PNAS
www.pnas.org

Supplementary Information for

Deactivation blocks proton pathways in the mitochondrial complex I

Michael Röpke[1], Daniel Riepl[2,a], Patricia Saura[2,a], Andrea Di Luca[2,a], Max E. Mühlbauer[1,2], Alexander Jussupow[1,2], Ana P. Gamiz-Hernandez[2], Ville R. I. Kaila[1,2,*]

[1] Department Chemie, Technische Universität München (TUM), Lichtenbergstraße 4, D-85747 Garching, Germany.
[2] Department of Biochemistry and Biophysics, Stockholm University, 10691, Stockholm, Sweden.
[a] contributed equally to this work.

* Corresponding author: Ville R. I. Kaila
Email: ville.kaila@dbb.su.se

This PDF file includes:

SI Methods

FEP calculations

Alchemical free energy perturbation (FEP) calculations were used to study the hydration free energy of the gating region in both *active* and *deactive* states, based on the last snapshots of MD simulations S19 and S23 (Table S1), respectively. To enhance the sampling, the simulation system comprised subunits ND1, ND2 (residues 1-150), ND3, ND4L and ND6, and all water molecules within 3 Å of these subunits. Water molecules close to the cavity between $Asp66^{ND3}$ and $Glu34^{ND4L}$ were removed to prevent interference with the created/annihilated water wire, comprising five water molecules between $Asp66^{ND3}$ and $Glu34^{ND4L}$, as observed in the unbiased MD simulations. $Asp66^{ND3}$ was modeled protonated, mimicking a state in which the water wire would enable pT to $Glu34^{ND4L}$. The system was embedded in a POPC/POPE/cardiolipin lipid membrane, re-solvated in TIP3P water and neutralized with NaCl to a concentration of 150 mM by adding 135 Na^+ and 65 Cl^- ions. The FEP-model comprised *ca*. 120,000 atoms and was relaxed for 20 ns with initial harmonic constraints with a force constant of 3 kcal mol^{-1} $Å^{-2}$ acting on the position of the all protein atoms and water molecules. The harmonic constraints on the protein were thereafter removed, followed by equilibration and FEP simulations. External forces were applied to prevent new water molecules from entering into the protein interior. All simulation parameters were analogous to those used in the complete complex I system setup (see Main text methods), unless otherwise stated.

The FEP calculations were conducted by applying flat-bottom distance constraints between the introduced water molecules as well as $Asp66^{ND3}$ and $Glu34^{ND4L}$ using the Colvars (1) module of NAMD. Bulk water molecules were prevented from entering into the protein interior by using a harmonic restraint introduced on the water oxygen positions with a force constant of 0.02 kcal mol^{-1} $Å^{-2}$. No constraints were applied on the protein atoms during production simulations to allow for the structural relaxation of the cavity upon introduction of the water chain between $Asp66^{ND3}$ and $Glu34^{ND4L}$. Sampling was done in 20 equidistant windows from λ=0 (wire completely decoupled) to λ=1 (wire fully coupled) in forward and backward directions with 1 ns equilibration and 10 ns production simulations for each window, with a total sampling for 220 ns for each FEP simulation. The convergence was further tested using two 44 ns replicas with 0.2 ns equilibration and 2 ns production per window. Benchmarking simulations (0.5 ns per window) were conducted to probe the effect of splitting vdW and electrostatics and the number of windows used. The FEP simulations included in total 2.5 μs MD sampling. All FEP simulations yielded similar results when compared to the full-scale simulations indicating robust results, with a standard deviation of *ca*. 1 kcal mol^{-1} for the statistical error. Estimation of the accuracy of the non-polarizable CHARMM36 force field in prediction of hydration free energies is outside the scope of the present work. To probe the transfer free energy from bulk to the protein interior, five water molecules were removed/created in a 50 x 50 x 50 Å water box, with a 150 mM NaCl concentration (11 Na^+/Cl^-) using the same FEP-protocol as for the complex I models, but without addition of constraints. All FEP simulations were conducted using the NAMD-FEP implementation and analyzed with the ParseFEP plugin (2) in VMD employing the Bennet acceptance ratio (BAR) method (3).

MD simulations with a hydronium species in the water chain

Simulations with an explicit H_3O^+ species were started from different snapshots of the deactive state simulation S23 with $Asp66^{ND3}$ in a deprotonated state (see Table S1) and the H_3O^+ modeled next to the $Asp66^{ND3}$ carboxylate in the cavity connecting to $Glu34^{ND4L}$. These states were propagated for 110 ns and the water content of the cavity was subsequently evaluated in a cumulative manner. Classical force field parameters for the hydronium ion were derived at the B3LYP-D3/def2-TZVP level using Turbomole 7.5 (15), converting the molecular Hessian into force constants using Hess2FF (26), and applying the RESP procedure (27) for calculation of atomic point charges. Lennard-Jones parameters were taken from the TIP3P water parameters implemented in the CHARMM36 (28) force field. The full set of parameters is reported in SI Appendix, Table S9.

DFT benchmarking and reaction pathway calculations

The quantum chemical calculations were benchmarked at the DFT level using B3LYP-D3 (4,5,6), CAM-B3LYP-D3 (6,7), CAMh-B3LYP-D3 (6,8), ωB97XD (6,9), TPSSh-D3 (6,10), and by using the correlated *ab initio* random phase approximation (RPA) (11) and domain-based local pair natural orbital coupled cluster with single-double and perturbative triples (DLPNO-CCSD(T)) (12). def2-TZVP basis sets (13) were used for the DFT calculations and aug-cc-pVTZ (14) basis sets at the correlated levels. Single point energy calculations were performed for model systems comprising 27, 131, and 381 atoms, constructed from DFT models 1-3 (see Table S4), and involving proton transfer reactions between carboxylates ($Asp66^{ND3}$, $Glu34^{ND4L}$, $Glu70^{ND4L}$) and water molecules. The DFT calculations of the large model systems were performed using an implicit polarizable medium with ε=4, whereas the calculations for small and intermediate models were done in gas phase. The benchmarking calculations were performed using TURBOMOLE v7.4-7.5 (15) for the DFT and RPA calculations, and ORCA v4.2 (16) for the DLPNO-CCSD(T) calculations. See Table S7 for the quantum chemical benchmarking results.

Reaction pathways of the proton transfer reactions in the DFT models were optimized using a minimum energy reaction pathway optimization approach (29) that resembles the zero-temperature string method. To this end, the initial guess of the reaction pathway was constructed by optimization of the structures with the proton on the donor and acceptor states, followed by interpolating 19 intermediate structures between the two states using the linear synchronous transit method. The reaction pathway was refined by constraining equally spaced intermediate structures with a quadratic potential (29) until the root-mean-square difference of the gradient varied less than 10^{-4} au in subsequent optimization steps. Numerical estimation of the Hessian showed one imaginary frequency (> 500 cm^{-1}) for the approximate transition state structure. See Ref. (29) for further methodological details of the reaction pathway optimization approach. The overall reaction process predicted by this approximate reaction pathway method resembles that obtained from our QM/MM umbrella sampling calculations, along the geometric reaction coordinate *R* (see methods).

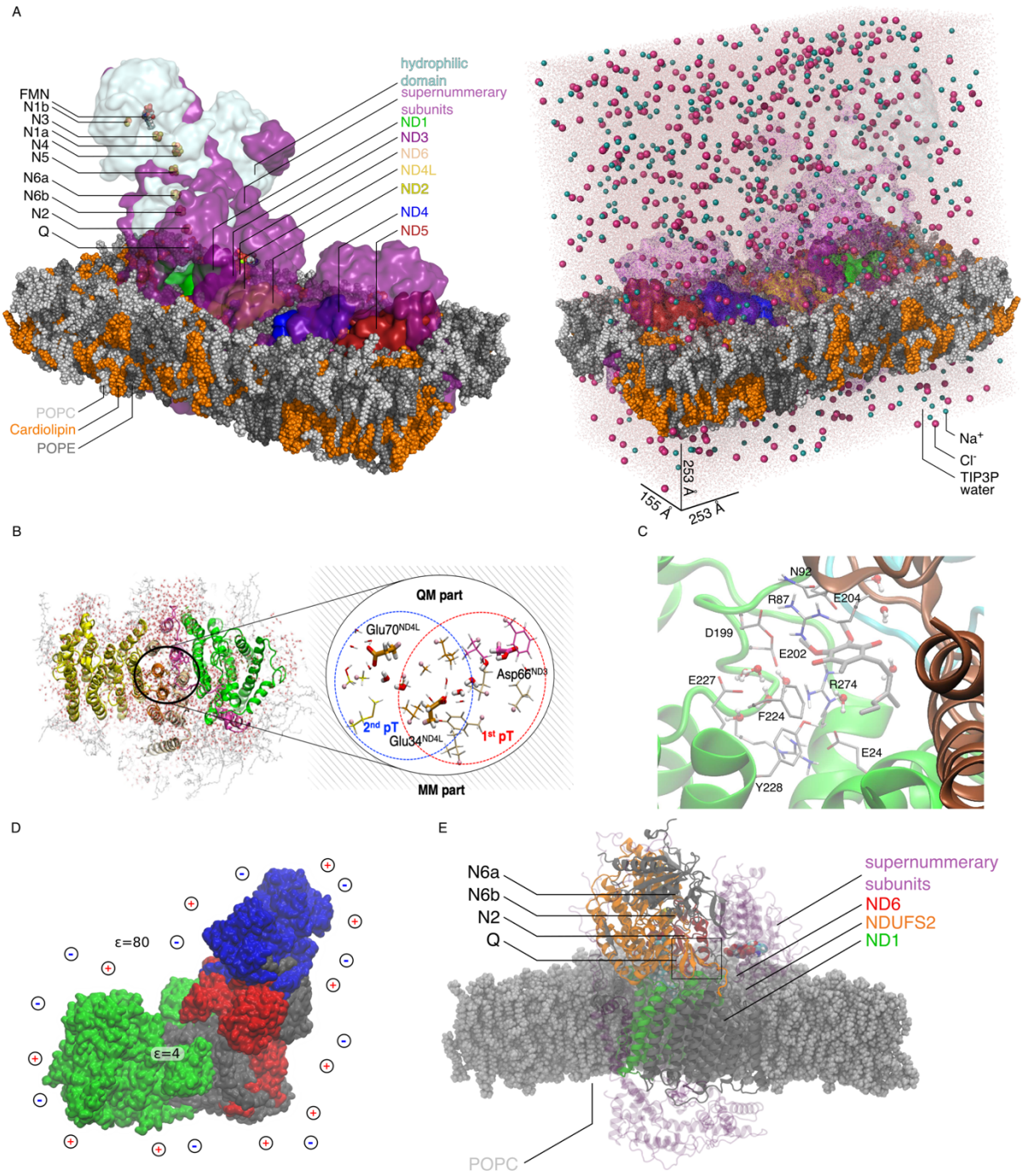


Fig. S1. Molecular models. A, MD simulation setup of the mitochondrial complex I. **B**, QM and QM/MM models for probing proton transfer at the ND3/ND4L/ND6/ND2 interface. **C**, QM region used in QM/MM MD simulations of binding site 2, with residues shown in licorice, modeled at the DFT level (see *Methods*). **D**, models used in PB/MC-electrostatic calculations. **E**, models used for exploration of the Q_{10} dynamics with atomistic- and coarse-grained MD.

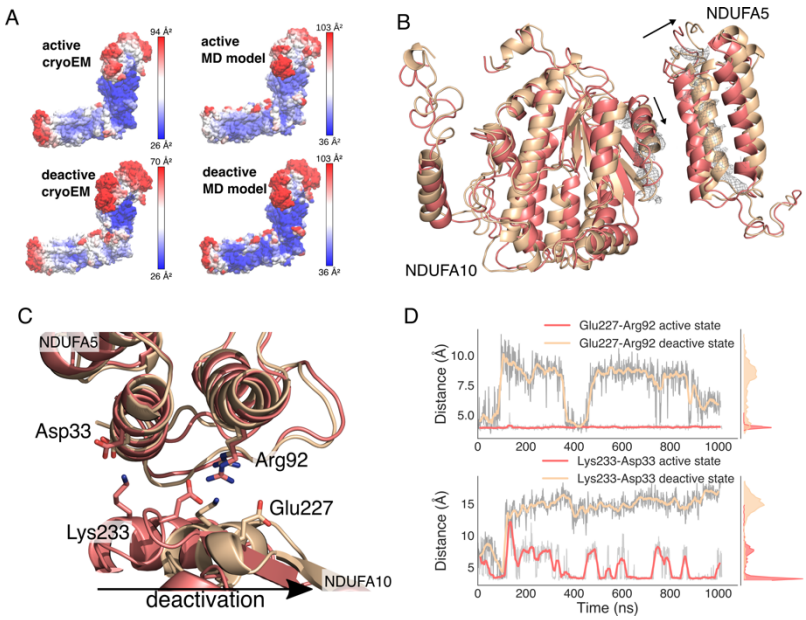


Fig. S2. Conformational changes linked to complex I deactivation. A, Comparison of dynamics inferred from MD simulations and B-factors obtained from cryoEM data. **B,** Conformational changes around the NDUFA10/5 subunit linked to the A/D transition in MD simulations. **C**, **D**, ion pairs at the NDUFA10/5 interface undergo conformational rearrangement during the MD simulations of the respective states.

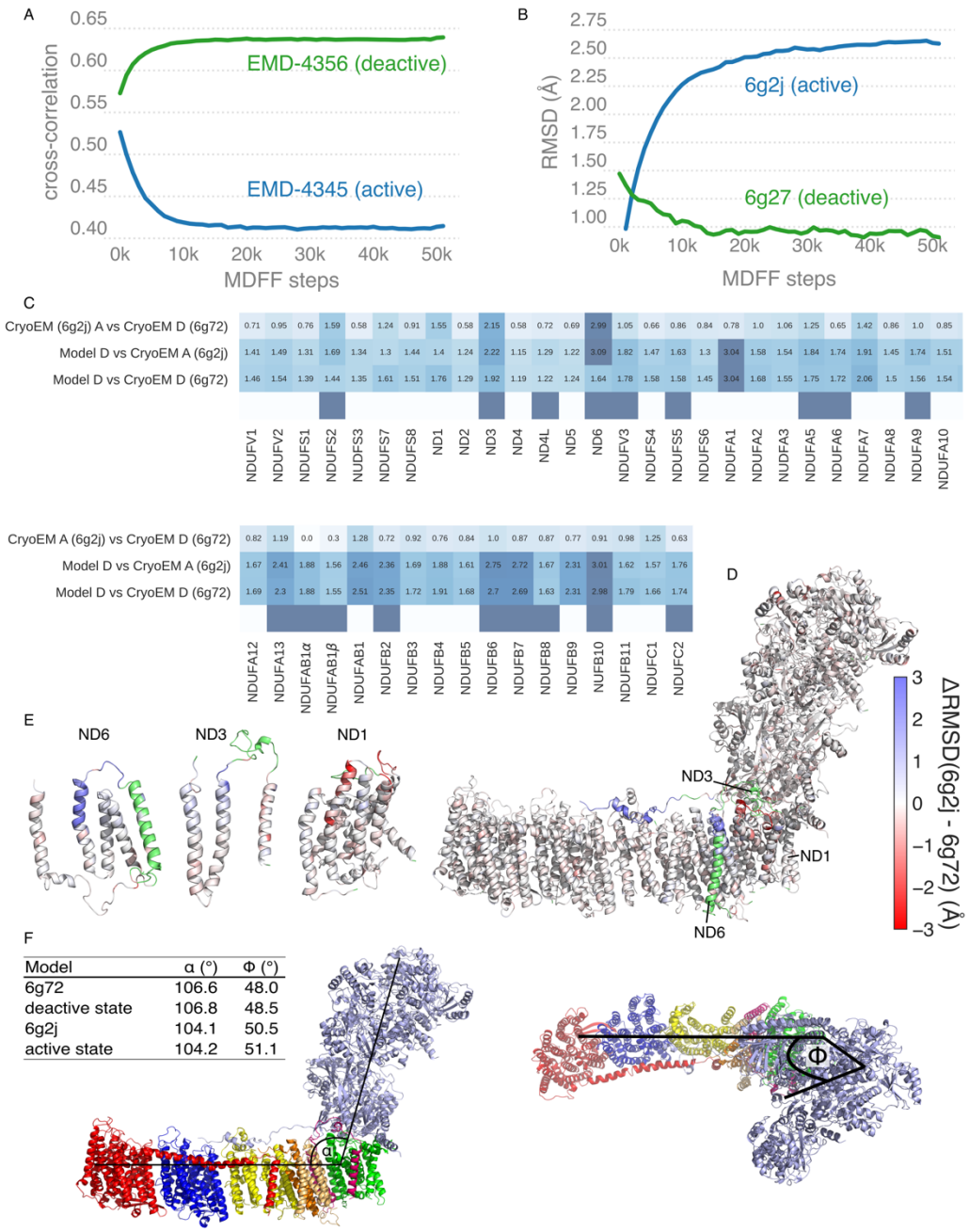


Fig. S3. Structural characterization of the deactive model. A, Cross-correlation between calculated density during the MDFF simulations and the cryoEM density map of the *active* (EMDB: 4345) and *deactive* state (EMDB: 4356). **B,** The *root-mean-square-deviation* (RMSD) of the *deactive* state model during MDFF relative to experimentally refined structures of the D (PDB ID: 6g72) and A (PDB ID: 6g2j) states. The RMSD was calculated on a backbone level for the resolved regions. **C,** Subunit-wise RMSD relative of the D state model relative to cryoEM-refined models of the D (PDB ID: 6g72) and A (PDB ID: 6g2j) states. **D** and **E,** RMSD difference between the D state model to the cryoEM-refined models of the D (PDB ID: 6g72) and A (PDB ID: 6g2j) states. Unobserved regions in either cryoEM models are colored green. **F,** Bending angles (α) and dihedral angles (ϕ) between the hydrophilic and membrane domains in the *active* and *deactive* states. The angles were determined using the principal axes of a subset of specific subunits for the respective domains (hydrophilic: NDUFV1, NDUFV2, ND1; membrane: ND5, ND4, ND2).

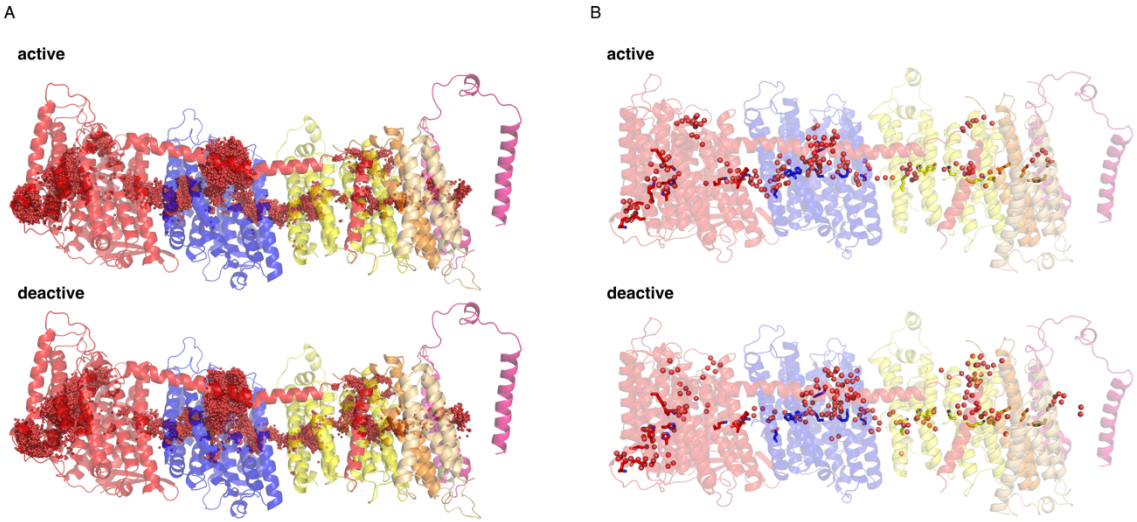


Fig. S4. Comparison of proton pathways in the *active* and *deactive* state MD simulations. A, Water dynamics (red spheres) shown as an ensemble average of the 2 μs MD simulations of the *active* state (*top*, simulations S1+S2) and *deactive* state (*bottom*, simulations S6+S7). **B**, Snapshot of individual buried water molecules (red spheres) after 1 μs MD simulation (*top: active* state simulations S1+S2; *bottom*: *deactive* state, simulations S6+S7).

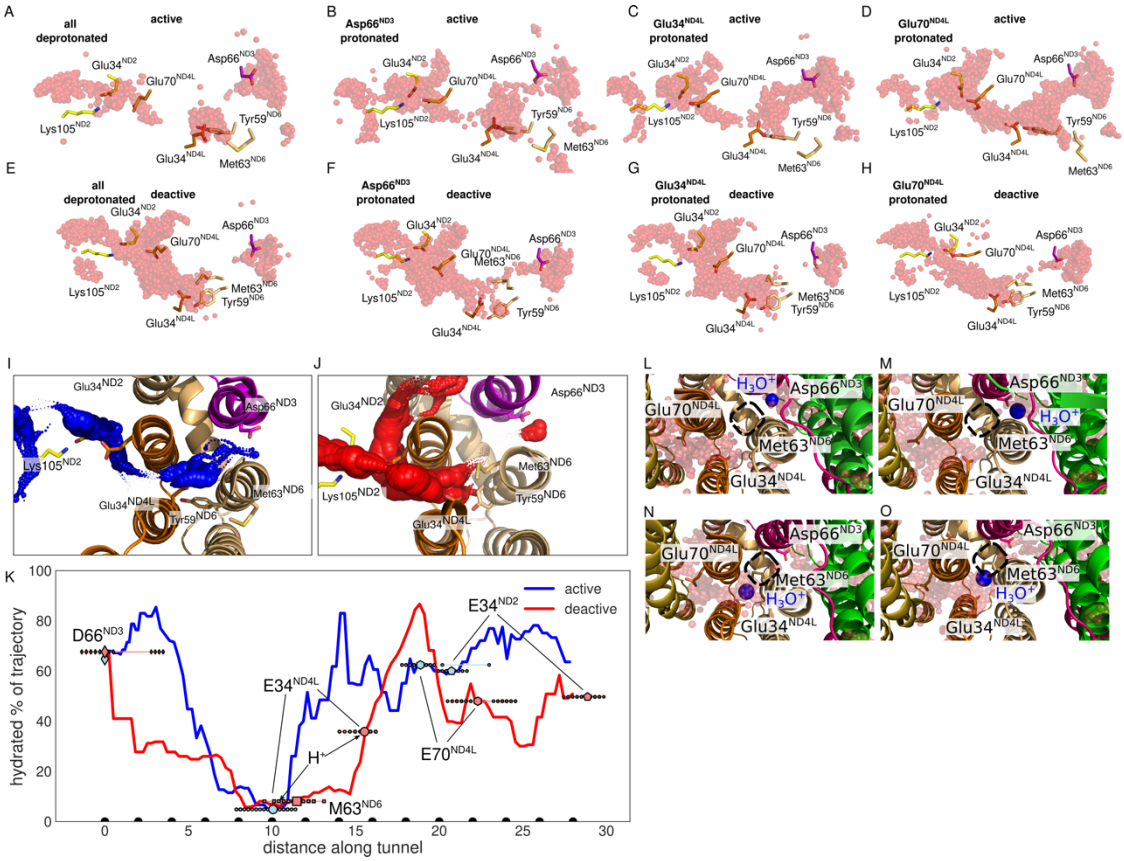


Fig. S5. MD simulations exploring proton transfer along the ND3/ND4L/ND6 region. The simulations show the effect of modeling different protonation states in Asp66^{ND3}, Glu34^{ND4L}/Tyr59^{ND6}, and Glu70^{ND4L}, mimicking proton transfer via the region in the *active* (left, A-D) and *deactive* (right, F-I) state simulations. **E**, **J**, Hydration average of the tunnels from MD simulations in the *active* (red) and *deactive* (blue) states with Asp66^{ND3} protonated. The shown diameter is proportional to the water content (see *Methods*). **K**, Hydration fraction along the tunnels connecting the acidic residues along the ND3/ND4L/ND6 gating region when Glu34^{ND4L} is protonated. The tunnel distances correspond to the beads shown in panel K. Functional residues along the tunnel are indicated with differently shaped markers (diamond: D66^{ND3}, circle: E34^{ND4L}, hexagon: E70^{ND4L}, pentagon: E34^{ND2}, (deactive only) square: M63^{ND6}). The mean position is indicated with a larger size marker. **L-O**, Hydration dynamics during 110 ns trajectories with a classically modelled H_3O^+ species, placed next to Asp66^{ND3} (panels L, M) or Glu34^{ND4L} (panels N, O) in the deactive state. The cavity, marked with a black frame, remains dry on the simulation timescale in the studied states, suggesting that the thermodynamic cost of hydrating this region could be high in the deactive state.

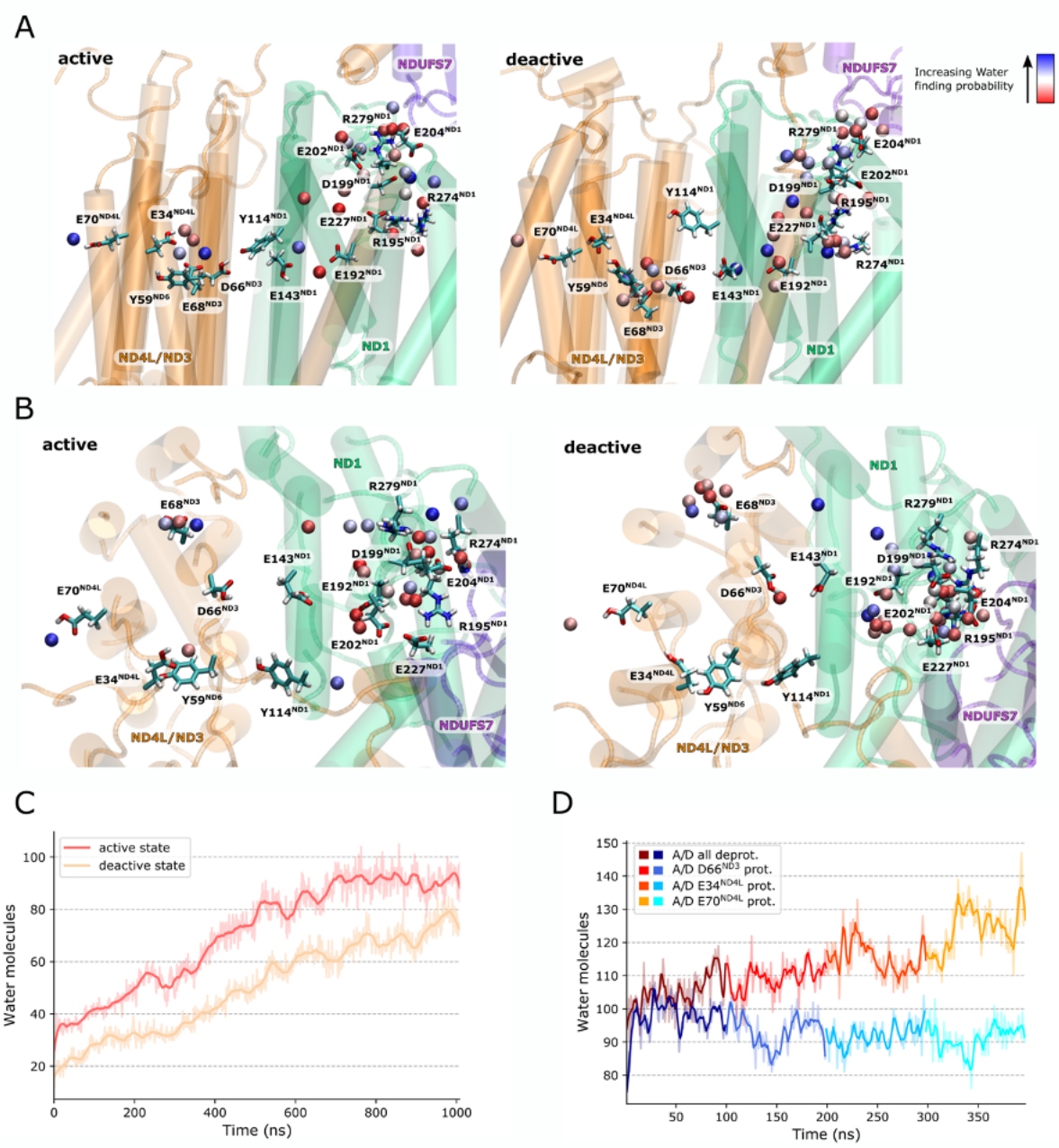


Fig. S6. Water cluster analysis and hydration dynamics. A, B, Top/side view of water cluster analysis for the ND1/ND3/ND4L/ND6 region obtained based on 1000 ns MD simulations of the *active* and *deactive* states (simulations S1 and S6, Table S1). The water shown are high probability water molecules identified in the cluster analysis (25). **C,** Hydration dynamics in the ND1/ND3/ND4L/ND6 region during *active* and *deactive* state MD simulations (simulations S1 and S6, Table S1). **D,** Hydration dynamics in the ND1/ND3/ND4L/ND6 region induced by shifting the proton during consecutive 100 ns steps along the Asp66^{ND3}, Glu34^{ND4L}, Glu70^{ND4L} (simulations S18-S21, Table S1) The simulations were initiated from 1000 ns frame shown in panel C.

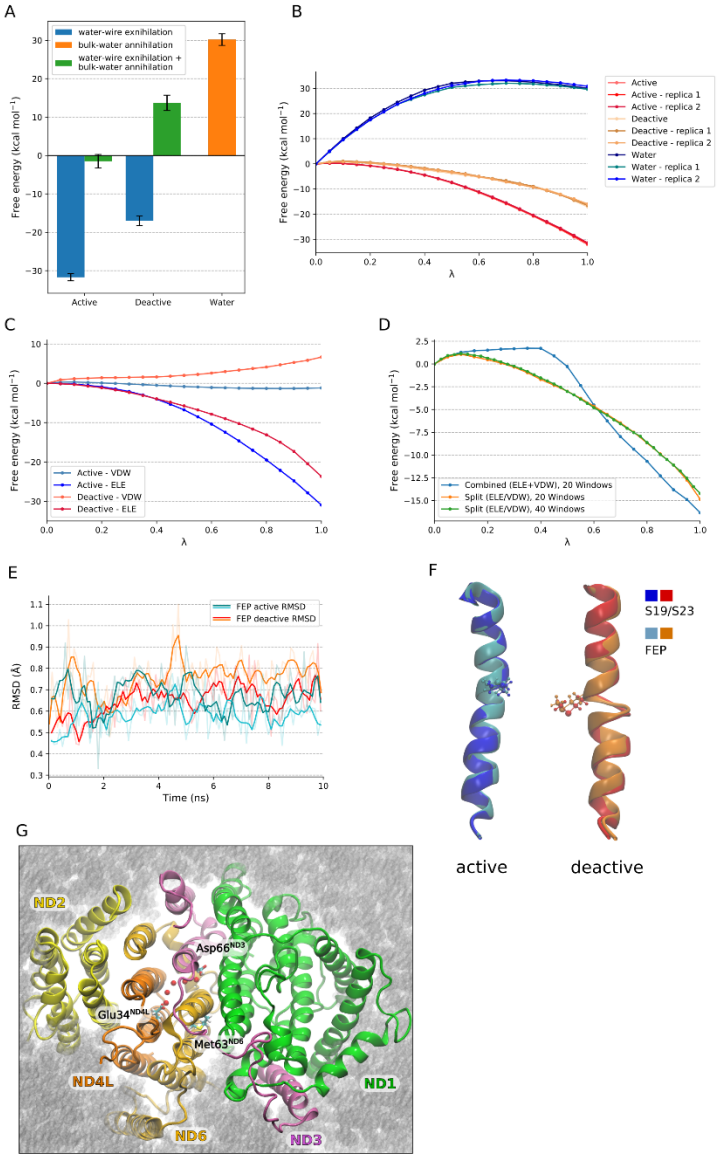


Fig. S7. Free energy perturbation (FEP) calculations. **A**, Free energy for introduction of a water wire (*exnihilation*) in the ND3/ND4L/ND6 region in the *active* and *deactive* states, and water removal (*annihilation*) from the bulk, using alchemical FEP. **B**, Free energy profiles from λ=0 (fully decoupled) to λ=1 (fully coupled) for all replicas. Note that intermediate λ values represent unphysical states. **C**, Decomposition of the free energy into vdW and electrostatic contributions. **D**, Benchmarking simulations to probe the effect of splitting the vdW and electrostatic coupling/decoupling into separate simulations and the dependence on number of sampling windows. The benchmarking calculations were performed on the *deactive* state model. **E**, Backbone RMSD of TM3-ND6 for the *active* and *deactive* FEP simulations (λ=1). The RMSD of each simulation is relative to the first frame and calculated by aligning the whole ND6 subunit, indicating small local rearrangement within the $TM3^{ND6}$ region. **F**, Comparison of the last snapshot of the *active* and *deactive* state FEP simulations (pale colors) with the last snapshots from simulations S19 and S23 (dark colors), respectively. $TM3^{ND6}$ shows small conformational changes, whereas larger conformational changes could be limited by the FEP simulations timescale.

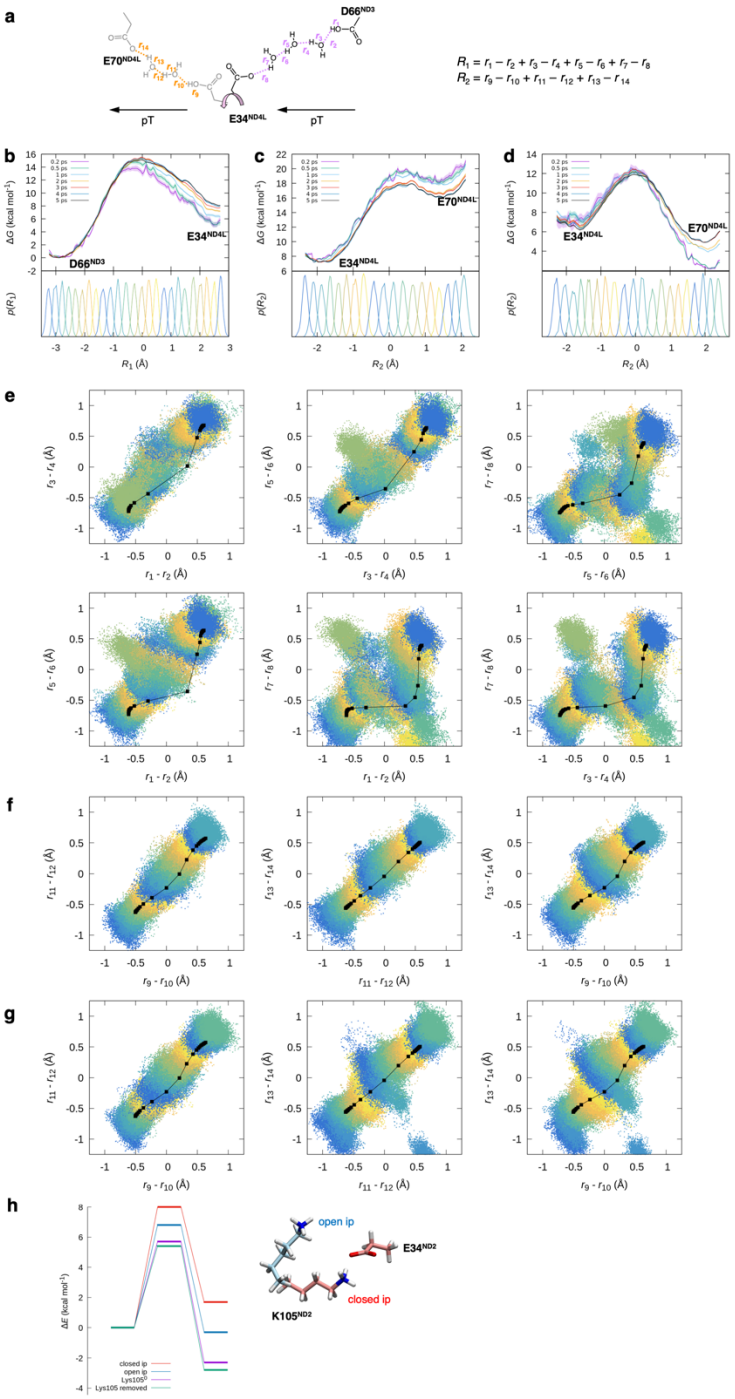


Fig. S8. Convergence of hybrid QM/MM free energy calculations and sampling of reaction coordinates. A, definition of reaction coordinates R_1 and R_2 used in the QM/MM umbrella sampling (US) simulations. **B, C, D,** convergence of QM/MM US calculations and overlap of reaction coordinates in the individual simulations. **E, F, G,** Sampled proton transfer distances in the QM/MM US simulations during the respective simulations shown in **B-D,** suggesting that the proton is transferred in a Grotthuss-type transfer mechanism (sampling along the diagonal). Results from DFT reaction pathway optimizations are shown as black dots. **H,** effect of modeling ND2 ion-pair opening in the DFT cluster models by conformational and/or protonation changes.

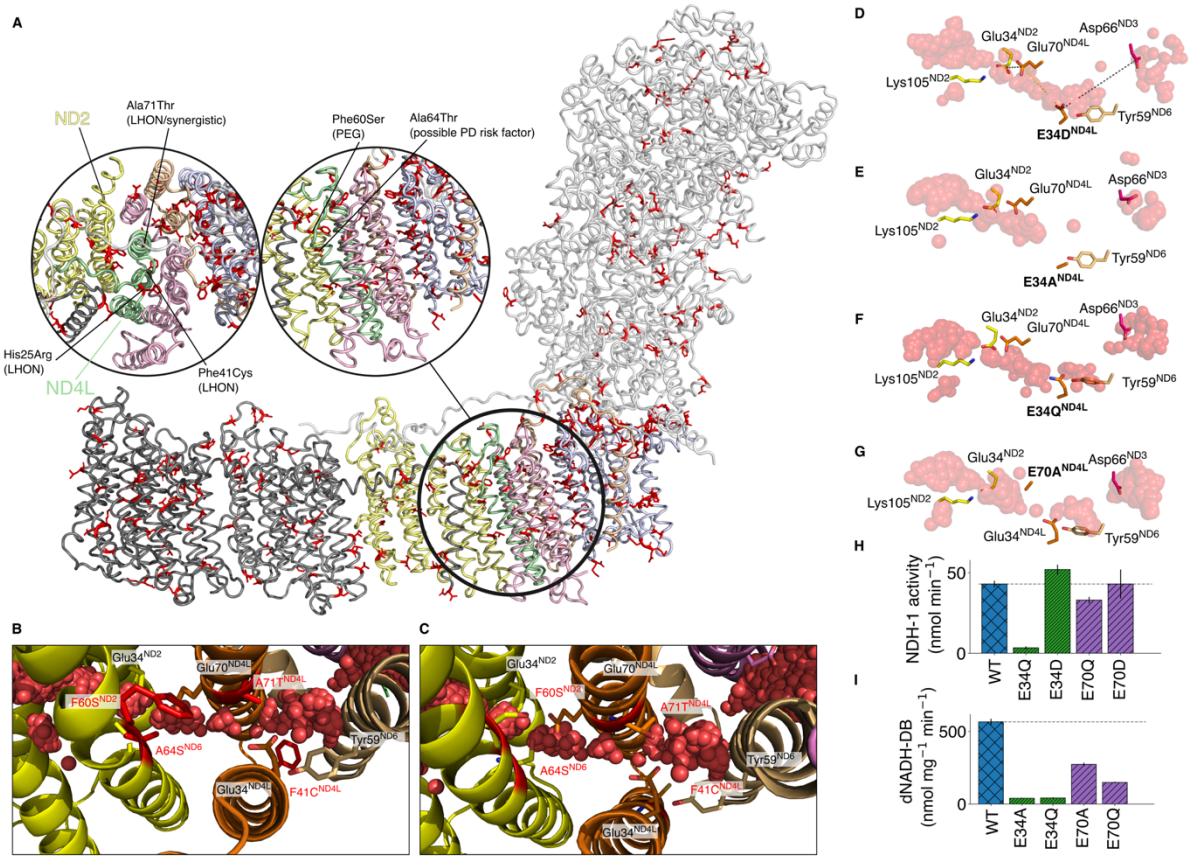


Fig. S9. Disease related mutations in the mitochondrial complex I. A, location of mitochondrial disease related mutations (residues marked in red) (22). **B,** Closeup of disease related mutations near the proton pathway at the ND3/ND4L/ND6 interface, and **C**, *in silico* models of the mutations. Introduced substitutions form contacts with water molecules or proton donor/acceptor groups within the wire. F41C^{ND4L} could lead to a formation of a disulfide bridge with Cys40^{ND6}, the backbone of which is located <10 Å in the wild type structure. See also Table S6. **D-G**, introduced single point mutations: **D**, E34D^{ND4L}, **E**, E34A^{ND4L}, **F**, E34Q^{ND4L}, **G**, E70A^{ND4L} perturb the proton pathways at the ND6/ND4L interface. **H, I**, effect of single point mutations of key residues on complex I activity. The experimental data was obtained from Refs. (23,24).

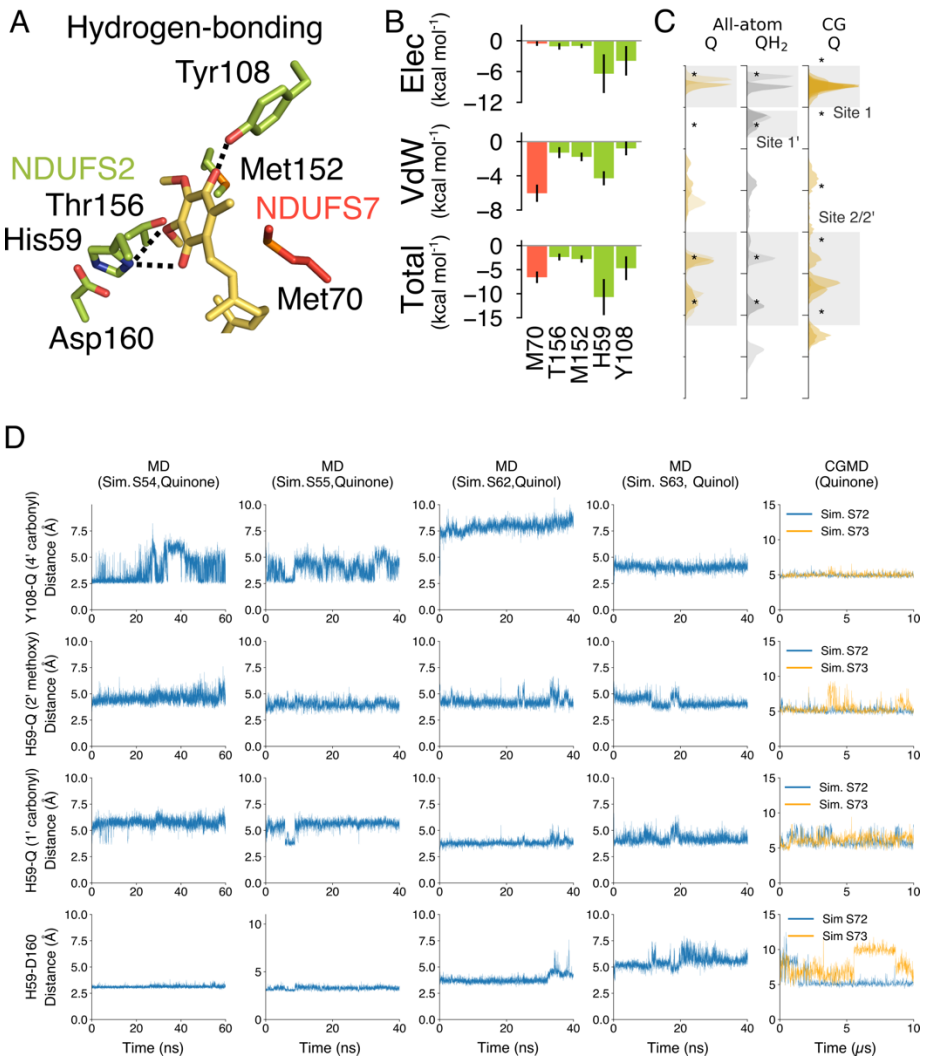


Fig. S10. Q_{10} dynamics explored by atomistic and coarse-grained MD. A, hydrogen-bonding binding mode, with the Q headgroup forming contacts with $Tyr108^{NDUFS2}$ and $His59^{NDUFS2}$. **B**, interaction energy decomposition of the binding energy contributions (in kcal mol^{-1}). **C**, Distribution of Q_{10}-$Tyr108^{NDUFS2}$ distances explored from classical (simulations S62-S77) and coarse-grained MD simulations (simulations S78-S87). The Q starting positions in the simulations are indicated with an asterisk. **D**, Dynamics of Q_{10} explored in independent MD (in blue, simulations S62, S53, S70, S71) and CGMD (simulations S78, S79) trajectories.

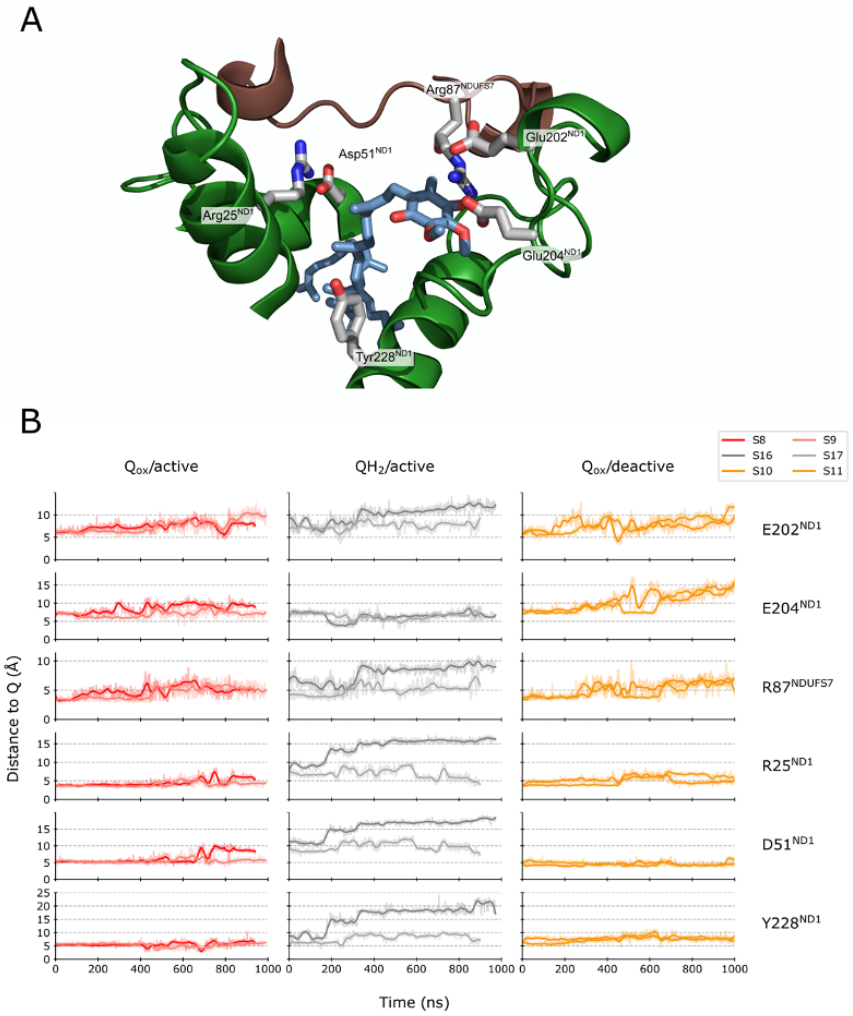


Fig. S11. Dynamics of Q_{10} in the membrane-bound binding region. **A,** Selection of key residues that form contact with Q_{10} during the *active*/*deactive* state simulations **B,** Distance plot between Q/QH_2 and the selected residues. The distances were measured from Glu(CD)/Asp(CG)/Arg(CZ)/Tyr(OH) to the closest carbonyl oxygen of Q_{10}.

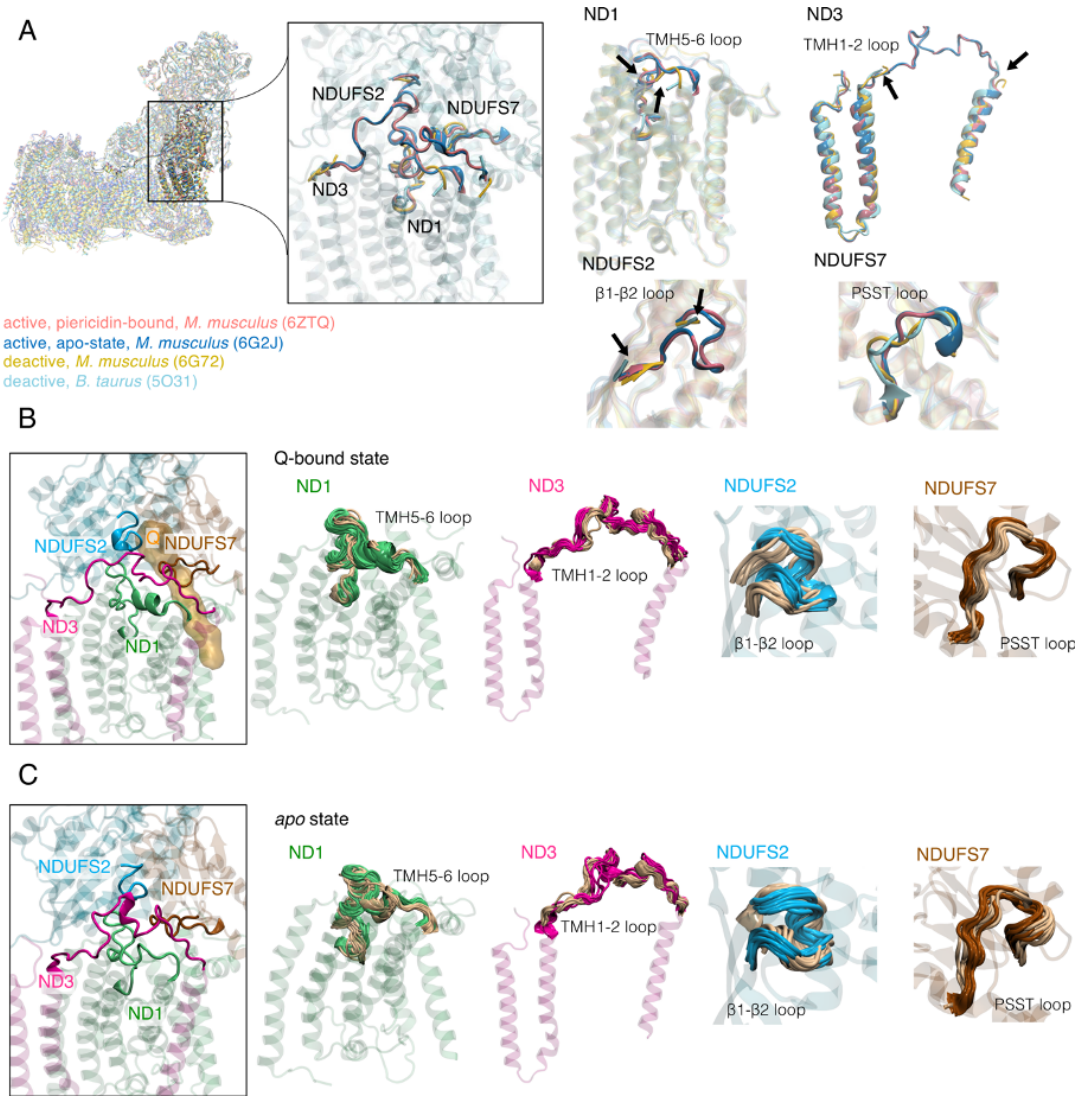


Fig. S12. Conformational changes around the Q tunnel linked to the A/D transition. A, unresolved loop regions/structural differences in cryoEM structures around ND1, ND3, NDUFS2, and NDUFS7 of the mammalian complex I (marked with arrows). **B, C,** Conformational dynamics around ND1, ND3, NDUFS2, and NDUFS7 in the *active* (subunit specific color) and the *deactive* (light brown) states MD simulations, with **B**, showing the Q-bound simulations (simulation S1/S6, Table S1), and **C,** showing the *apo*-state simulations (simulation S12, S14, Table S1).

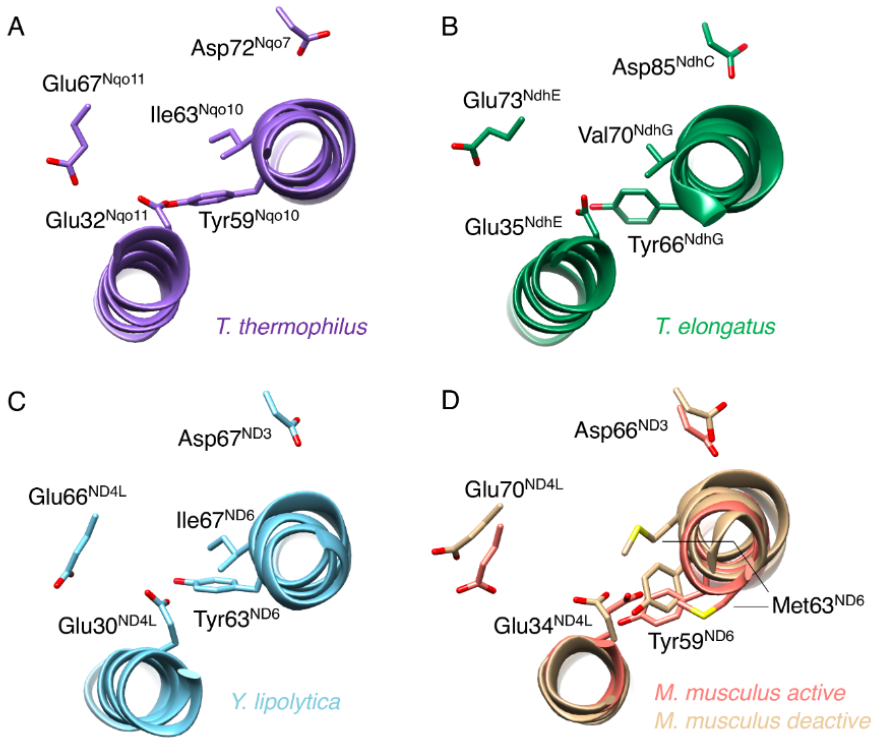


Fig. S13. Conformational changes around TM3$^{ND3/ND6}$ and the ND3/ND4L/ND6 gating region in **A**, *Thermus thermophilus* (purple, Nqo11/Nqo10, PDB ID: 4HEA(17)), **B**, *Thermosynechococcus elongatus* (green, NdhE/NdhG, PDB ID: 6TJV (18)), **C**, *Yarrowia lipolytica* (blue, ND4L/ND6, PDB ID: 6RFR (19)), and **D**, mouse complex I *active* (red) and *deactive* (pale) models derived in this study.

Table S1. Overview of classical MD simulations.

Simulation	State	Length	Starting point	Comment
S1	active	1000 ns	CryoEM after equilibration	Q site 1
S2	active	500 ns	CryoEM after equilibration	Q site 1
S3	-	200 steps	CryoEM structure active + map deactive	MDFF-pre minimization
S4	-	50,000 steps	S3	MDFF run
S5	deactive	2000 steps	S4	MDFF post-minimization
S6	deactive	1000 ns	S5	Q site 1
S7	deactive	670 ns	S5	Q site 1
S8	active	1000 ns	CryoEM after equilibration	Q site 2
S9	active	1000 ns	CryoEM after equilibration	Q site 2
S10	deactive	1000ns	S5	Q site 2
S11	deactive	1000ns	S5	Q site 2
S12	active	1000 ns	CryoEM after equilibration	*apo*-state
S13	active	1000 ns	CryoEM after equilibration	*apo*-state
S14	deactive	1000 ns	S5	*apo*-state
S15	deactive	1000 ns	S5	*apo*-state
S16	active	1000 ns	CryoEM after equilibration	QH_2 site 2
S17	active	1000 ns	CryoEM after equilibration	QH_2 site 2

Simulation	State	Length/Steps	Starting point	Comment $D66^{ND3}/E34^{ND4L}/E70^{ND4L}$
S18	active	100 ns	S1	DDD
S19	active	100 ns	S18	PDD
S20	active	100 ns	S19	DPD
S21	active	100 ns	S20	DDP
S22	deactive	100 ns	S6	DDD
S23	deactive	100 ns	S22	PDD
S24	deactive	100 ns	S23	DPD
S25	deactive	100 ns	S24	DDP

Simulation	State	Length/Steps	Starting point	Comment
S26	active	10 ns	S1 after 962 ns	A64S
S27	active	10 ns	S1 after 962 ns	A71T
S28	active	10 ns	S1 after 962 ns	E34A
S29	active	10 ns	S1 after 962 ns	E34D
S30	active	10 ns	S1 after 962 ns	E34Q
S31	active	10 ns	S1 after 962 ns	E70A
S32	active	10 ns	S1 after 962 ns	E70Q
S33	active	10 ns	S1 after 962 ns	F41C
S34	active	10 ns	S1 after 962 ns	F60S
S35	active	10 ns	S1 after 990 ns	A64S
S36	active	10 ns	S1 after 990 ns	A71T
S37	active	10 ns	S1 after 990 ns	E34A
S38	active	10 ns	S1 after 990 ns	E34D
S39	active	10 ns	S1 after 990 ns	E34Q
S40	active	10 ns	S1 after 990 ns	E70A
S41	active	10 ns	S1 after 990 ns	E70Q
S42	active	10 ns	S1 after 990 ns	F41C
S43	active	10 ns	S1 after 990 ns	F60S
S44	deactive	10 ns	S6 after 746 ns	A64S
S45	deactive	10 ns	S6 after 746 ns	A71T
S46	deactive	10 ns	S6 after 746 ns	E34A
S47	deactive	10 ns	S6 after 746 ns	E34D
S48	deactive	10 ns	S6 after 746 ns	E34Q
S49	deactive	10 ns	S6 after 746 ns	E70A
S50	deactive	10 ns	S6 after 746 ns	E70Q
S51	deactive	10 ns	S6 after 746 ns	F41C
S52	deactive	10 ns	S6 after 746 ns	F60S
S53	deactive	10 ns	S6 after 981 ns	A64S
S54	deactive	10 ns	S6 after 981 ns	A71T
S55	deactive	10 ns	S6 after 981 ns	E34A
S56	deactive	10 ns	S6 after 981 ns	E34D

S57	deactive	10 ns	S6 after 981 ns	E34Q
S58	deactive	10 ns	S6 after 981 ns	E70A
S59	deactive	10 ns	S6 after 981 ns	E70Q
S60	deactive	10 ns	S6 after 981 ns	F41C
S61	deactive	10 ns	S6 after 981 ns	F60S

Simulation	**State**	**Length/Steps**	**Starting point**	**Comment**
S62	Q/active	60 ns	Resolved piericidin model	Q site 1
S63	Q/active	40 ns	Resolved piericidin model	Q site 1
S64	Q/active	40 ns	Docked Q based on Ref. 5	Q site 1’
S65	Q/active	40 ns	Docked Q based on Ref. 5	Q site 1’
S66	Q/active	40 ns	Docked Q based on Ref. 5	Q site 2
S67	Q/active	40 ns	Docked Q based on Ref. 5	Q site 2
S68	Q/active	40 ns	Docked Q based on Ref. 5	Q site 2’
S69	Q/active	40 ns	Docked Q based on Ref. 5	Q site 2’
S70	QH_2/active	40 ns	Docked Q based on Ref. 5	Q site 1
S71	QH_2/active	40 ns	Docked Q based on Ref. 5	Q site 1
S72	QH_2/active	40 ns	Docked Q based on Ref. 5	Q site 1’
S73	QH_2/active	40 ns	Docked Q based on Ref. 5	Q site 1’
S74	QH_2/active	40 ns	Docked Q based on Ref. 5	Q site 2
S75	QH_2/active	40 ns	Docked Q based on Ref. 5	Q site 2
S76	QH_2/active	40 ns	Docked Q based on Ref. 5	Q site 2’
S77	QH_2/active	40 ns	Docked Q based on Ref. 5	Q site 2’
S78	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 1
S79	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 1
S80	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 1’
S81	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 1’
S82	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 2
S83	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 2
S84	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 2’
S85	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 2’
S86	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 2’/out
S87	Q/active CGMD	10 000 ns	Docked Q based on Ref. 5	Q site 2’/out
	Total:	14 985 ns (MD) 100 μs (CGMD)		

Table S2. Non-standard protonation states determined based on PBE/MC calculations (see main text *Methods*). HSE – ε-protonated (neutral) His, HSD - δ-protonated (neutral) His, HSP – ε/δ-protonated (charged) His.

Simulation	Residues	Subunit
S1-S17, S26-S61	His19(HSP), His53(HSE), His67(HSE), His124(HSE), His160(HSE)	NDUFS3
	His43(HSE), His421(HSE), His437(HSE), His494(HSE), His549(HSE)	NDUFS1
	His59(HSP), His190(HSP), His27(HSE), His55(HSE), His150(HSE), His200(HSE), His348(HSE), His398(HSE)	NDUFS2
	His42(HSE), His99(HSE)	NDUFV2
	His 261(HSP)	NDUFV1
	His67(HSP), His146(HSP)	NDUFS8
	His213(HSP), His338(HSP), His170(HSE), His184(HSE), His422(HSE), His430(HSE), His440(HSE)	ND4
	Glu34, Glu70	ND4L
	His50 (HSP), His190(HSP)	NDUFB7
	His111(HSP)	NDUFA6
	Asp26, His35(HSE)	NDUFAB1β
	His93(HSE), His171(HSE), His287(HSE), Glu262, Glu143	ND1
	His29(HSE), His109(HSE), His534(HSE)	ND5
	Asp66	ND3
	His316(HSE)	ND2
	His114(HSE), His151(HSE), His167(HSE), His257(HSE)	NDUFA10
	His2(HSE), His3(HSE), His58(HSE), His87(HSE), His131(HSE), His134(HSE), His240(HSE), His250(HSE), His321(HSE)	NDUFA9
	His29(HSE)	NDUFS4
	His13(HSE), His54(HSE)	NDUFS6
	His22(HSE)	NDUFV3
	His21(HSE), His41(HSE)	NDUFA5
	His75(HSE), His105(HSE), His129(HSE)	NDUFA6
	His30(HSE), His31(HSE), His99(HSE), His163(HSE)	NDUFA8
	His21(HSE), His108(HSE)	NDUFA11
	His27(HSE)	NDUFA1
	His59(HSE), His117(HSE)	NDUFC2
	His45(HSE), His97(HSE), His98(HSE)	NDUFS5
	His9(HSE), His12(HSE)	NDUFB1
	His135(HSE)	NDUFB5
	His82(HSE), His88(HSE)	NDUFB6
	His84(HSE), His94(HSE), His148(HSE)	NDUFB8
	His59(HSE)	NDUFB4
	His12(HSE), His76(HSE), His169(HSE)	NDUFB9
	His60(HSE), His91(HSE)	NDUFB7
	His55(HSE), His64(HSE)	NDUFB10
	His17(HSE), His87(HSE)	NDUFA12
	His38(HSE)	NDUFV3
S18, S22 (as in S1 except)	Lys135	ND2

	Lys237, Glu378	ND4
	Lys336, Lys392, Asp393	ND5
	Glu66	ND3
	Glu34	ND4L
S19, S23 (as in S1 except)	Glu34, Lys135	ND2
	Lys237, Glu123	ND4
	Lys336, Glu145	ND5
S20, S24 (as in S1 except)	Glu143	ND1
	Glu34, Glu70	ND4L
S21, S25 (as in S1 except)	Glu34, Glu70	ND4L

Table S3. Secondary structure restraints used for modeling the D state during MDFF.

Subunit	Residues	Strength (kcal mol^{-1} rad^{-2})
ND3	1-25, 30-34, 54-78, 84-109	200
NDUFV2	24-38, 42-59, 62-74, 76-87, 97-125, 136-143, 153-163, 165-179, 184-187	200
NDUFV1	32-38, 44-64, 76-80, 91-96, 105-112, 114-130, 132-139, 142-160, 173-180, 183-198, 214-236, 238-260, 265-283, 290-299, 315-326, 332-338, 342-358, 362-382, 385-402, 405-436	200
NDUFS7	35-54, 63-74, 78-83, 94-100, 104-117, 121-137, 145-149, 153-157, 162-179, 181-189	200
NDUFS1	8-21, 25-34, 55-60, 65-68, 76-81, 83-100, 114-124, 145-149, 157-167, 172-176, 181-185, 195-203, 207-211, 218-229, 236-253, 264-274, 282-286, 290-320, 324-338, 342-346, 358-362, 367-373, 375-381, 388-402, 404-410, 426-435, 437-446, 448-481, 491-494, 496-517, 520-523, 541-548, 558-563, 570-575, 579-584, 595-607, 615-632, 641-652, 675-693	200
ND1	1-33, 45-57, 67-86, 99-123, 125-156, 161-170, 177-192, 217-243, 251-277, 281-312	200
NDUFS8	15-25, 29-45, 65-70, 83-89, 93-101, 105-115, 122-128, 132-137, 146-175	200
ND6	2-22, 24-74, 85-109, 140-146, 148-171	200
NDUFS3	15-33, 36-44, 46-53, 55-67, 73-83, 89-98, 103-112, 116-119, 125-139, 141-144, 171-175, 184-188, 206-210	200
ND4L	1-21, 25-86	200
ND5	1-23, 31-58, 61-105, 112-134, 136-157, 162-194, 197-203, 209-224, 226-267, 270-293, 297-319, 321-350, 366-381, 387-402, 405-433, 447-472, 483-509, 517-579, 583-606	200
ND4	1-18, 21-40, 49-81, 88-112, 114-137, 140-172, 189-223, 225-249, 252-278, 281-338, 352-367, 372-414, 430-450	200
ND2	2-23, 25-46, 49-80, 91-107, 112-145, 150-171, 174-194, 197-222, 226-236, 239-253, 258-301, 309-317, 324-332, 336-343	200
NDUFA10	4-11, 14-18, 23-29, 34-52, 78-85, 91-119, 121-126, 130-141, 145-159, 167-173, 175-187, 195-224, 230-241, 247-274, 285-299, 316-320	200
NDUFA9	2-6, 20-24, 28-42, 44-50, 56-60, 67-73, 79-87, 89-94, 105-125, 127-133, 143-160, 162-173, 178-185, 190-197, 200-220, 226-231, 234-249, 254-272, 280-289, 309-322, 338-341	200
NDUFS2	7-14, 22-26,47-51, 64-78, 86-94, 96-100, 102-106, 111-127, 131-185, 206-229, 232-239, 246-254, 256-264, 268-273, 277-281, 292-318, 336-355, 360-381, 385-421, 424-430	200
NDUFS4	26-30, 33-39, 54-60, 63-67, 72-76, 82-107, 119-123	200

NDUFS6	25-29, 39-47, 49-53, 55-60, 70-75,81-84, 89-94	200
NDUFA2	1-171	200
NDUFAB1α	7-23, 43-58, 64-70, 72-82	200
NDUFAB1β	7-21, 35-41, 43-59, 63-67, 72-85	200
NDUFA5	21-37, 42-62, 64-72, 76-98	200
NDUFA6	27-53, 58-73, 78-98, 102-111, 120-127	200
NDUFA8	21-50, 52-92, 97-114	200
NDUFA11	4-13, 19-44, 50-82, 88-107, 109-134	200
NDUFA13	32-99, 129-137	200

Additional restraints centered on the MDFF setup, for modeling the D state.

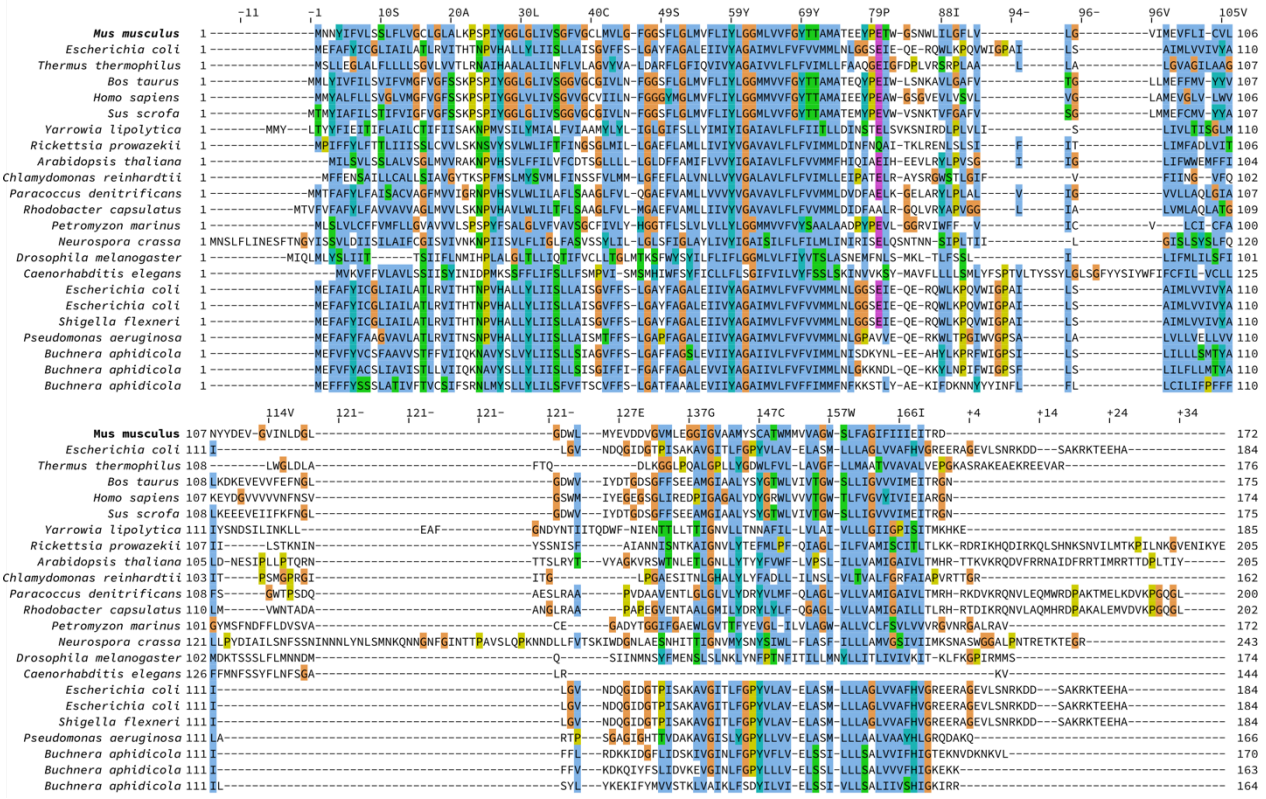
Subunit	**Residues**	**Atoms**	**Strength (kcal mol^{-1} Å^{-1})**
ND6	65-90	backbone	70
	67, 69	CB, CZ	70
	64	CG	70
ND3	52-58	backbone	70

Table S4. Overview of QM/MM MD and QM/MM free energy simulations. QM/MM model 1 – QM region shown in Fig. S1C; QM/MM model 2 – QM region in Fig. S1B. QM models 3-5 were constructed from QM models 1-2 for the DFT benchmarking (see SI Methods).

Simulation	State	Length	Starting point	Comment
S88	QH_2	2.5 ps	QM/MM model 1	Q site 2/N_{QM} =208
S89	QH_2	2.5 ps		Q site 2/N_{QM} =208
S90	pT from $Asp66^{ND3}$ →$Glu34^{ND4L}$	25 x 5 ps = 125 ps	QM/MM model 2	N_{QM} = 139
S91	pT from $Glu34^{ND4L}$ → $Glu70^{ND4L}$	19 x 5 ps = 95 ps	QM/MM model 3	$Glu^-/Lys^{+\,(ND2)}$ closed N_{QM} =90
S92	pT from $Glu34^{ND4L}$ →$Glu70^{ND4L}$	19 x 5 ps = 95 ps		$Glu^-/Lys^{0\,(ND2)}$ open N_{QM} =90
		Total: 320 ps		

Model	Residues in QM region
QM/MM model 1 (*N*=208 atoms)	$Glu24^{ND1}$ (protonated), $Arg195^{ND1}$, $Asp199^{ND1}$, $Glu202^{ND1}$, $Glu204^{ND1}$, $Ser209^{ND1}$, $Arg274^{ND1}$, $Phe224^{ND1}$, $Glu227^{ND1}$, $Tyr228^{ND1}$, $Arg279^{ND1}$, $Gln92^{NDUFS7}$, $Arg87^{NDUFS7}$, QH_2/QH^-, 9 H_2O
QM/MM model 2 (*N*=139 atoms)	$Asp66^{ND3}$, $Glu34^{ND4L}$, $Tyr59^{ND6}$, $Phe62^{ND3}$, $Leu63^{ND3}$, $Leu32^{ND6}$, $Leu57^{ND6}$, $Ile58^{ND6}$, $Val37^{ND4L}$, $Ala71^{ND4L}$, 6 H_2O,
QM/MM model 3 (*N*=90 atoms)	$Glu34^{ND4L}$, $Glu70^{ND4L}$, $Tyr59^{ND6}$, $Ala71^{ND4L}$, $Val37^{ND4L}$, $Ala67^{ND4L}$, $Met68^{ND2}$, $Ala64^{ND2}$, $Ser67^{ND2}$, 3 H_2O
DFT model 1 (*N*=250 atoms)	$Asp66^{ND3}$, $Glu34^{ND4L}$, $Tyr59^{ND6}$, $Phe62^{ND3}$, $Leu63^{ND3}$, $Leu67^{ND3}$, $Leu32^{ND6}$, $Leu57^{ND6}$, $Ile58^{ND6}$, $Met63^{ND6}$, $Leu30^{ND4L}$, $Leu31^{ND4L}$, $Val37^{ND4L}$, $Leu38^{ND4L}$, $Ala71^{ND4L}$, $Leu75^{ND4L}$, $Leu106^{ND2}$, 8 H_2O
DFT model 2 (*N*=190 atoms)	$Glu34^{ND4L}$, $Glu70^{ND4L}$, $Glu34^{ND2}$, $Lys105^{ND2}$, $Tyr59^{ND6}$, $Leu33^{ND4L}$, $Val37^{ND4L}$, $Ala67^{ND4L}$, $Ala71^{ND4L}$, $Leu37^{ND2}$, $Leu38^{ND2}$, $Gln63^{ND2}$, $Ala64^{ND2}$, $Ser67^{ND2}$, $Met68^{ND2}$, 7 H_2O
DFT model 3 (*N*=381 atoms)	$Asp66^{ND3}$, $Glu34^{ND4L}$, $Tyr59^{ND6}$, $Phe62^{ND3}$, $Leu63^{ND3}$, $Leu67^{ND3}$, $Leu32^{ND6}$, $Leu57^{ND6}$, $Ile58^{ND6}$, $Met63^{ND6}$, $Leu30^{ND4L}$, $Leu31^{ND4L}$, $Val37^{ND4L}$, $Leu38^{ND4L}$, $Ala71^{ND4L}$, $Leu75^{ND4L}$, $Leu106^{ND2}$, $Glu70^{ND4L}$, $Glu34^{ND2}$, $Lys105^{ND2}$, $Tyr59^{ND6}$, $Leu33^{ND4L}$, $Val37^{ND4L}$, $Ala67^{ND4L}$, $Ala71^{ND4L}$, $Leu37^{ND2}$, $Leu38^{ND2}$, $Gln63^{ND2}$, $Ala64^{ND2}$, $Ser67^{ND2}$, $Met68^{ND2}$, 15 H_2O
DFT model 4 (*N*=131 atoms)	$Asp66^{ND3}$, $Glu34^{ND4L}$, $Tyr59^{ND6}$, $Phe62^{ND3}$, $Leu63^{ND3}$, $Leu67^{ND3}$, $Leu32^{ND6}$, $Ala71^{ND4L}$, $Ile58^{ND6}$, 7 H_2O
DFT model 5 (*N*=27 atoms)	$Asp66^{ND3}$, $Glu34^{ND4L}$, 3 H_2O

Table S5. Multiple sequence alignment (20) of ND6 (top) and ND4L (bottom) from different species.



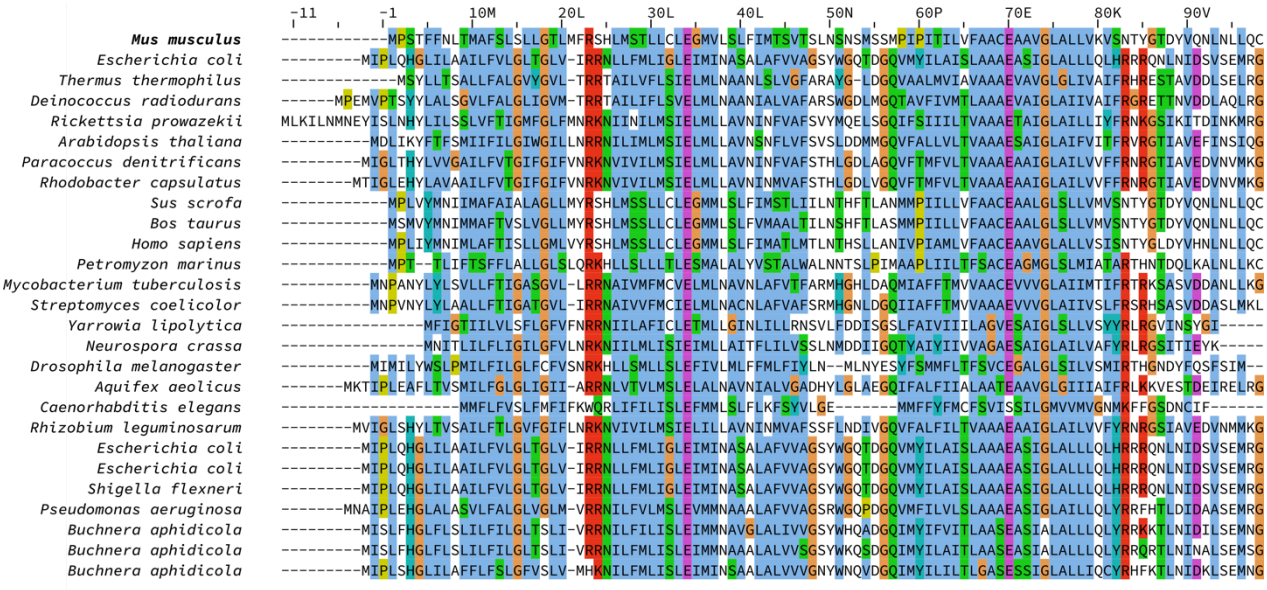


Table S6. Disease related mutations of subunits ND1/ND3/ND4L/ND6, encoded by mitochondrial DNA as contained in MITOMAP (22).

Nucleotide Position	Locus	Nucleotide Change	Reported Disease	residue	from AA	to AA
3308	MT-ND1	T-G	Sudden Infant Death	1	M	*
3308	MT-ND1	T-C	MELAS / DEAF enhancer /hypertension / LVNC /putative LHON	1	M	T
3310	MT-ND1	C-T	Diabetes / HCM	2	P	S
3316	MT-ND1	G-A	Diabetes / LHON / PEO	4	A	T
3335	MT-ND1	T-C	LHON	10	I	T
3336	MT-ND1	T-C	Carotid atherosclerosis risk			
3337	MT-ND1	G-A	Cardiomyopathy	11	V	M
3340	MT-ND1	C-T	Encephaloneuromyopathy	12	P	S
3376	MT-ND1	G-A	LHON MELAS overlap	24	E	K
3380	MT-ND1	G-A	MELAS	25	R	Q
3388	MT-ND1	C-A	Materally Inherited Nonsyndromic Deafness	28	L	M
3391	MT-ND1	G-A	LHON	29	G	S
3394	MT-ND1	T-C	LHON / Diabetes /CPT deficiency / high altitude adaptation	30	Y	H
3395	MT-ND1	A-G	LHON / HCM with hearing loss	30	Y	C
3396	MT-ND1	T-C	NSHL / MIDD			
3397	MT-ND1	A-G	ADPD / Possibly LVNC-cardiomyopathy associated	31	M	V
3398	MT-ND1	T-C	DMDF+HCM / GDM / possibly LVNC-cardiomyopathy associated	31	M	T
3399	MT-ND1	A-T	Gestational Diabetes (GDM)	31	M	I
3407	MT-ND1	G-A	HCM / Muscle involvement	34	R	H
3407	MT-ND1	G-A	HCM / Muscle involvement	34	R	H
3418	MT-ND1	A-G	AMegL	38	N	D
3421	MT-ND1	G-A	MIDD	39	V	I
3460	MT-ND1	G-A	LHON	52	A	T
3472	MT-ND1	T-C	LHON	56	F	L
3481	MT-ND1	G-A	MELAS / Progressive Encephalomyopathy	59	E	K
3488	MT-ND1	T-C	LHON	61	L	P
3496	MT-ND1	G-T	LHON	64	A	S
3497	MT-ND1	C-T	LHON	64	A	V
3551	MT-ND1	C-T	LHON	82	A	V
3632	MT-ND1	C-T	LHON	109	S	F
3634	MT-ND1	A-G	LHON	110	S	G
3635	MT-ND1	G-A	LHON	110	S	N
3644	MT-ND1	T-C	BD-associated	113	V	A
3667	MT-ND1	T-G	Peripheral neuropathy of T2 diabetes	121	W	G
3688	MT-ND1	G-A	Leigh Syndrome	128	A	T
3697	MT-ND1	G-A	MELAS / LS / LDYT / BSN	131	G	S
3700	MT-ND1	G-A	LHON	132	A	T
3713	MT-ND1	T-C	LHON	136	V	A
3733	MT-ND1	G-C	LHON	143	E	Q
3733	MT-ND1	G-A	LHON	143	E	K
3736	MT-ND1	G-A	LHON	144	V	I
3745	MT-ND1	G-A	LHON / high altitude variant	147	A	T
3769	MT-ND1	C-G	LHON	155	L	V
3781	MT-ND1	T-C	LHON	159	S	P
3796	MT-ND1	A-G	Adult-Onset Dystonia	164	T	A
3833	MT-ND1	T-A	PEG	176	L	Q

3866	MT-ND1	T-C	LHON +limb claudication	187	I	T
3890	MT-ND1	G-A	Progressive Encephalomyopathy / LS /Optic Atrophy	195	R	Q
3919	MT-ND1	T-C	LHON	205	S	P
3946	MT-ND1	G-A	MELAS	214	E	K
3949	MT-ND1	T-C	MELAS	215	Y	H
3958	MT-ND1	G-A	LHON	218	G	S
3959	MT-ND1	G-A	MELAS	218	G	D
3995	MT-ND1	A-G	MELAS	230	N	S
4081	MT-ND1	T-C	LHON	259	F	L
4123	MT-ND1	A-T	LHON	273	I	F
4132	MT-ND1	G-A	NAION-associated	276	A	T
4136	MT-ND1	A-G	LHON	277	Y	C
4142	MT-ND1	G-A	Developmental delay, seizure, hypotonia	279	R	Q
4160	MT-ND1	T-C	LHON	285	L	P
4163	MT-ND1	T-C	LHON	286	M	T
4171	MT-ND1	C-A	LHON / Leigh-like phenotype	289	L	M
4216	MT-ND1	T-C	LHON / Insulin Resistance/possible adaptive high altitude variant	304	Y	H
10086	MT-ND3	A-G	Hypertensive end-stage renal disease	10	N	D
10158	MT-ND3	T-C	Leigh Disease / MELAS	34	S	P
10191	MT-ND3	T-C	Leigh Disease /Leigh-like Disease / ESOC	45	S	P
10197	MT-ND3	G-A	Leigh Disease / Dystonia/ Stroke / LDYT	47	A	T
10237	MT-ND3	T-C	LHON	60	I	T
10254	MT-ND3	G-A	Leigh Disease	66	D	N
10398	MT-ND3	A-G	PD protective factor /longevity / altered cell pH / metabolic syndrome /breast cancer risk / LS risk / ADHD / cognitive decline / SCA2 age of onset	114	T	A
10398	MT-ND3	A-A	Invasive Breast Cancer risk factor AD PD BD lithium response Type 2DM			
10543	MT-ND4L	A-G	LHON	25	H	R
10591	MT-ND4L	T-G	LHON	41	F	C
10652	MT-ND4L	T-C	BD / MDD-associated			
10663	MT-ND4L	T-C	LHON	65	V	A
10680	MT-ND4L	G-A	LHON / synergistic combo 10680A + 12033G + 14258A	71	A	T
14163	MT-ND6	C-T	Possible deafness factor	174	N	H
14258	MT-ND6	G-A	LHON synergistic combo 10680A + 12033G + 14258A also combo 14258A +14582G	174	N	H
14279	MT-ND6	G-A	LHON	174	N	H
14319	MT-ND6	T-C	PD, early onset	174	N	H
14325	MT-ND6	T-C	LHON	174	N	H
14340	MT-ND6	C-T	SNHL	174	N	H
14430	MT-ND6	A-G	Thyroid Cancer	174	N	H
14439	MT-ND6	G-A	Mitochondrial Respiratory Chain Disorder	174	N	H
14453	MT-ND6	G-A	MELAS / Leigh Disease	174	N	H
14459	MT-ND6	G-A	LDYT / Leigh Disease /dystonia / carotid atherosclerosis risk	174	N	H
14482	MT-ND6	C-G	LHON	174	N	H
14482	MT-ND6	C-A	LHON	174	N	H
14484	MT-ND6	T-C	LHON	174	N	H

14487	MT-ND6	T-C	Dystonia / Leigh Disease/ ataxia / ptosis /epilepsy	174	N	H
14495	MT-ND6	A-G	LHON	174	N	H
14498	MT-ND6	T-C	LHON	174	N	H
14502	MT-ND6	T-C	LHON	174	N	H
14568	MT-ND6	C-T	LHON	174	N	H
14577	MT-ND6	T-C	MIDM	174	N	H
14582	MT-ND6	A-G	LHON synergistic combo 14258A + 14582G	174	N	H
14596	MT-ND6	A-T	LHON	174	N	H
14600	MT-ND6	G-A	Leigh Disease w/ optic atrophy	174	N	H
14668	MT-ND6	C-T	Depressive Disorder associated	174	N	Y

Table S7. Benchmarking the quantum chemical calculations using different density functionals against *ab initio* random-phase approximation (RPA) and domain-based local pair natural orbital coupled cluster theory with single-double and perturbative triples (DLPNO-CCSD(T)). **Top**: 381 atom DFT-model system (see Table S4) at the def2-TZVP/ε=4 level. **Middle**: 131 atom DFT-model system (see Table S4) at the def2-TZVP level and RPA/aug-cc-pVTZ level. **Bottom:** 27 atom DFT-model system, comprising an array of three water molecules, between two carboxylates (side chains of $D66^{ND3}$ and $E34^{ND4L}$), at the def2-TZVP level for the DFT calculations and using aug-cc-pVTZ basis set for the DLPNO-CCSD(T) and RPA calculations. Energies are given in kcal mol^{-1}.

N=381 atoms **Proton position**	**B3LYP-D3**	**CAM-B3LYP-D3**	**CAMh-B3LYP-D3**	**ωB97XD**	**TPSSh-D3**
D66	0.0	0.0	0.0	0.0	0.0
TS1	15.4	14.9	15.8	17.3	14.3
E34	5.9	5.6	6.4	6.0	6.3
TS2	16.4	15.9	17.1	18.0	16.1
E70	11.3	10.9	12.0	11.4	11.8

N=131 atoms	**RPA**	**B3LYP-D3**	**CAM-B3LYP-D3**	**CAMh-B3LYP-D3**	**ωB97XD**	**TPSSh-D3**
Reactant state	0.0	0.0	0.0	0.0	0.0	0.0
Transition state	20.4	15.9	15.5	16.5	18.4	15.0
Product state	7.9	6.8	6.4	7.4	7.4	7.2

N=27 atoms	**DLPNO-CCSD(T)**	**RPA**	**B3LYP-D3**	**CAM-B3LYP-D3**	**CAMh-B3LYP-D3**	**ωB97XD**	**TPSSh-D3**
Reactant state	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Transition state	23.5	23.4	19.3	18.7	19.3	21.3	17.9
Product state	0.1	0.2	-0.4	-0.8	-0.6	0.1	-0.9

Table S8. Subunit naming in different complex I isoforms.

Mouse	***T. thermophilus***	***E. coli***	**Cyanobacteria**
ND1	Nqo8	NuoH	NdhA
ND2	Nqo14	NuoN	NdhB
ND3	Nqo7	NuoA	NdhC
ND4	Nqo13	NuoM	NdhD
ND4L	Nqo11	NuoK	NdhE
ND5	Nqo12	NuoL	NdhF
ND6	Nqo10	NuoJ	NdhG
NDUFS1	Nqo3	NuoG	-
NDUFS2	Nqo4	NuoDa	NdhH
NDUFS3	Nqo5	NuoCa	NdhJ
NDUFS7	Nqo6	NuoB	NdhK
NDUFS8	Nqo9	NuoI	NdhI
NDUFV1	Nqo1	NuoF	-
NDUFV2	Nqo2	NuoE	-

Table S9. Force field parameters for H_3O^+ used in the classical simulations in the deactive state.

Atom	$\boldsymbol{q}$ (e)	e_i (kcalmol^{-1})	$R_{min}/2$ (Å)
O	-0.521	-0.1521	1.7682
H	+0.507	-0.46	0.2245
Bond	$\boldsymbol{k_b}$ (kcal mol^{-1} Å^{-2})	$\boldsymbol{r_0}$ (Å)	
O-H	450	1.05	
Angle	$\boldsymbol{k_\theta}$ (kcal mol^{-1} rad^{-2})	$\boldsymbol{\theta_0}$ (deg)	
H-O-H	55	113.3	

SI References

1. G. Fiorin, M. L. Klein, J. Hénin, Using collective variables to drive molecular dynamics simulations, *Mol. Phys.*, **111**, 3345-3362 (2013).
2. P. Liu, F. Dehez, W. Cai, C. Chipot, A toolkit for the analysis of free-energy perturbation calculations, *J. Chem. Theor. Comput.*, **8**, 2606-2616 (2012).
3. C. H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J. Comput. Phys.* **22**, 245–268 (1976).
4. A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. *Chem. Phys.* **98**, 5648–5652 (1993).
5. C. Lee, W. Yang, R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
6. S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104-1–19 (2010).
7. T. Yanai, D. P. Tew, N. C. Handy, A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **393**, 51–57 (2004).
8. Shao, Y. Mei, D. Sundholm, V. R. I. Kaila, Benchmarking the Performance of Time-Dependent Density Functional Theory Methods on Biochromophores. *J. Chem. Theory Comput.* **16**, 587–600 (2020).
9. J. Da Chai, M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008).
10. J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **91**, 146401 (2003).
11. F. Furche, Molecular tests of the random phase approximation to the exchange-correlation energy functional. *Phys. Rev. B - Condens. Matter Mater. Phys.* **64** (2001).
12. C. Riplinger, B. Sandhoefer), A. Hansen, F. Neese, Natural triple excitations in local coupled cluster calculations with pair natural orbitals *J. Chem. Phys.* **139**, 134101 (2013).
13. F. Weigend, R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
14. R. A. Kendall, T. H. Dunning, R. J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **96**, 6796–6806 (1992).
15. R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.*, **162**, 165–169 (1989).
16. F. Neese, F. Wennmohs, U. Becker, C. Riplinger, The ORCA quantum chemistry program package. *J. Chem. Phys.* **152**, 224108 (2020).
17. R. Baradaran, J. M., Berrisford, G. S. Minhas, L.A. Sazanov, Crystal structure of the entire respiratory complex I, *Nature*, **494**, 443–448, (2013).
18. J. M. Schuller, *et al*., Redox-coupled proton pumping drives carbon concentration in the photosynthetic complex I. *Nat. Commun.* **11**, 494, (2020).
19. K. Parey, *et al.* High-resolution cryo-EM structures of respiratory complex I: Mechanism, assembly, and disease, *Sci. Adv*. **5**, eaax9484 (2019).
20. F. Sievers, *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol. Syst. Biol*. **7** (2011).
21. J. Warnau, *et al.* Redox-Coupled Quinone Dynamics in the Respiratory Complex I. *Proc. Natl. Acad. Sci. USA*, **115,** E8413–E8420 (2018).
22. M. T. Lott, *et al.* MtDNA variation and analysis using Mitomap and Mitomaster. *Curr. Protoc. Bioinforma.* **44**, 1.23.1-1.23.26 (2013).
23. M. Kervinen, J. Pätsi, M., Finel, I. E. Hassinen, A Pair of Membrane-Embedded Acidic Residues in the NuoK Subunit of Escherichia coli NDH-1, a Counterpart of the ND4L Subunit of the Mitochondrial Complex I, Are Required for High Ubiquinone Reductase Activity, *Biochemistry* **43**, 773–781 (2004).

24. C. Kao, E. Nakamaru-Ogiso, A. Matsuno-Yagi, T. Yagi, Characterization of the membrane domain subunit NuoK (ND4L) of the NADH-quinone oxidoreductase from Escherichia coli, *Biochemistry* **44**, 9545–9554 (2005).
25. E. D. López, *et al.*, WATCLUST: A tool for improving the design of drugs based on protein-water interactions. *Bioinformatics* **31**, 3697–3699 (2015).
26. K. Nilsson, D. Lecerof, E. Sigfridsson, U. Ryde, An automatic method to generate force-field parameters for hetero-compounds. *Acta Cryst. D* **59**, 274–289 (2003).
27. C.I. Bayly, P. Cieplak, W.D. Cornell, P.A. Kollman, A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints For Determining Atom-Centered Charges: The RESP Model. *J. Phys. Chem*. **97**, 10269–10280(1993).
28. R. B. Best, *et al.*, Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ_1 and $\chi 2$ Dihedral Angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
29. P. Plessow, Reaction path optimization without NEB springs or interpolation algorithms, *J. Chem. Theory Comput*. **9,** 1305–1310 (2013).