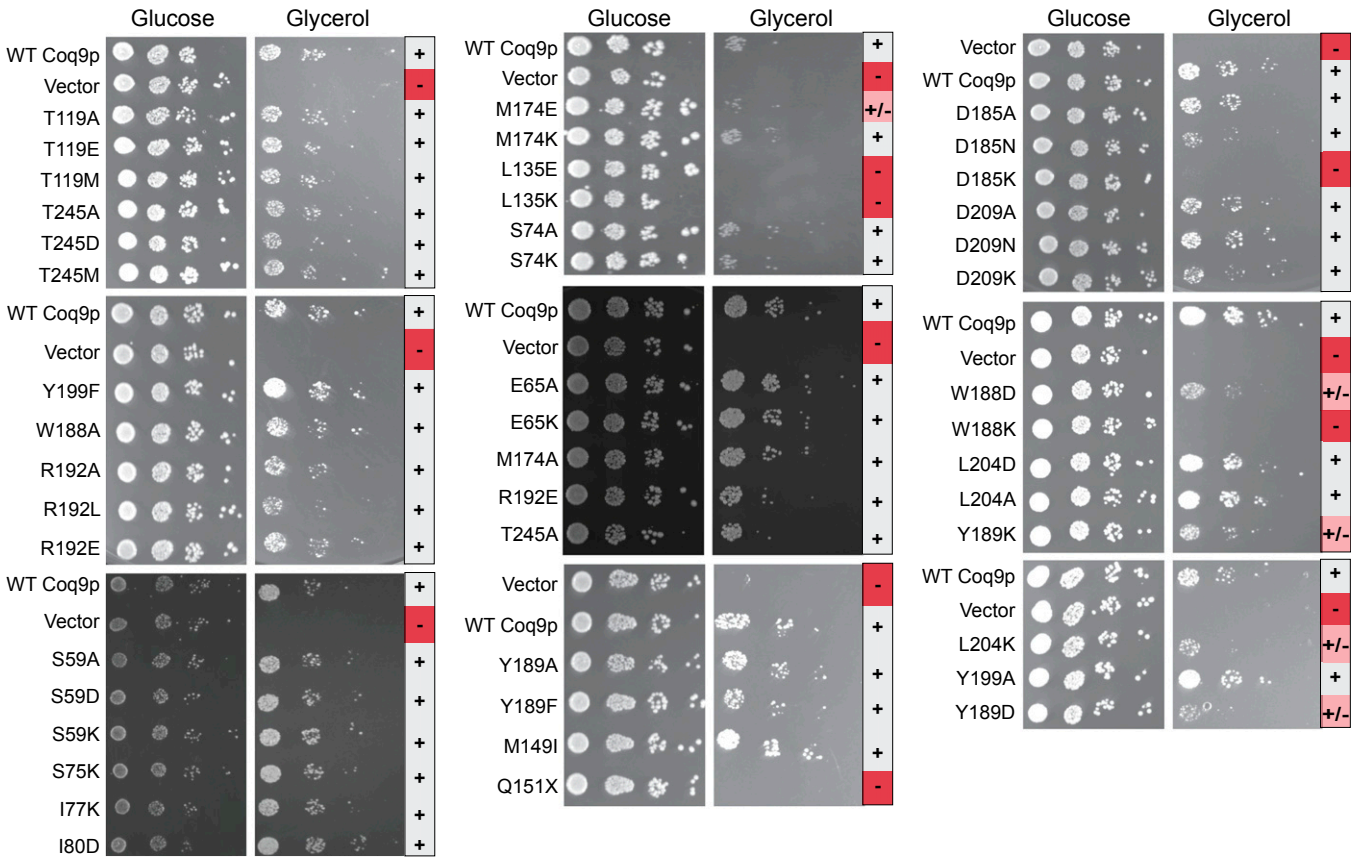


Fig. S5. Yeast genetic complementation assay identifies key Coq9p functional residues (related to Fig. 5). *Saccharomyces cerevisiae* genetic complementation assay with dilutions of *Δcoq9* strains harboring Coq9p point mutations grown on glucose- or glycerol-containing media. The glycerol growth phenotype is denoted as respiration competent (+), diminished competence (+/–), or incompetent (–).

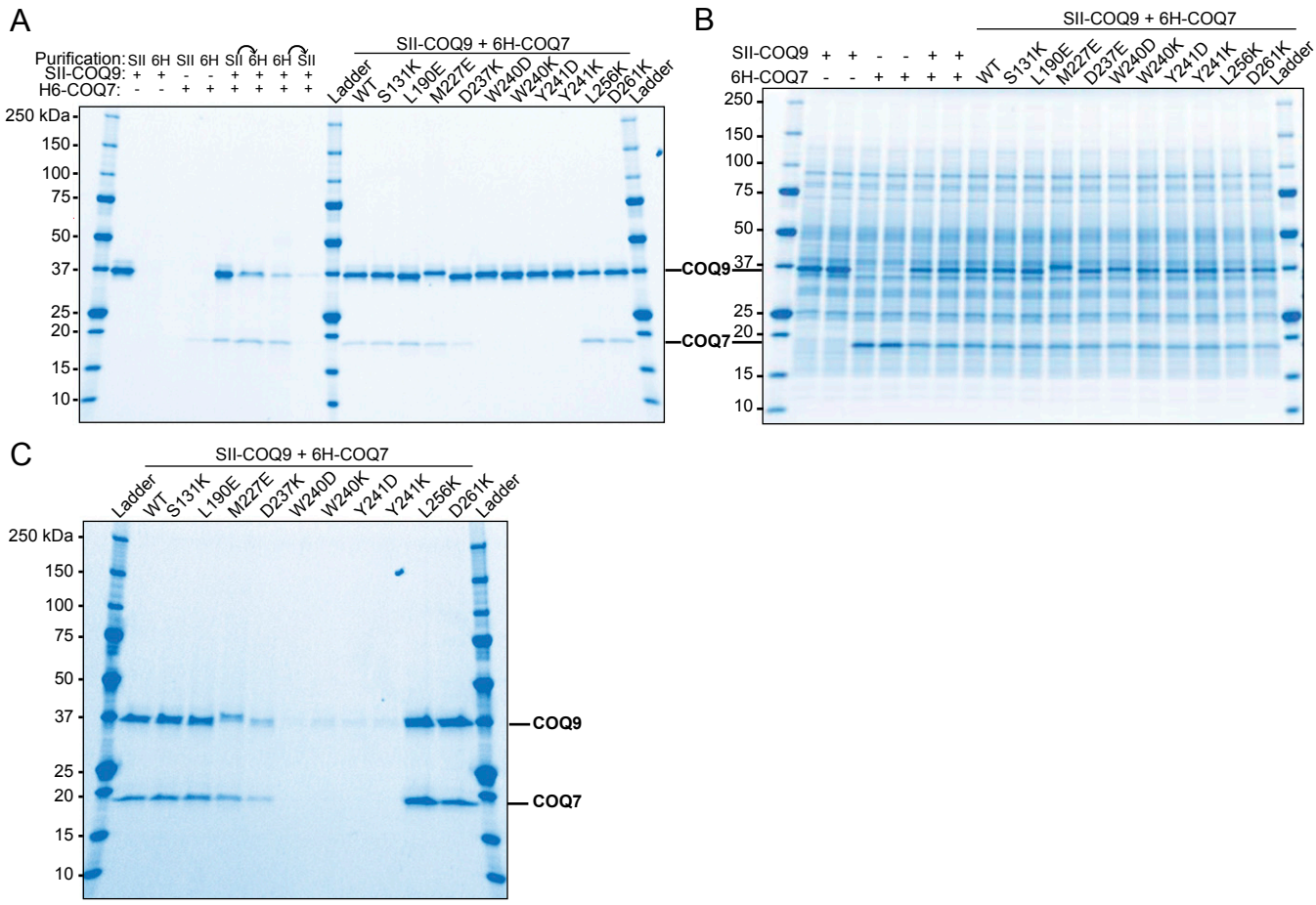


Fig. S6. Residues within the COQ9 surface patch are required for the COQ9–COQ7 interaction (related to Fig. 6). (*A*) Tagged SII-COQ9$^{N\Delta45}$ and H6-COQ7$^{N\Delta38}$ cotranslated in cell-free wheat germ extract (1), purified using the SII or 6H affinity tag, or purified sequentially by both tags as indicated by arrows. Point mutants of SII-COQ9$^{N\Delta45}$ were cotranslated with 6H-COQ7$^{N\Delta38}$ and purified using the SII-COQ9$^{N\Delta45}$ affinity tag. (*B*) Total translation before purification. (*C*) The elution from the SII purification purified using the 6H affinity tag on 6H-COQ7$^{N\Delta38}$.

1. Makino S, Beebe ET, Markley JL, Fox BG (2014) Cell-free protein synthesis for functional and structural studies. *Methods Mol Biol* 1091:161–178.

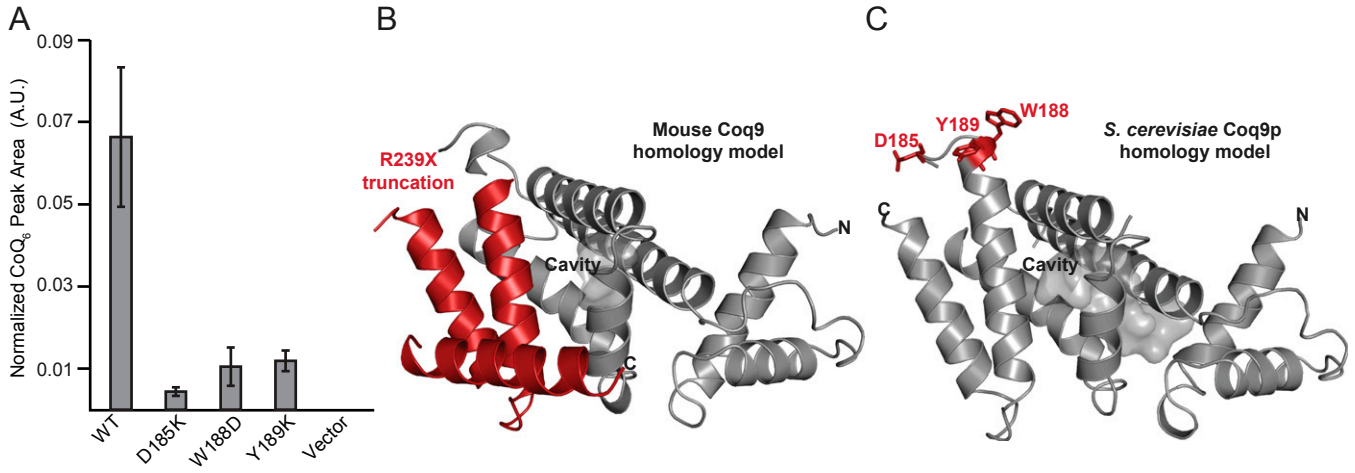


Fig. S7. COQ9 truncations and point mutants in mice and yeast (related to Fig. 6). (*A*) Coenzyme Q_6 levels normalized to added Q_{10} standard and detected by targeted LC-MS/MS in Δ*coq9 S. cerevisiae* harboring plasmid for mutant Coq9p with 95% confidence intervals indicated by error bars ($n = 6$). (*B*) Mouse Coq9 homology model with the Coq9^{R239X} truncated region shown in red. (*C*) Yeast Coq9p homology model with the Coq9p surface patch residues predicted to disrupt the Coq7p interaction shown in red.

Other Supporting Information Files

Dataset S1 (XLSX)