Electronic Supplementary Material (ESI) for Chemical Communications.

Supporting information for

Biophysical and *in silico* characterization of NrtA: A protein-based host for aqueous nitrate and nitrite recognition

Ke Ji[a], Kiheon Baek[a], Weicheng Peng[ab], Kevin A. Alberto[a], Hedieh Torabifard[a], Steven O. Nielsen*[a], and Sheel C. Dodani*[a]

[a]Departments of Chemistry and Biochemistry and [b]Biological Sciences, The University of Texas at Dallas, Richardson, Texas 75080, USA

*sheel.dodani@utdallas.edu, steven.nielsen@utdallas.edu

Methods

General. Reagents and chemicals were purchased from Sigma-Aldrich, VWR International, or Thermo Fisher Scientific and were used as received, unless otherwise stated.

Multiple sequence alignment, percent identity matrix, and protein structures. Clustal Omega was used to generate the multiple sequence alignment and the percent identity matrix in Table S1.[1] All protein structures were generated using PyMOL v2.4.1.

Plasmid design. The genes encoding wild-type NrtA (UniProt ID: P73452) and the K269A variant were codon optimized for expression in *Escherichia coli* and inserted between the NdeI and XhoI restriction sites in the pET-21a(+) vector with a C-terminal polyhistidine-tag and stop codon from the vector (GenScript). The protein and DNA sequences are shown in Figures S1–S3.

Protein expression and purification. A freshly prepared plasmid encoding wild-type NrtA or the K269A variant was transformed into electrocompetent *E. cloni* EXPRESS BL21 (DE3) cells. Cells were plated on LB-Miller agar plates containing 100 µg/mL ampicillin and incubated for 15 h at 37 °C. Following this, a single colony was picked for protein expression into 60 mL of Terrific Broth (TB) media containing 100 µg/mL ampicillin and incubated overnight at 37 °C with shaking at 250 rpm. The next day, 1.2 L of TB media containing 100 µg/mL ampicillin was inoculated with 48 mL of the overnight culture and incubated at 37 °C for 1.5 h with shaking at 250 rpm. Once the OD_{600} reached 1.5, protein expression was induced with 1.2 mL of 1 M isopropyl β-d-1-thiogalactopyranoside for a final concentration of 1 mM. The flasks were cooled in an ice-water bath for 5 min with shaking, prior to incubation for 24 h at 18 °C with shaking at 200 rpm. The next day, cells were collected by centrifugation at 3,000*g* (5810 R, Eppendorf) for 30 min at 4 °C and resuspended in pre-chilled 20 mM HEPES buffer at pH 7.5 with 200 mM NaCl, 5 mM $MgCl_2$, 30 µg/mL deoxyribonuclease I, and one Pierce protease inhibitor XL capsule. The resuspended cells were stored at -20 °C until further use.

For protein purification, the cell pellets were thawed overnight at 4 °C and lysed using sonication at 30% amplitude, 15 s pulse on, and 45 s pulse off for 5 min (QSonica) in an ice-water bath. The lysates were clarified by ultracentrifugation at 37,000*g* (Optima XPN-80 Ultracentrifuge, Beckman Coulter) for 35 min at 4 °C. The resulting supernatant was loaded onto a pre-equilibrated 5 mL column pre-packed with a nickel-nitrilotriacetic acid agarose resin (HisTrap, GE Healthcare) via a sample pump (NGC Chromatography System, Bio-Rad Laboratories). The column was then washed with 20 column volumes (CV) of 20 mM HEPES buffer at pH 7.5 with 200 mM NaCl and 30 mM imidazole. The polyhistidine-tagged protein was eluted with a 0% to 100% linear gradient of a 20 mM HEPES buffer at pH 7.5 with 200 mM NaCl, and 500 mM imidazole over 20 CV. The eluted fractions with an absorbance at 280 nm were pooled and loaded on a HiPrep 26/10 Desalting column with 20 mM HEPES at pH 7.7 with 400 mM NaCl. The eluted fractions with an absorbance at 280 nm were pooled and concentrated to less than 10 mL using an EMD Millipore Amicon Ultra-15 Centrifugal Filter Unit with a 30 kDa molecular weight cut-off.

The resulting protein samples were loaded onto a pre-equilibrated HiLoad 26/600 Superdex 200 size-exclusion column (GE Healthcare). Pure monomeric protein was then eluted with 20 mM HEPES buffer at pH 7.7 with 400 mM NaCl using an elution flow rate of 0.8 mL/min. Eluted fractions with an absorbance at 280 nm were pooled and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit with a 30 kDa molecular weight cut-off. The concentrated protein was dialyzed against 20 mM HEPES buffer at pH 7.7 with 400 mM NaCl (final sample to buffer ratio: 1:10,000) using Slide-A-Lyzer Dialysis Cassette with 10 kDa molecular weight cut-off at 10 °C.

After 8 h, the dialysis buffer was switched to 20 mM HEPES buffer at pH 7.5 with 100 mM NaCl for differential scanning fluorimetry (DSF) experiments or 20 mM HEPES buffer at pH 7.7 with 50 mM NaCl for isothermal titration calorimetry (ITC) experiments for 14 h (final sample to buffer ratio: 1:10,000). For ITC experiments with Tris and sodium phosphate buffers, the protein sample in HEPES buffer at pH 7.5 with 50 mM NaCl was dialyzed against 20 mM Tris or sodium phosphate buffers at pH 7.7 with 50 mM NaCl at 10 °C for 14 h (final sample to buffer ratio: 1:10,000). The next day, the samples were clarified using centrifugation at 21,000*g* (5424 R, Eppendorf) for 15 min at 4°C. Protein samples were stored at 4 °C until further use.

Protein concentration determination. The concentrations of wild-type NrtA and the K269A variant were determined in 20 mM HEPES buffer at pH 7.7 with 50 mM NaCl using the absorbance of tryptophan residues at 280 nm according to the Beer-Lambert law, $A = \varepsilon lc$ where A is the absorbance, ε is the extinction coefficient, l is the optical path length in cm, and c is the concentration. The extinction coefficient of 85,370 $M^{-1}cm^{-1}$ was determined using the ProtParam tool in ExPASy.[2] Absorption spectra were recorded from 250–400 nm with a 1 nm step size using an Agilent Cary 7000 UV−Vis-NIR spectrophotometer in the Molecular and Protein Characterization Core at UT Dallas. Purified protein samples were diluted at least 40-fold to ensure that the absorbance values were below 0.3. All measurements were carried out in a quartz cuvette with a 10 mm pathlength (Hellma Analytics) at room temperature in triplicate for each protein preparation.

SDS-PAGE. General methods were adapted from a previous protocol with the following modifications.[3] Briefly, 100 µL of 2-mercaptoethanol was added to 900 µL of 4x Laemmli Sample Buffer (Bio-Rad Laboratories). Following this, 16 µL of 5 µM purified protein samples were mixed with 4 µL of 4x Laemmli Sample Buffer and boiled at 95 °C for 5 min. The samples were then loaded onto a gel prepared from TGX Stain-free FastCast Acrylamide Solutions (Bio-Rad Laboratories). The loaded gel was placed in a Mini-PROTEAN Tetra Cell (Bio-Rad Laboratories) and merged in the running buffer prepared from a TGS 10X Solution (Research Products International) with an applied voltage of 200 mV for 60 min using a PowerPac Basic Power Supply (Bio-Rad Laboratories). The gel was visualized and analyzed using the ChemiDoc Touch Imaging System and Image Lab Software for PC Version 6.0 (Bio-Rad Laboratories, Figure S4).

Differential scanning fluorimetry (DSF). General DSF methods were adapted from a previous protocol.[4] Briefly, purified protein samples were diluted to 10 µM in 20 mM HEPES buffer at pH 7.5 with 100 mM NaCl and SYPRO Orange Protein Gel Stain (5X). Fifteen microliters of 20 µM protein and 15 µL of 0.002, 0.01. 0.02, 0.1, 0.2, 1, 2, 20, 100, or 200 mM anion stock solutions in the same buffer were dispensed in triplicate into a MicroAmp Fast Optical 96-Well Reaction Plate. The final protein concentration was 10 µM with 0.001, 0.005. 0.01, 0.05, 0.1, 0.5, 1, 10, 50, and 100 mM sodium nitrate, nitrite, or gluconate. All the sample plates were sealed by MicroAmp Optical Adhesive Film and centrifuged at 800*g* (5810 R, Eppendorf) at room temperature to remove any air bubbles. The experiments were carried out using QuantStudio 6 Flex Real-Time PCR System in the Genome Center at UT Dallas. The ROX dye was selected to be the reporter with the x4-m4 default excitation and emission filters, and the passive reference and quencher were set to none. Data was collected over a 25 to 99 °C temperature gradient up to 30 min with 3 °C/min increments. The resulting data was analyzed using the Protein Thermal Shift Software v1.3. To compare samples, the raw fluorescence was normalized for Figures S5 and S6. For each sample, the unfolding temperature (T_m) was determined using the values of the first derivative of the fluorescence (Tables S2, S3). The average of six technical replicates from two independent measurements with the standard deviation is reported for one protein preparation.

Isothermal titration calorimetry (ITC). All ITC experiments were performed using a Low Volume Affinity ITC (TA Instruments). All sodium nitrate and sodium nitrite solutions were prepared in 20 mM HEPES buffer, Tris, or sodium phosphate buffer at pH 7.7 with 50 mM NaCl to minimize any buffer mismatch with the protein samples. The nitrate and nitrite concentrations used in the syringe were kept at 150 μM and 300 μM, respectively. For each anion, a series of protein concentrations were screened for optimal signal over background. The final protein concentrations in the sample cell were 15 μM for nitrate and 28 μM for nitrite. For every experiment, the first injection was 0.4 μL followed by 24 injections of 1.5 μL. The stirring rate was set to 200 rpm. The dilutional heat was measured by titrating sodium nitrate or sodium nitrite into each buffer at each temperature tested. All ITC experiments were performed in duplicate with the same protein batch and a single-site binding model was used to generate the sigmoidal fit for the two independent experiments. Each experiment was integrated using the NanoAnalyze software (TA Instruments) and the baseline of each peak was adjusted manually. All thermodynamic parameters, including the dissociation constant (K_d), stoichiometry (n), the binding enthalpy (ΔH), the binding entropy (TΔS), and binding Gibbs free energy (ΔG), were determined using the NanoAnalyze software based on following equations:

$$Q = V\Delta H[L_B] = V[P]\frac{n\Delta HK[L]}{1 + K[L]}$$

$$K_d = \frac{1}{K}$$

$$\Delta G = -RT\ln K = \Delta H - T\Delta S$$

where Q is the cumulative amount of heat released as the result of ligand binding, V is the volume of the sample cell, ΔH is the enthalpy of binding, $[L_B]$ is the concentration of ligand bound to protein, $[P]$ is the protein concentration, n is the stoichiometry, K is the association binding constant, and $[L]$ is the free ligand concentration.[5]

All thermodynamic parameters are reported as an average of two protein preparations, each measured in duplicate with the standard deviation (Figures S7–S13; Tables S4, S5). The heat capacity change (ΔC_p) upon ion binding was obtained from the linear regression of plotting ΔH versus temperature.

Molecular dynamics (MD) simulations and principal component analysis (PCA). The CHARMM36 protein force field parameters were used for the simulations, and the TIP3P model was used for water molecules.[6, 7] The apo NrtA model was generated by deleting the nitrate ion in the crystal structure, the nitrate-bound NrtA was created by using the crystal structure (PDB ID: 2G29), and the nitrite-bound NrtA was created by replacing nitrate with nitrite according to their centers of mass.[8] The nitrate and nitrite force fields were obtained from previous studies and used without further modification.[9, 10] The system was solvated and neutralized the charge with the total number of ~47,000 atoms within a unit cell of 85 x 74 x 75 $Å^3$ according to VMD (http://www.ks.uiuc.edu/Research/vmd/).[11] The entire system was energy-minimized for 1,000 steps. To reduce any undesired protein movement after energy minimization, gradual heating followed by equilibration was adapted from a previous study with the following modifications: the system was heated from 0 K to 100 K over 12.5 ps and then from 100 K to 310 K in the NVT ensemble over 125 ps, using 10.0 $kcal \cdot mol^{-1} \cdot Å^{-2}$ harmonic restraints applied to protein heavy atoms.[12] The system was then equilibrated with 5.0, 2.5, and 0.1 $kcal \cdot mol^{-1} \cdot Å^{-2}$ harmonic restraints on protein heavy atoms for 100 ps in the NPT ensemble at 310 K, respectively.[12]

MD simulations were carried out at 310K for 600 ns for the apo form and 400 ns for the nitrate and nitrite-bound forms via the Nanoscale Molecular Dynamics (NAMD) software package on Lonestar5 (v.2.10) and Longhorn (v.2.14) system of the Texas Advanced Computing Center (http://www.ks.uiuc.edu/Research/namd/).[13] In order to increase the sampling space, the second and third run were performed starting with the coordinates saved from previous simulations at 50 ns and 100 ns, respectively.

MDAnalysis and scikit-learn libraries were used for processing all simulation data except solvent accessible surface areas (SASAs).[14] The root mean square deviation (RMSD) for alpha carbon (Cα) atoms were calculated to determine the equilibrated state of each trajectory. For all one-dimensional RMSD plots, the starting structure of each simulation was used as a reference (Figure S14). Based on the one and two-dimensional RMSD results, the first 500 ns of each simulation for the apo form and 300 ns of each simulation for the nitrate and nitrite-bound forms were removed. Therefore, only the last 100 ns corresponding to the equilibrated portion of each trajectory was used for analysis (Figures S14, S15).

The distances between the anions and the binding residues were measured every 500 ps using the collective variables module in NAMD (Figures S16–S30; Table S6).[15] Representative snapshots from the nitrate and nitrite-bound trajectories are shown in Figure S31. The dynamic cross-correlation matrix analysis (DCCM) of the Cα atoms for apo, nitrate-bound, and nitrite-bound forms are shown in Figure S32 and Supplementary information.xlsx. The anticorrelated regions were projected on the crystal structure as shown in Figure S33.[16] SASAs were measured using VMD for each equilibrated trajectory, and the time-averaged values for are summarized in Table S7.[11]

Principal component analysis (PCA) was processed by employing the diagonalized covariance matrix of the Cα Cartesian coordinates from each trajectory. Data were normalized before processing PCA, and the cosine content of the first principal component (PC1) was calculated for evaluating the validity of sampling. To analyze the conformations sampled in each trajectory of the apo, nitrate-bound, and nitrite-bound forms, the PC1 was plotted versus PC2 (Figure S34). Trajectories were projected on the eigenvectors for obtaining the collective motions. After obtaining the collective motion, root mean square fluctuation (RMSF) of the Cα atoms were measured (Figure S35). To compare the results between all the datasets, only the residues that had Cα RMSF values in PC1 direction greater than 0.4 Å and were common to all trajectories in each form were selected (Tables S8–S11).

Supporting Figures and Tables

Table S1. Multiple sequence alignment and percent identity matrix of *Synechocystis sp. PCC 6803* (UniProt ID: P73452), *Synechococcus elongatus* (UniProt ID: P38043), and *Phormidium laminosum* (UniProt ID: Q51880). Residues within 4 Å of nitrate ion were selected according to the NrtA crystal structure from *Synechocystis sp. PCC 6803* (PDB ID: 2G29).

Organism	Nitrate-binding residues										% Identity
	71	102	124	155	190	196	222	239	240	269	
Synechocystis sp. PCC 6803	L	W	L	Q	T	H	P	V	G	K	100
Synechococcus elongatus	L	W	L	Q	T	H	P	V	G	K	56
Phormidium laminosum	P	W	L	Q	T	H	P	V	G	K	67

30 TSPTTTSTGTGTGSSTDQAISPLVEGENAPEVTTAKLGFIALTDAAPLIIAKEKGFYAKY 89
90 GMPDVEVLKQASWGTTRDNLVLGSASGGIDGAHILTPMPYLITMGTVTDGKPTPMYILAR 149
150 LNVNGQGIQLGNNYKDLKVGTDAAPLKEAFAKVTDPKVAMTFPGGTHDMWIRYWLAAGGM 209
210 EPGKDFSTIVVPPAQMVANVKVNAMESFCVGEPWPLQTVNQGVGYQALTTGQLWKDHPEK 269
270 AFGMRADWVDQNPKAAKALLMAVMEAQQWCDQAENKEEMCQILSKREWFKVPFEDIIDRS 329
330 KGIYNFGNGQETFEDQEIMQKYWVDNASYPYKSHDQWFLTENIRWGYLPASTDTKAIVDK 389
390 VNREDLWREAAQALEVPADQIPSSPSRGIETFFDGITFDPENPQAYLDSLKIKSIKA 446
447 LEHHHHHH*

Figure S1. Amino acid sequence for NrtA from *Synechocystis sp. PCC 6803* (UniProt ID: P73452) used in this study. The K269 position for mutagenesis is highlighted in red.

CATATGACCAGCCCGACCACCACCAGCACCGGTACCGGTACCGGTAGCAGCACCGACCAGGC
GATCAGCCCGCTGGTTGAGGGTGAAAACGCGCCGGAAGTTACCACCGCGAAACTGGGTTTCAT
TGCGCTGACCGACGCGGCGCCGCTGATCATTGCGAAGGAAAAAGGCTTTTACGCGAAATATGG
TATGCCGGATGTTGAGGTGCTGAAGCAAGCGAGCTGGGGTACCACCCGTGACAACCTGGTTCT
GGGTAGCGCGAGCGGTGGCATTGATGGTGCGCACATTCTGACCCCGATGCCGTACCTGATCA
CCATGGGTACCGTGACCGATGGCAAGCCGACCCCGATGTATATCCTGGCGCGTCTGAACGTTA
ACGGCCAGGGTATTCAACTGGGCAACAACTACAAGGACCTGAAAGTGGGTACCGATGCGGCG
CCGCTGAAGGAAGCGTTCGCGAAAGTGACCGACCCGAAGGTTGCGATGACCTTTCCGGGTGG
CACCCACGATATGTGGATCCGTTATTGGCTGGCGGCGGGTGGCATGGAGCCGGGTAAAGATTT
CAGCACCATTGTGGTTCCGCCGGCGCAGATGGTGGCGAACGTGAAGGTTAACGCGATGGAAA
GCTTTTGCGTTGGCGAGCCGTGGCCGCTGCAGACCGTGAACCAAGGCGTTGGTTACCAGGCG
CTGACCACCGGCCAACTGTGGAAAGACCACCCGGAGAAGGCGTTCGGTATGCGTGCGGACTG
GGTGGATCAAAACCCGAAGGCGGCGAAAGCGCTGCTGATGGCGGTTATGGAAGCGCAGCAAT
GGTGCGATCAGGCGGAGAACAAAGAGGAAATGTGCCAAATCCTGAGCAAACGTGAATGGTTCA
AGGTTCCGTTTGAGGACATCATTGATCGTAGCAAGGGCATCTACAACTTCGGCAACGGTCAGG
AGACCTTCGAAGATCAGGAGATCATGCAAAAGTACTGGGTGGACAACGCGAGCTACCCGTATA
AGAGCCACGATCAATGGTTCCTGACCGAAAACATCCGTTGGGGTTACCTGCCGGCGAGCACCG
ACACCAAAGCGATTGTGGATAAGGTTAACCGTGAAGACCTGTGGCGTGAAGCGGCGCAGGCG
CTGGAAGTGCCGGCGGACCAAATTCCGAGCAGCCCGAGCCGTGGCATCGAAACCTTCTTTGAC
GGTATTACCTTTGATCCGGAGAACCCGCAGGCGTATCTGGATAGCCTGAAGATCAAAAGCATTA
AAGCGCTCGAGCACCACCACCACCACCACTGA

Figure S2. The nucleotide sequence encoding wild-type NrtA (UniProt ID: P73452) used in this study. The following regions are highlighted: restriction sites (orange), NrtA (black), His-tag (purple), and stop codon (green).

CATATGACCAGCCCGACCACCACCAGCACCGGTACCGGTACCGGTAGCAGCACCGACCAGGC
GATCAGCCCGCTGGTTGAGGGTGAAAACGCGCCGGAAGTTACCACCGCGAAACTGGGTTTCAT
TGCGCTGACCGACGCGGCGCCGCTGATCATTGCGAAGGAAAAAGGCTTTTACGCGAAATATGG
TATGCCGGATGTTGAGGTGCTGAAGCAAGCGAGCTGGGGTACCACCCGTGACAACCTGGTTCT
GGGTAGCGCGAGCGGTGGCATTGATGGTGCGCACATTCTGACCCCGATGCCGTACCTGATCA
CCATGGGTACCGTGACCGATGGCAAGCCGACCCCGATGTATATCCTGGCGCGTCTGAACGTTA
ACGGCCAGGGTATTCAACTGGGCAACAACTACAAGGACCTGAAAGTGGGTACCGATGCGGCG
CCGCTGAAGGAAGCGTTCGCGAAAGTGACCGACCCGAAGGTTGCGATGACCTTTCCGGGTGG
CACCCACGATATGTGGATCCGTTATTGGCTGGCGGCGGGTGGCATGGAGCCGGGTAAAGATTT
CAGCACCATTGTGGTTCCGCCGGCGCAGATGGTGGCGAACGTGAAGGTTAACGCGATGGAAA
GCTTTTGCGTTGGCGAGCCGTGGCCGCTGCAGACCGTGAACCAAGGCGTTGGTTACCAGGCG
CTGACCACCGGCCAACTGTGGAAAGACCACCCGGAGGCGGCGTTCGGTATGCGTGCGGACTG
GGTGGATCAAAACCCGAAGGCGGCGAAAGCGCTGCTGATGGCGGTTATGGAAGCGCAGCAAT
GGTGCGATCAGGCGGAGAACAAAGAGGAAATGTGCCAAATCCTGAGCAAACGTGAATGGTTCA
AGGTTCCGTTTGAGGACATCATTGATCGTAGCAAGGGCATCTACAACTTCGGCAACGGTCAGG
AGACCTTCGAAGATCAGGAGATCATGCAAAAGTACTGGGTGGACAACGCGAGCTACCCGTATA
AGAGCCACGATCAATGGTTCCTGACCGAAAACATCCGTTGGGGTTACCTGCCGGCGAGCACCG
ACACCAAAGCGATTGTGGATAAGGTTAACCGTGAAGACCTGTGGCGTGAAGCGGCGCAGGCG
CTGGAAGTGCCGGCGGACCAAATTCCGAGCAGCCCGAGCCGTGGCATCGAAACCTTCTTTGAC
GGTATTACCTTTGATCCGGAGAACCCGCAGGCGTATCTGGATAGCCTGAAGATCAAAAGCATTA
AAGCGCTCGAGCACCACCACCACCACCACTGA

Figure S3. The nucleotide sequence encoding the K269A variant. The following regions are highlighted: restriction sites (orange), NrtA (black), K269A mutation site (red), His-tag (purple), and stop codon (green).

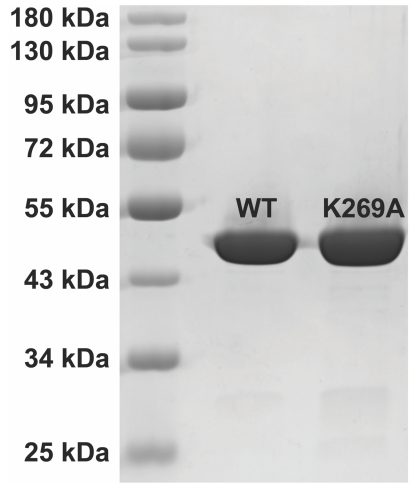

Figure S4. Stain-free SDS-PAGE gel of purified wild-type NrtA and the K269A variant. The expected molecular weight for wild-type NrtA is 47.0 kDa and the K269A variant is 46.9 kDa.

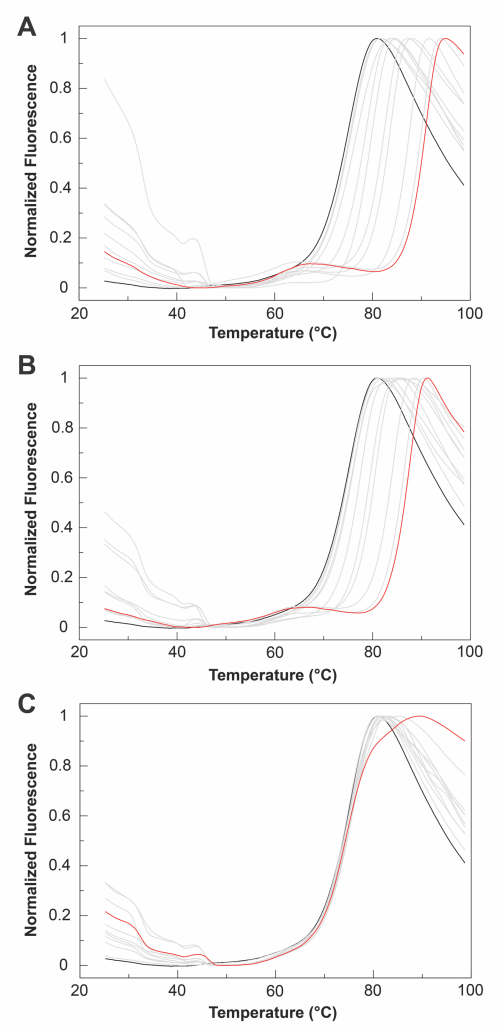

Figure S5. Representative thermal unfolding of wild-type NrtA with 0.001, 0.005. 0.01, 0.05, 0.1, 0.5, 1, 10, 50, and 100 mM sodium (A) nitrate, (B) nitrite, and (C) gluconate as monitored by SYPRO Orange using differential scanning fluorimetry (DSF). All experiments were carried out in 20 mM HEPES buffer at pH 7.5 with 100 mM NaCl. For each ion concentration, the unfolding temperature (T_m) is reported in Table S2.

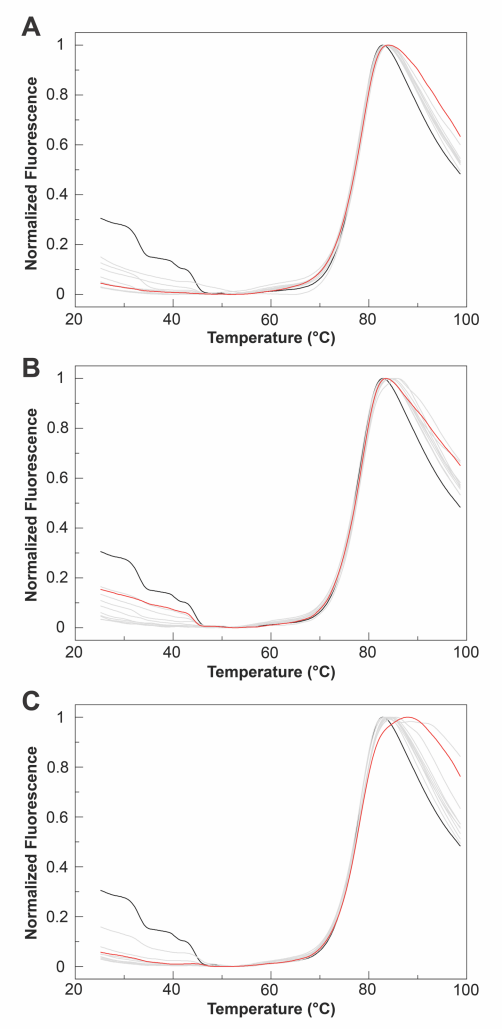

Figure S6. Representative thermal unfolding of the K269A variant with 0.001, 0.005. 0.01, 0.05, 0.1, 0.5, 1, 10, 50, and 100 mM sodium (A) nitrate, (B) nitrite, and (C) gluconate as monitored by SYPRO Orange using DSF. All experiments were carried out in 20 mM HEPES buffer at pH 7.5 with 100 mM NaCl. For each ion concentration, the unfolding temperature (T_m) is reported in Table S3.

Table S2. The T_m and ΔT_m values of wild-type NrtA obtained from DSF in Figure S5. The average of six technical replicates collected over two independent experiments with standard deviation is reported for one protein preparation.

Concentration (mM)	T_m (°C)			ΔT_m (°C)		
	Nitrate	**Nitrite**	**Gluconate**	**Nitrate**	**Nitrite**	**Gluconate**
0		75.19 ± 0.18			-	
0.001	75.37 ± 0.44	75.27 ± 0.75	75.00 ± 0.50	0.18 ± 0.48	0.08 ± 0.77	-0.19 ± 0.53
0.005	75.76 ± 0.55	75.61 ± 0.55	75.39 ± 0.92	0.57 ± 0.58	0.42 ± 0.58	0.20 ± 0.94
0.01	76.03 ± 0.57	75.22 ± 0.24	74.95 ± 0.45	0.84 ± 0.60	0.03 ± 0.30	-0.24 ± 0.49
0.05	78.91 ± 0.41	76.23 ± 0.62	75.66 ± 0.35	3.72 ± 0.45	1.04 ± 0.65	0.47 ± 0.39
0.1	79.90 ± 0.27	77.19 ± 0.42	75.32 ± 0.62	4.71 ± 0.32	2.00 ± 0.46	0.13 ± 0.65
0.5	82.51 ± 0.13	79.48 ± 0.21	75.19 ± 0.46	7.32 ± 0.22	4.29 ± 0.28	0.00 ± 0.49
1	83.64 ± 0.16	80.61 ± 0.06	75.71 ± 2.32	8.45 ± 0.24	5.42 ± 0.19	0.52 ± 2.33
10	87.63 ± 0.16	84.21 ± 0.08	75.78 ± 0.19	12.44 ± 0.24	9.02 ± 0.20	0.59 ± 0.26
50	90.24 ± 0.13	86.57 ± 0.08	74.80 ± 0.14	15.05 ± 0.22	11.38 ± 0.20	-0.39 ± 0.23
100	90.93 ± 0.08	87.61 ± 0.09	74.26 ± 0.37	15.74 ± 0.20	12.42 ± 0.20	-0.93 ± 0.41

Table S3. The T_m and ΔT_m values for the K269A variant obtained from DSF in Figure S6. The average of six technical replicates collected over two independent experiments with standard deviation is reported for one protein preparation.

Concentration (mM)	T_m (°C)			ΔT_m (°C)		
	Nitrate	Nitrite	Gluconate	Nitrate	Nitrite	Gluconate
0		78.34 ± 0.50			-	
0.001	78.30 ± 0.06	78.59 ± 0.40	78.69 ± 0.27	0.37 ± 0.50	0.28 ± 0.64	-0.37 ± 0.57
0.005	78.62 ± 0.42	78.57 ± 0.49	78.34 ± 0.32	0.25 ± 0.65	0.20 ± 0.70	0.03 ± 0.59
0.01	78.47 ± 0.15	78.47 ± 0.68	78.32 ± 0.36	0.32 ± 0.52	0.23 ± 0.84	0.42 ± 0.62
0.05	78.52 ± 0.12	78.49 ± 0.47	78.57 ± 0.28	0.05 ± 0.51	0.30 ± 0.69	0.18 ± 0.57
0.1	78.44 ± 0.25	78.54 ± 0.29	78.62 ± 0.61	-0.09 ± 0.56	0.18 ± 0.58	0.15 ± 0.79
0.5	78.25 ± 0.20	78.52 ± 0.29	78.49 ± 0.34	0.10 ± 0.54	0.20 ± 0.58	0.28 ± 0.61
1	78.39 ± 0.14	78.64 ± 0.20	78.52 ± 0.33	0.18 ± 0.52	0.15 ± 0.54	0.23 ± 0.60
10	78.66 ± 0.15	78.57 ± 0.16	78.76 ± 0.08	0.13 ± 0.52	0.13 ± 0.53	-0.02 ± 0.51
50	78.59 ± 0.11	78.54 ± 0.15	78.37 ± 0.20	0.28 ± 0.51	0.23 ± 0.52	0.00 ± 0.54
100	78.71 ± 0.25	78.62 ± 0.22	77.97 ± 0.30	-0.04 ± 0.56	0.25 ± 0.55	0.35 ± 0.58

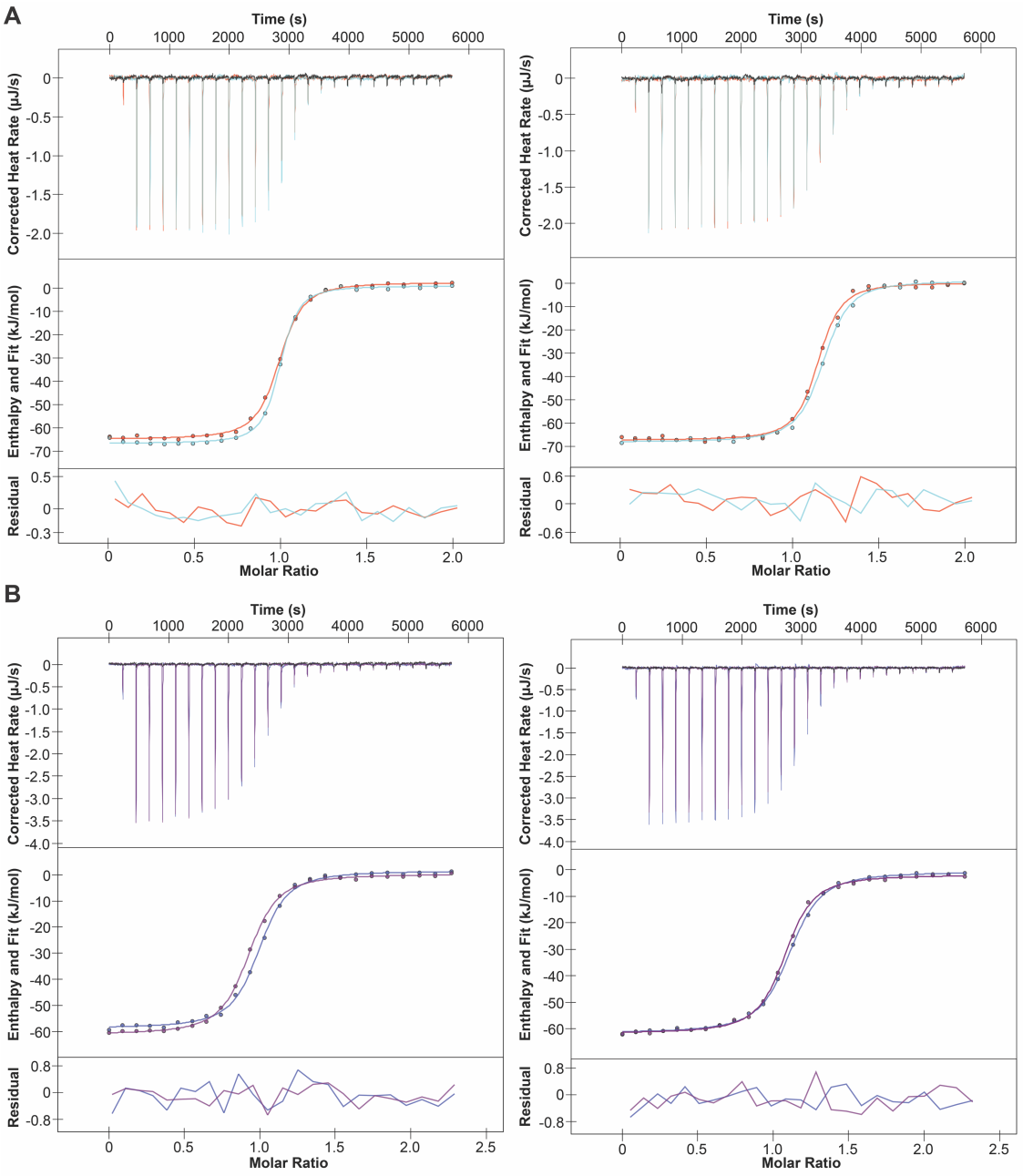

Figure S7. Isothermal titration calorimetry (ITC) analysis of (A) nitrate and (B) nitrite binding to NrtA at 20 °C. In each panel, the thermogram (top), isotherm (middle), and residual (bottom) plot is shown. Titration of the anion into the buffer alone (black) is included in each thermogram. All experiments were carried out in 20 mM HEPES buffer at pH 7.7 with 50 mM NaCl. Data is shown for two protein preparations, each measured in duplicate.

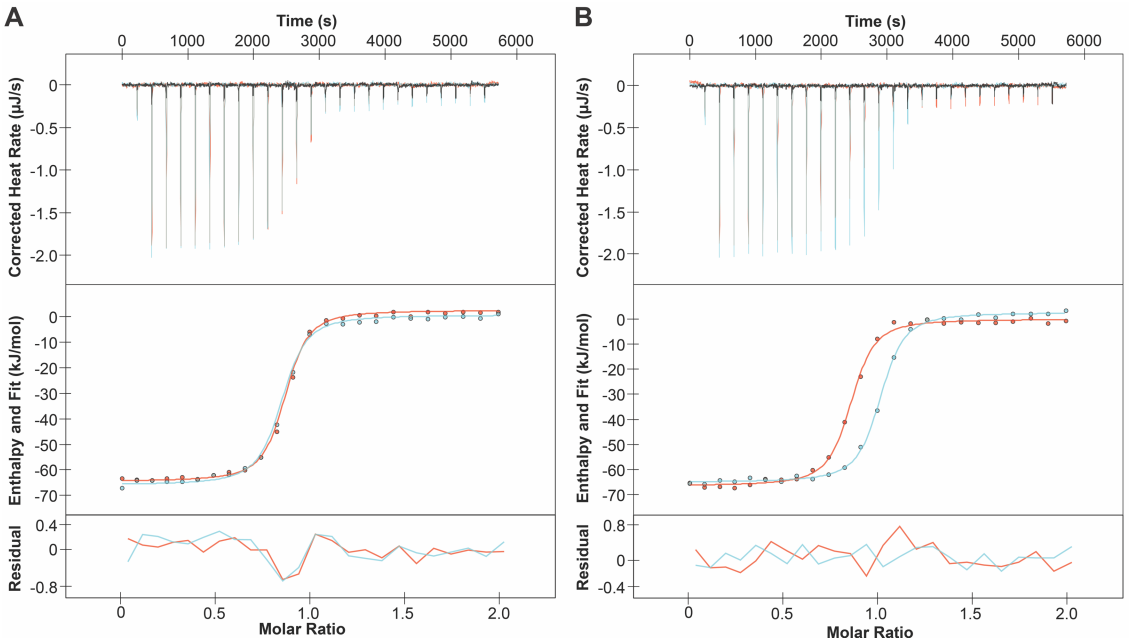

Figure S8. ITC analysis of nitrate binding to NrtA at 20 °C in 20 mM (A) Tris or (B) sodium phosphate buffer at pH 7.7 with 50 mM NaCl. In each panel, the thermogram (top), isotherm (middle), and residual (bottom) plot is shown. Titration of the anion into the buffer alone (black) is included in each thermogram. Data is shown for one protein preparation measured in duplicate.

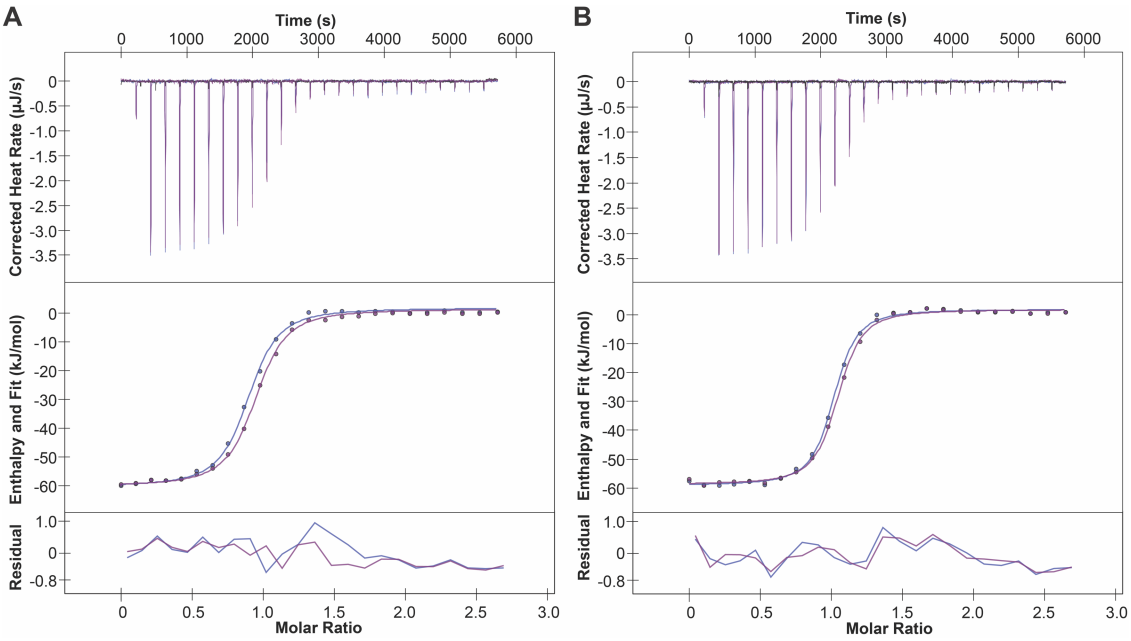

Figure S9. ITC analysis of nitrite binding to NrtA at 20 °C in 20 mM (A) Tris or (B) sodium phosphate buffer at pH 7.7 with 50 mM NaCl. In each panel, the thermogram (top), isotherm (middle), and residual (bottom) plot is shown. Titration of the anion into the buffer alone (black) is included in each thermogram. Data is shown for one protein preparation measured in duplicate.

Table S4. Thermodynamic parameters for nitrate and nitrite binding to NrtA at 20 °C in 20 mM HEPES, Tris, and sodium phosphate buffers at pH 7.7 with 50 mM NaCl.

Anion	Buffer	K_d (nM)	ΔH_{exp} (kJ·mol^{-1})	TΔS (kJ·mol^{-1})	ΔG (kJ·mol^{-1})	n
Nitrate	HEPES	57 ± 21	-68.1 ± 1.2	-27.4 ± 1.7	-40.7 ± 0.9	1.06 ± 0.01
	Tris	65 ± 7	-66.7 ± 0.3	-26.4 ± 0.0	-40.3 ± 0.3	0.98 ± 0.01
	Phosphate	60 ± 9	-67.6 ± 0.2	-27.0 ± 0.4	-40.6 ± 0.4	1.05 ± 0.11
Nitrite	HEPES	233 ± 41	-60.7 ± 1.3	-23.4 ± 1.8	-37.2 ± 0.4	1.05 ± 0.05
	Tris	265 ± 4	-61.6 ± 0.3	-24.8 ± 0.4	-36.9 ± 0.0	1.01 ± 0.04
	Phosphate	141 ± 8	-60.8 ± 0.3	-22.3 ± 0.1	-38.5 ± 0.1	1.07 ± 0.02

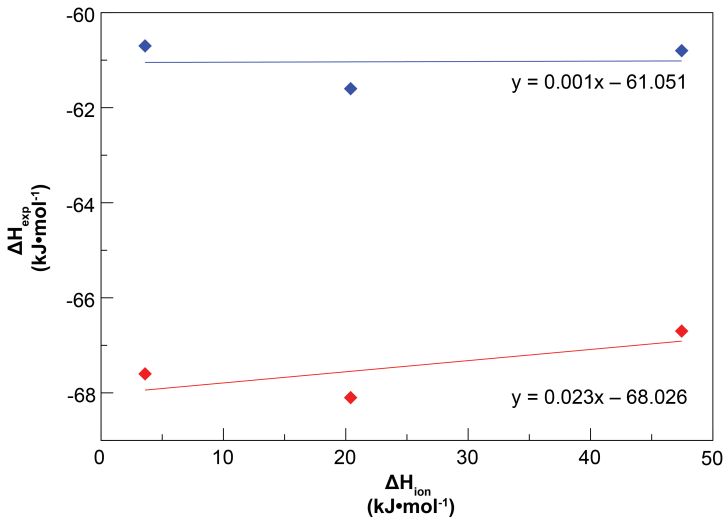

Figure S10. Plot of the experimentally determined enthalpies of NrtA binding to nitrate (red) and nitrite (blue) versus the ionization enthalpies of HEPES, Tris, and sodium phosphate buffers.[17] Each data set is fitted to a linear regression model.

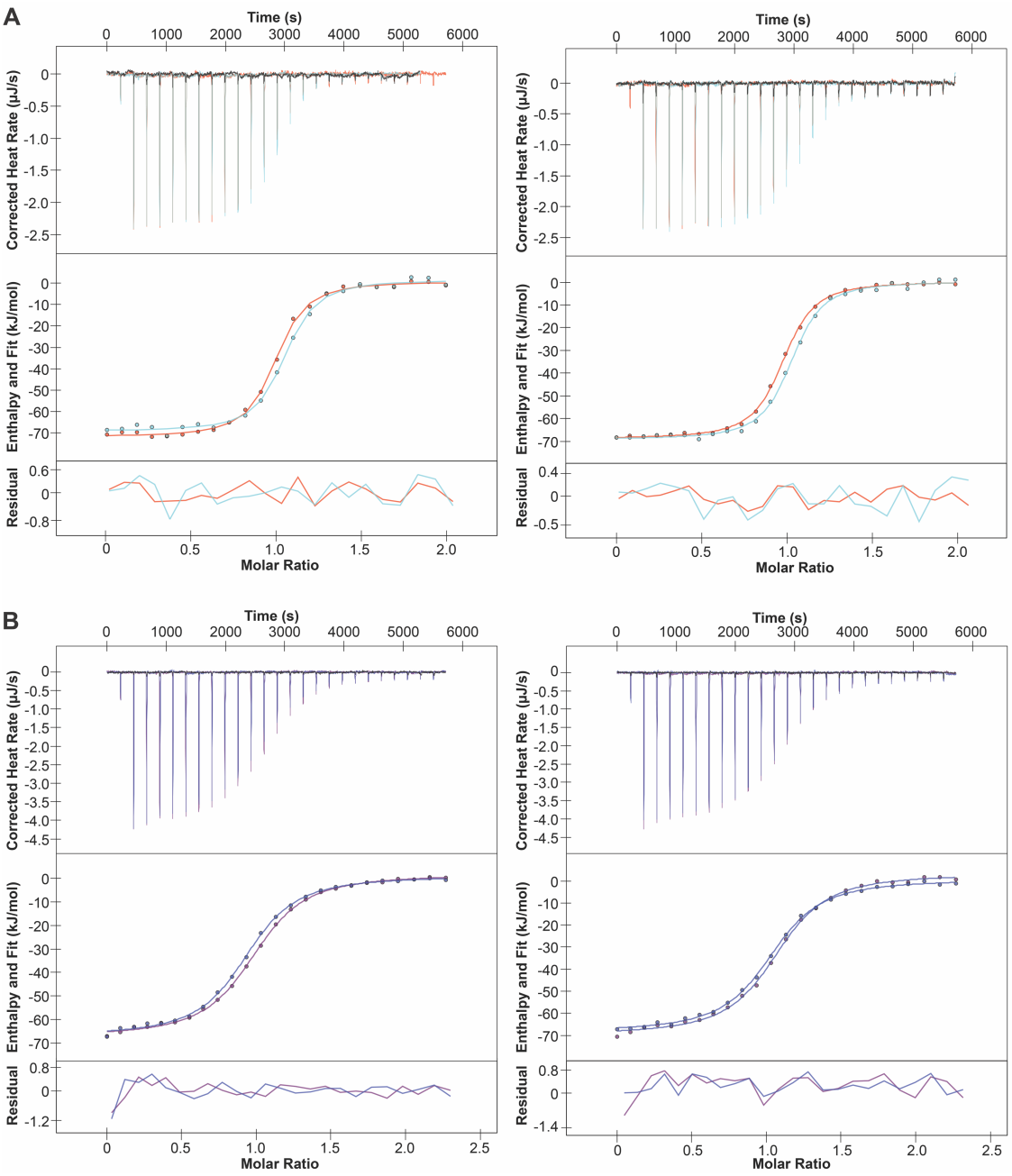

Figure S11. ITC analysis of (A) nitrate and (B) nitrite binding to NrtA at 37 °C. In each panel, the thermogram (top), isotherm (middle), and residual (bottom) plot is shown. Titration of the anion into the buffer alone (black) is included in each thermogram. All experiments were carried out in 20 mM HEPES buffer at pH 7.7 with 50 mM NaCl. Data is shown for two protein preparations, each measured in duplicate.

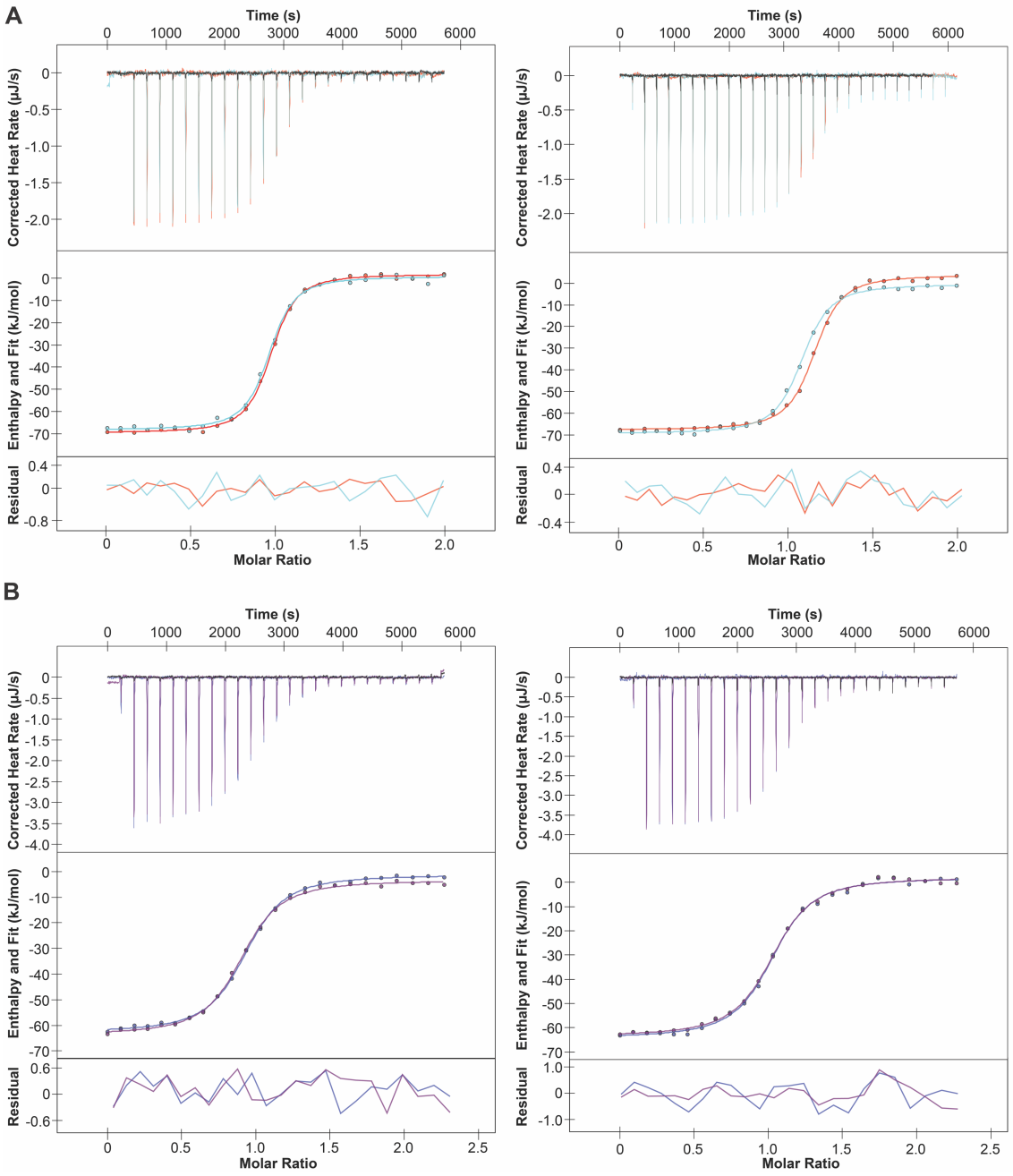

Figure S12. ITC analysis of (A) nitrate and (B) nitrite binding to NrtA at 30 °C. In each panel, the thermogram (top), isotherm (middle), and residual (bottom) plot is shown. Titration of the anion into the buffer alone (black) is included in each thermogram. All experiments were carried out in 20 mM HEPES buffer at pH 7.7 with 50 mM NaCl. Data is shown for two protein preparations, each measured in duplicate.

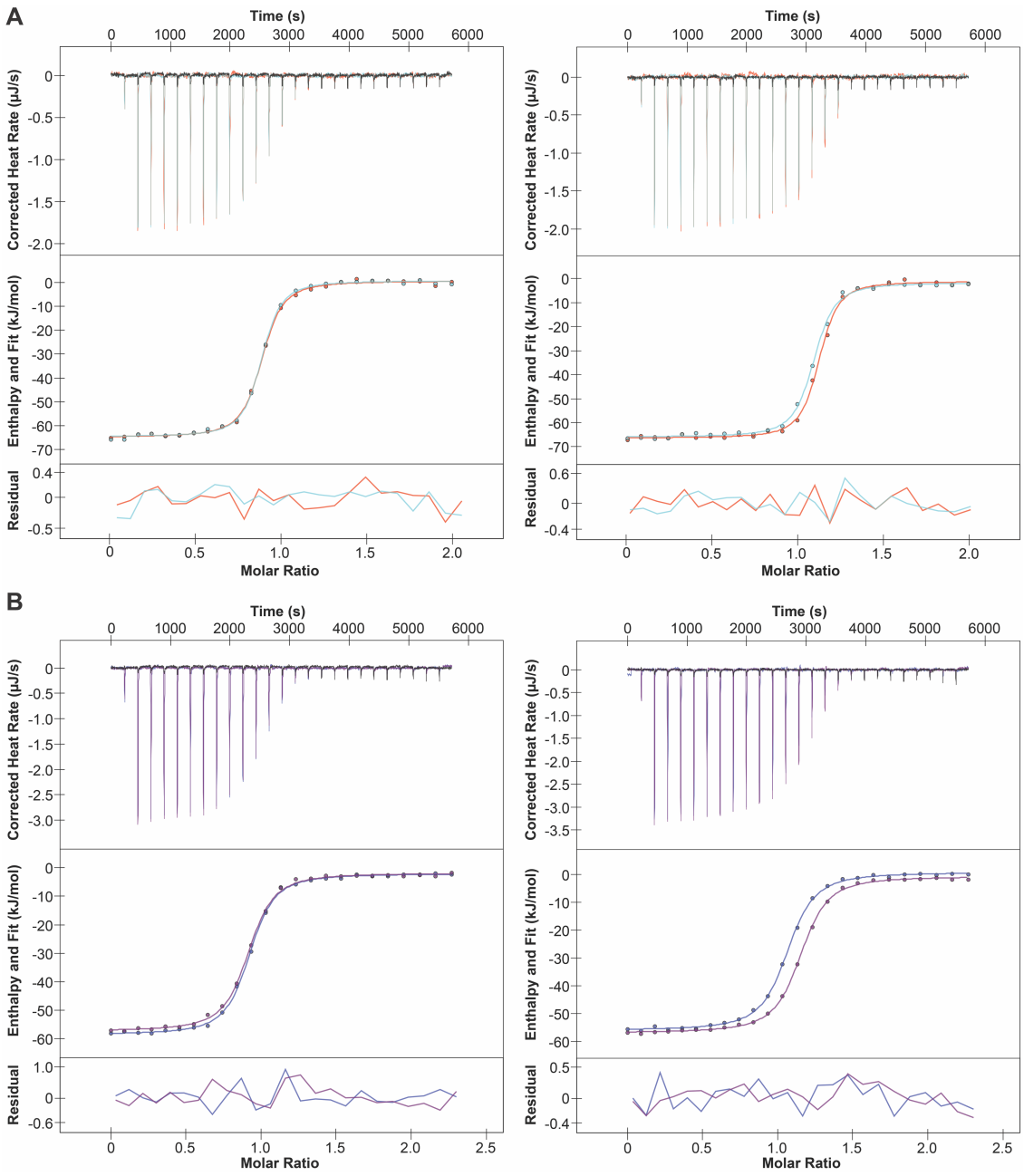

Figure S13. ITC analysis of (A) nitrate and (B) nitrite binding to NrtA at 10 °C. In each panel, the thermogram (top), isotherm (middle), and residual (bottom) plot is shown. Titration of the anion into the buffer alone (black) is included in each thermogram. All experiments were carried out in 20 mM HEPES buffer at pH 7.7 with 50 mM NaCl. Data is shown for two protein preparations, each measured in duplicate.

Table S5. Thermodynamic parameters for nitrate and nitrite binding to NrtA at different temperatures. All experiments were carried out in 20 mM HEPES buffer at pH 7.7 with 50 mM NaCl. The average of two protein preparations, each measured in duplicate with the standard deviation is shown.

Anion	Temperature (°C)	K_d (nM)	ΔH (kJ·mol^{-1})	TΔS (kJ·mol^{-1})	ΔG (kJ·mol^{-1})	ΔC_p (J·mol^{-1} K^{-1})	n
Nitrate	10	48 ± 8	-65.0 ± 0.7	-25.3 ± 0.5	-39.7 ± 0.4	-209 ± 45	1.02 ± 0.02
	20	57 ± 23	-68.0 ± 1.2	-27.2 ± 1.7	-40.7 ± 1.0		1.06 ± 0.02
	30	88 ± 9	-70.3 ± 2.5	-29.3 ± 2.3	-41.0 ± 0.3		1.06 ± 0.04
	37	115 ± 8	-70.5 ± 1.2	-29.3 ± 1.2	-41.2 ± 0.2		1.06 ± 0.05
Nitrite	10	187 ± 16	-56.3 ± 0.5	-19.8 ± 0.8	-36.5 ± 0.3	-463 ± 96	1.05 ± 0.06
	20	230 ± 33	-60.6 ± 1.1	-23.4 ± 1.5	-37.3 ± 0.4		1.05 ± 0.05
	30	473 ± 10	-63.5 ± 1.0	-26.8 ± 0.9	-36.7 ± 0.1		1.02 ± 0.02
	37	753 ± 12	-69.7 ± 2.9	-33.3 ± 2.8	-36.4 ± 0.0		1.05 ± 0.04

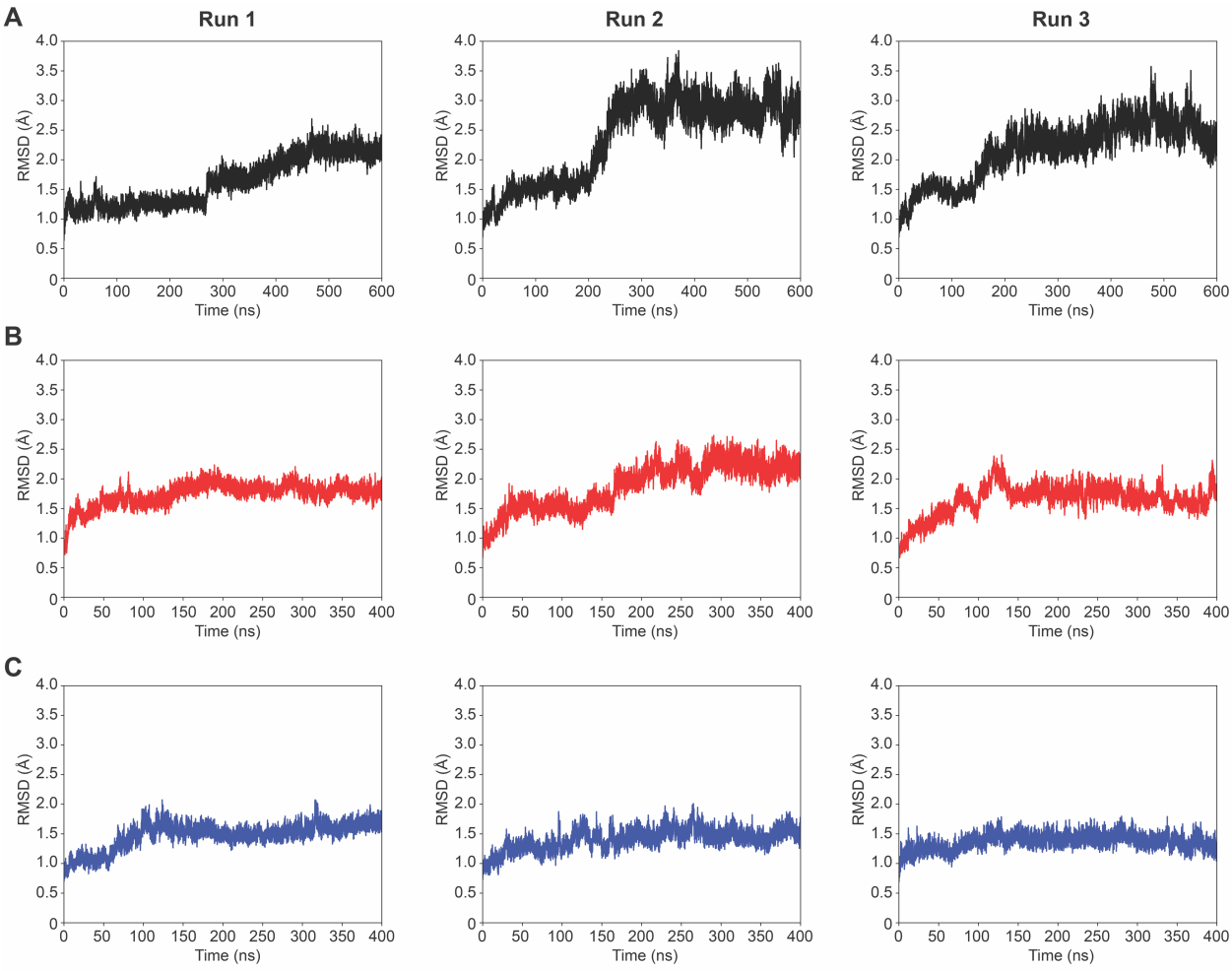

Figure S14. One-dimensional root mean square deviation (RMSD) plots of the Cα atoms from the MD simulations for (A) apo, (B) nitrate-bound, and (C) nitrite-bound NrtA.

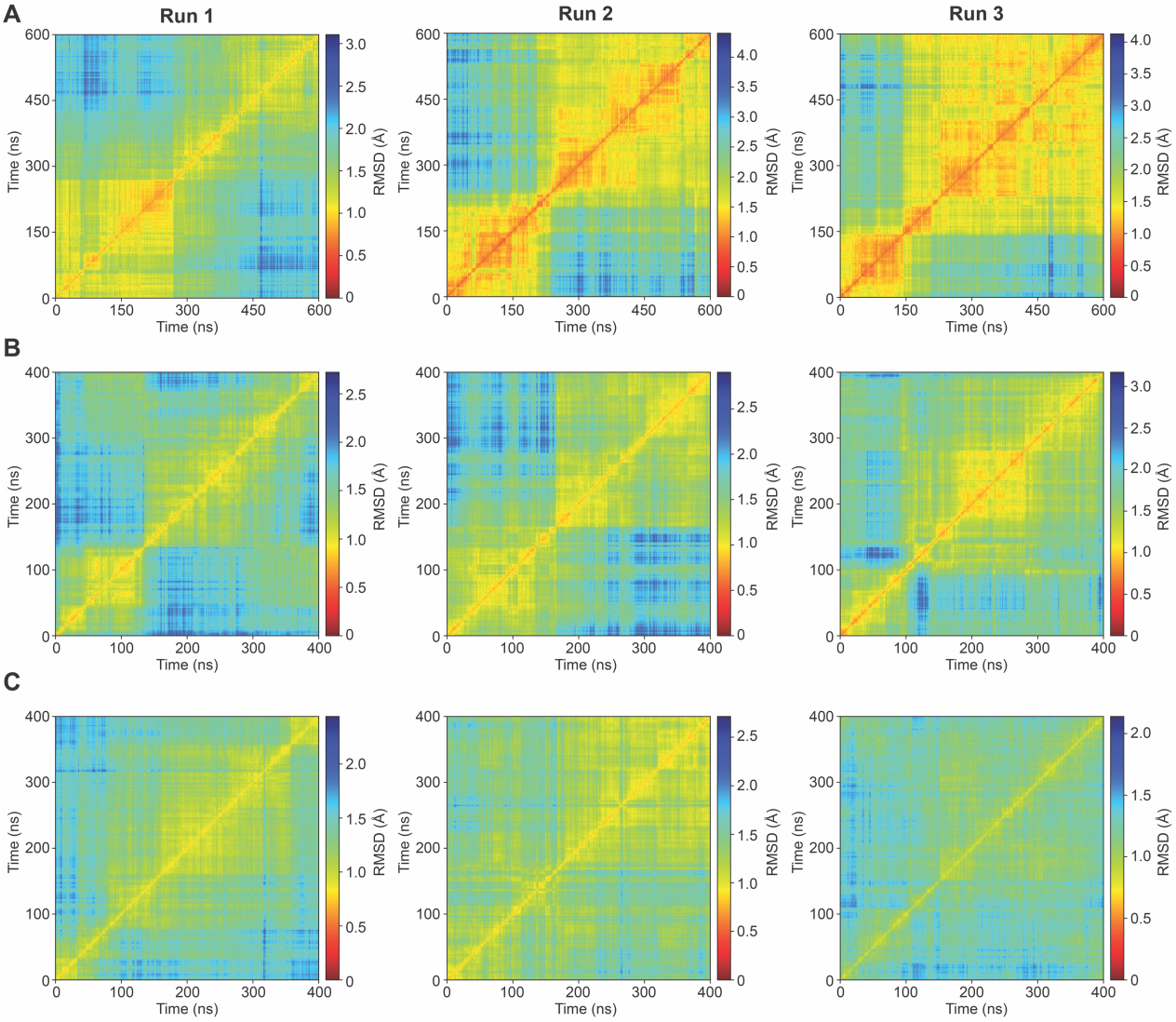

Figure S15. Two-dimensional RMSD plots of the Cα atoms from the MD simulations for (A) apo, (B) nitrate-bound, and (C) nitrite-bound NrtA.

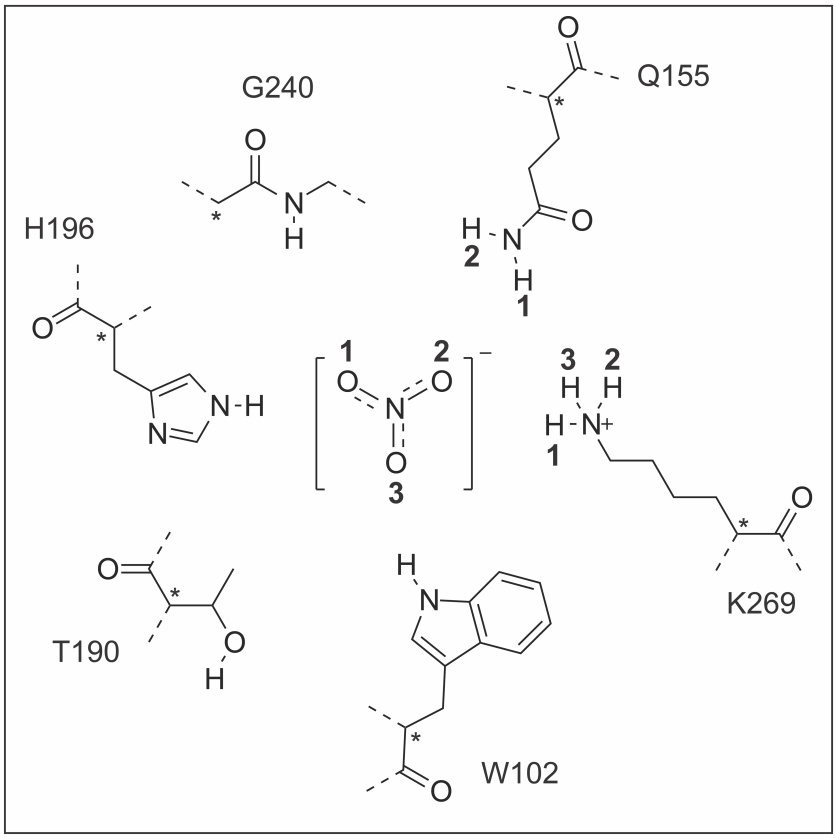

Figure S16. Schematic representation of the nitrate binding site in NrtA (PDB ID: 2G29). Polar residues within 4 Å of the nitrate ion are labelled with the single letter amino acid abbreviation and corresponding sequence number. For the distance analysis of the MD simulations in Figures S17–S30, atoms are numbered and Cα atoms are labelled with an asterisk (*), where appropriate.

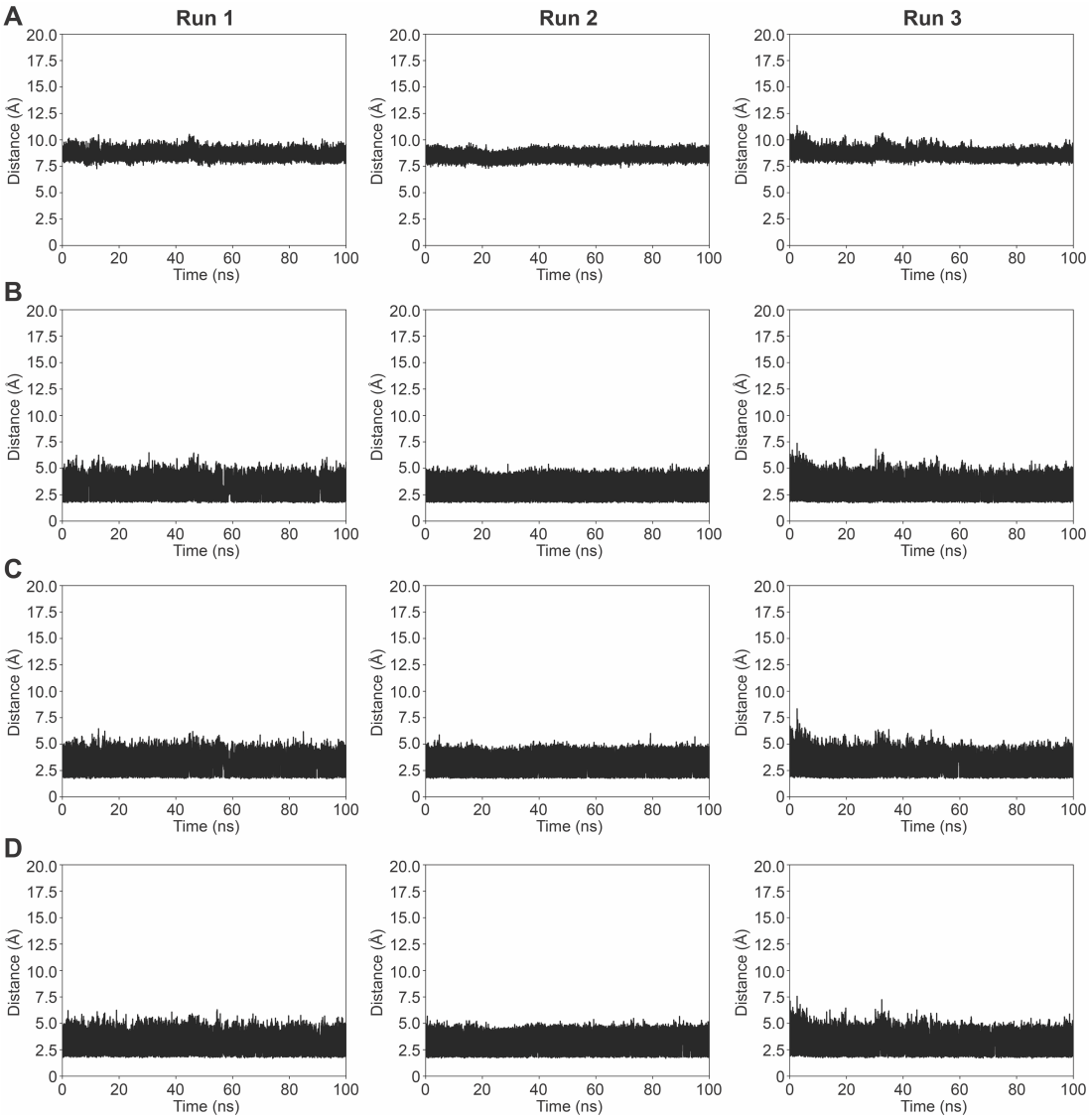

Figure S17. Analysis of the distances between residue W102 and nitrate as a function of time for three independent MD simulations. Distances are shown for (A) Cα of W45 and N of nitrate, (B) NH of W102 and O1 of nitrate, (C) NH of W102 and O2 of nitrate, and (D) NH of W102 and O3 of nitrate.

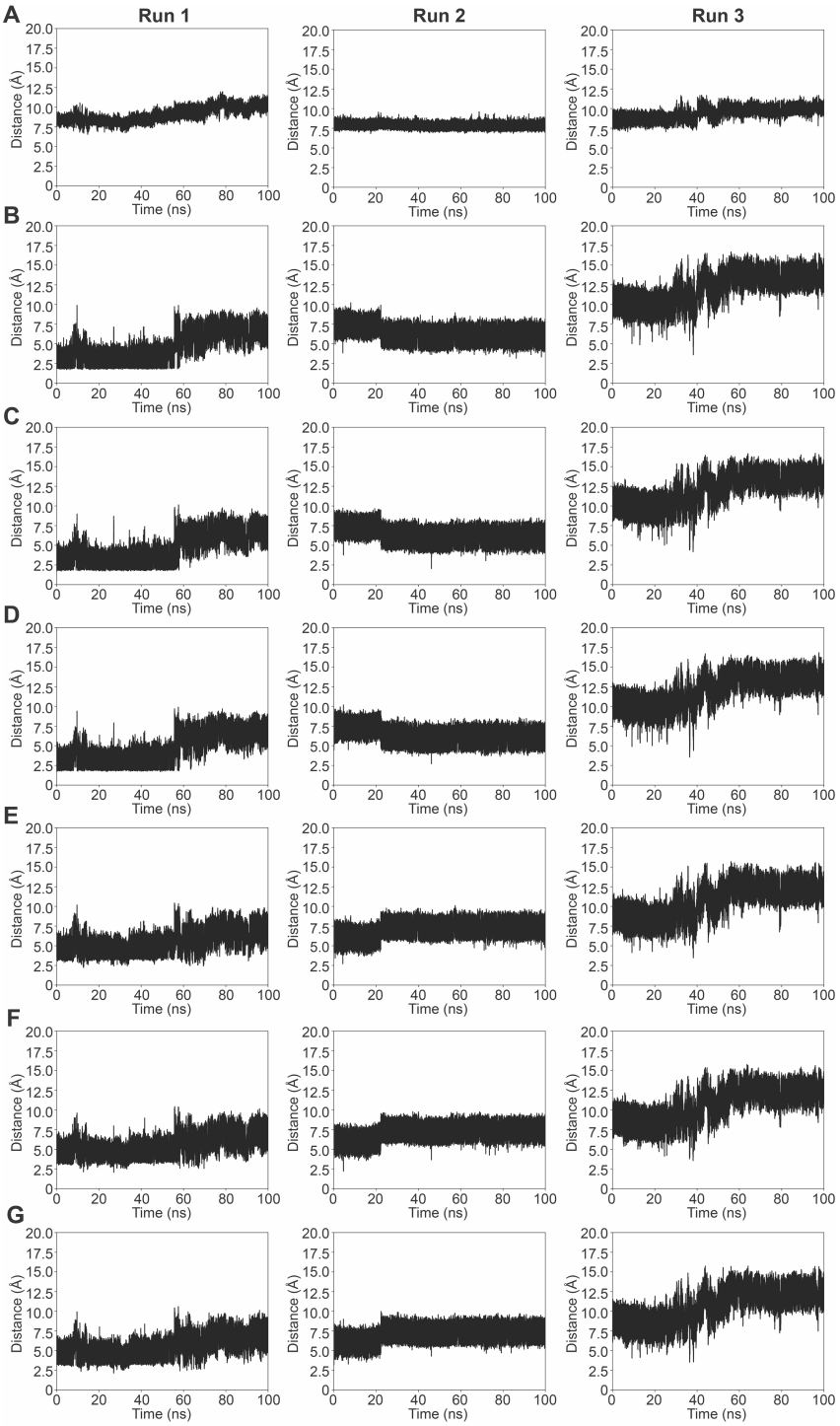

Figure S18. Analysis of the distances between residue Q155 and nitrate as a function of time for three independent MD simulations. Distances are shown for (A) Cα of Q155 and N of nitrate, (B) NH1 of Q155 and O1 of nitrate, (C) NH1 of Q155 and O2 of nitrate, (D) NH1 of Q155 and O3 of nitrate, (E) NH2 of Q155 and O1 of nitrate, (F) NH2 of Q155 and O2 of nitrate, and (G) NH2 of Q155 and O3 of nitrate.

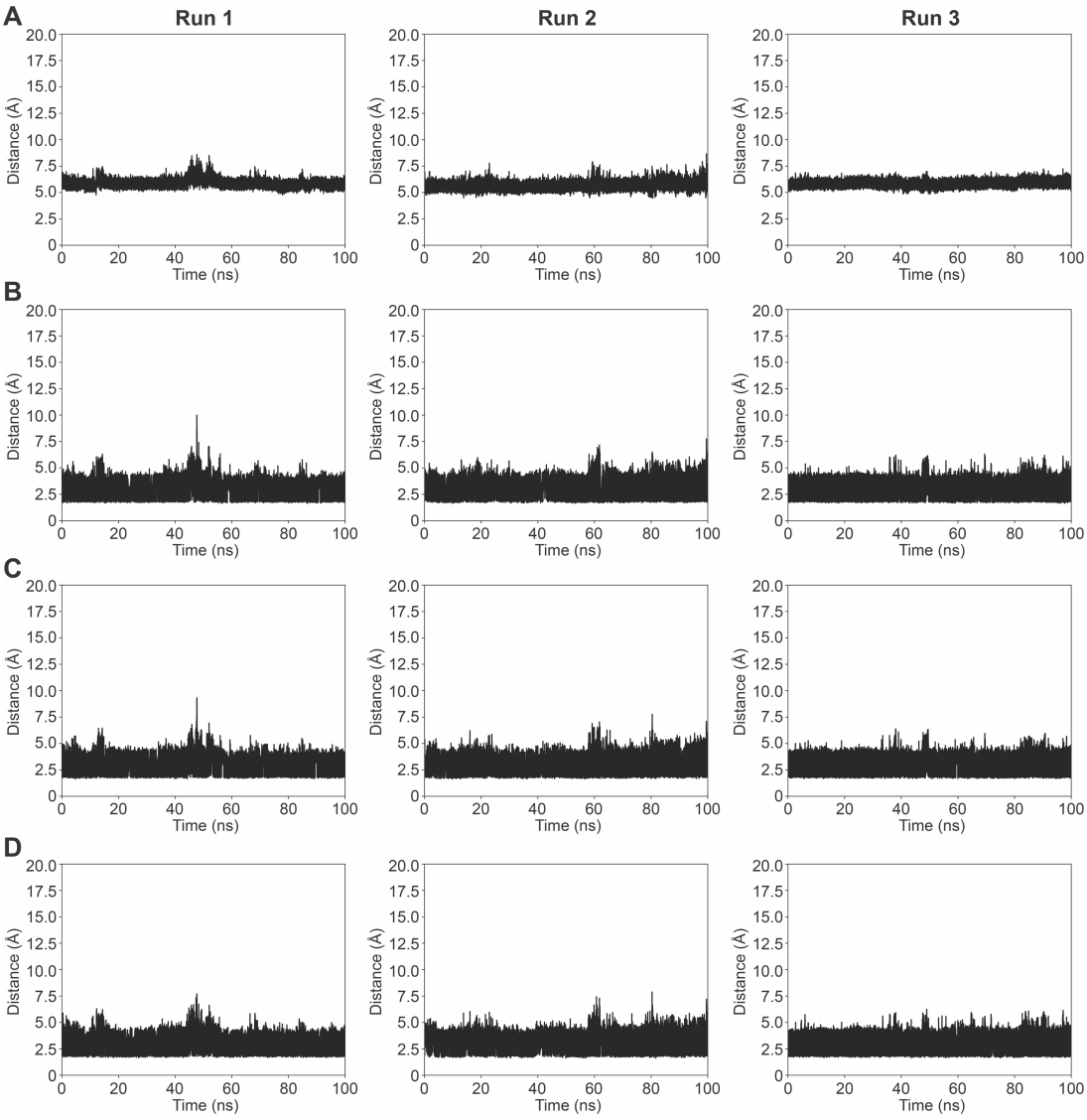

Figure S19. Analysis of the distances between residue T190 and nitrate as a function of time for three independent MD simulations. Distances are shown for (A) Cα of T190 and N of nitrate, (B) OH of T190 and O1 of nitrate, (C) OH of T190 and O2 of nitrate, and (D) OH of T190 and O3 of nitrate.

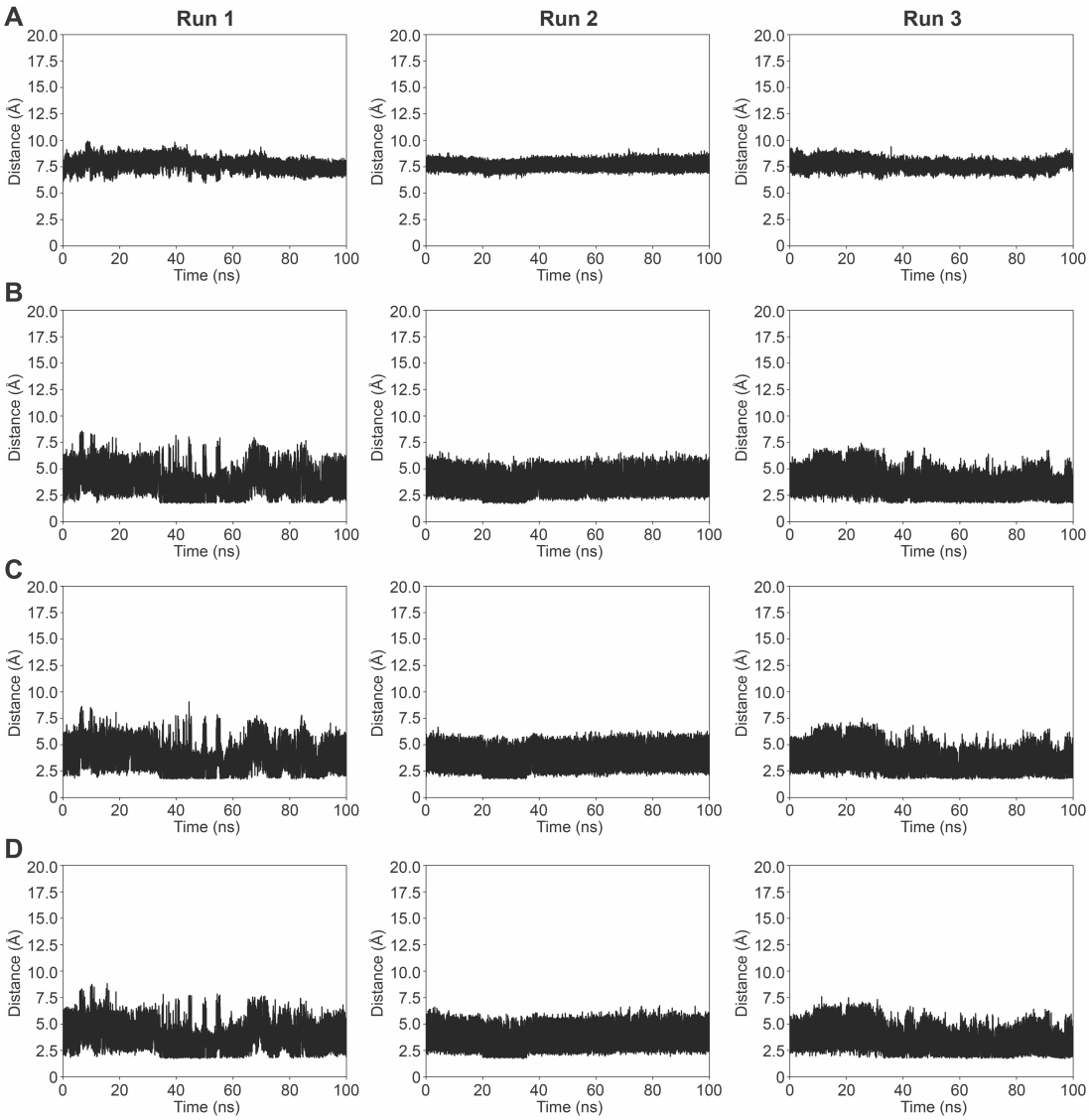

Figure S20. Analysis of the distances between residue H196 and nitrate as a function of time for three independent MD simulations. Distances are shown for (A) Cα of H196 and N of nitrate, (B) NH of H196 and O1 of nitrate, (C) NH of H196 and O2 of nitrate, and (D) NH of H196 and O3 of nitrate.

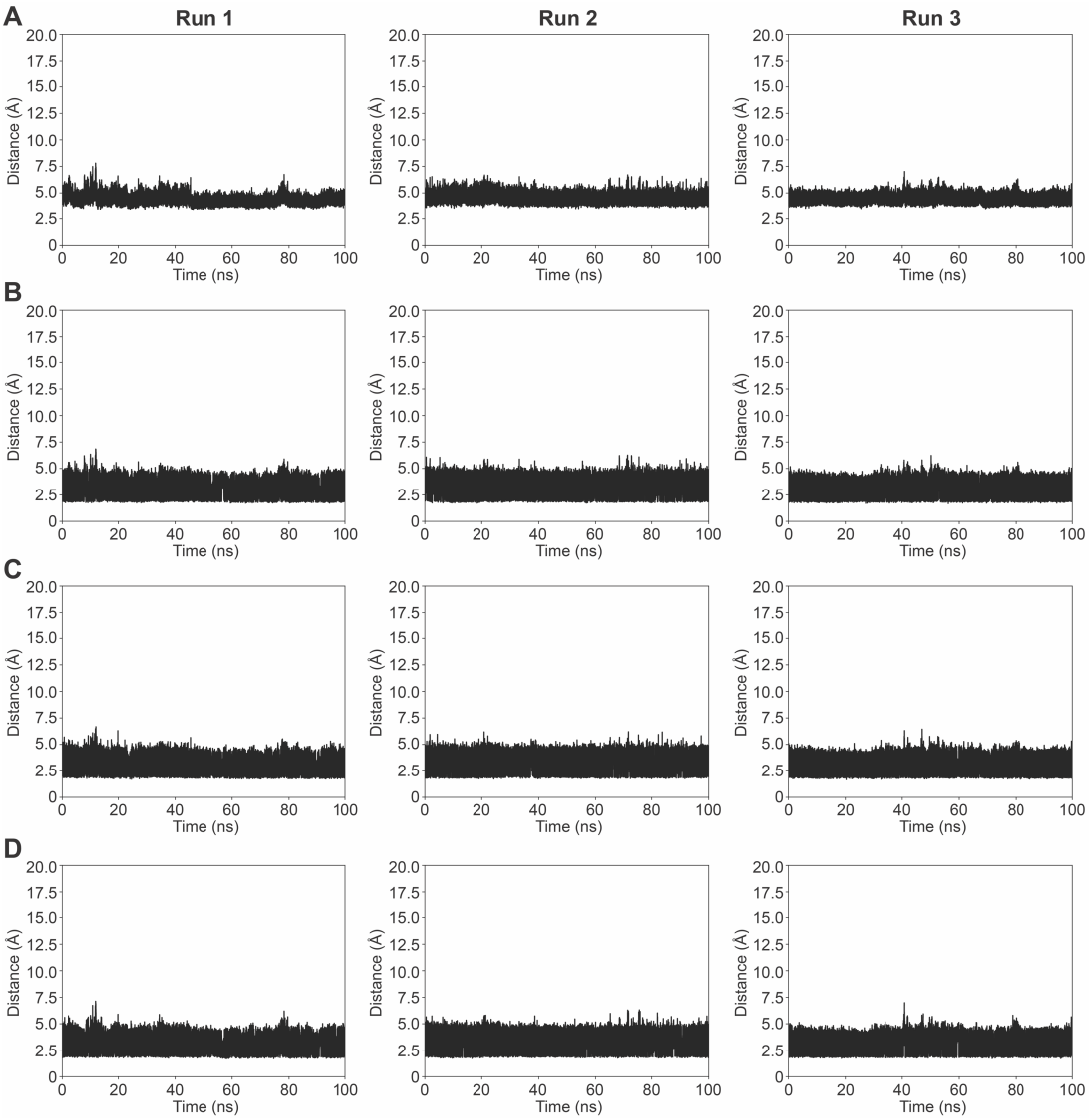

Figure S21. Analysis of the distances between residue G240 and nitrate as a function of time for three independent MD simulations. Distances are shown for (A) Cα of G240 and N of nitrate, (B) NH of G240 and O1 of nitrate, (C) NH of G240 and O2 of nitrate, and (D) NH of G240 and O3 of nitrate.

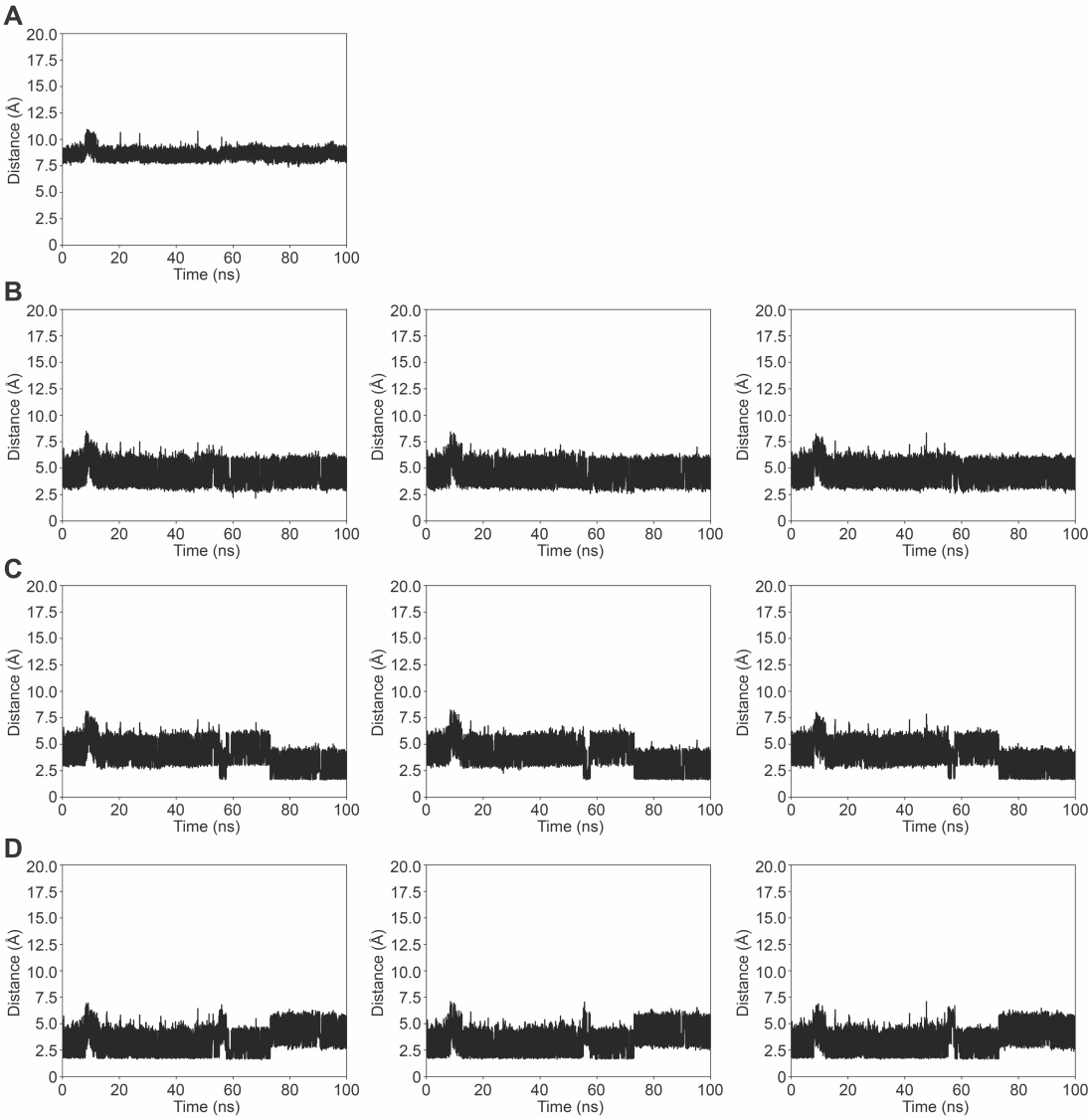

Figure S22. Analysis of the distances between residue K269 and nitrate as a function of time for MD simulation run 1. Distances are shown for (A) Cα of K269 and N of nitrate, (B) NH1 of K269 and O1, O2, O3 of nitrate (left to right), (C) NH2 of K269 and O1, O2, O3 of nitrate (left to right), and (D) NH3 of K269 and O1, O2, O3 of nitrate (left to right).

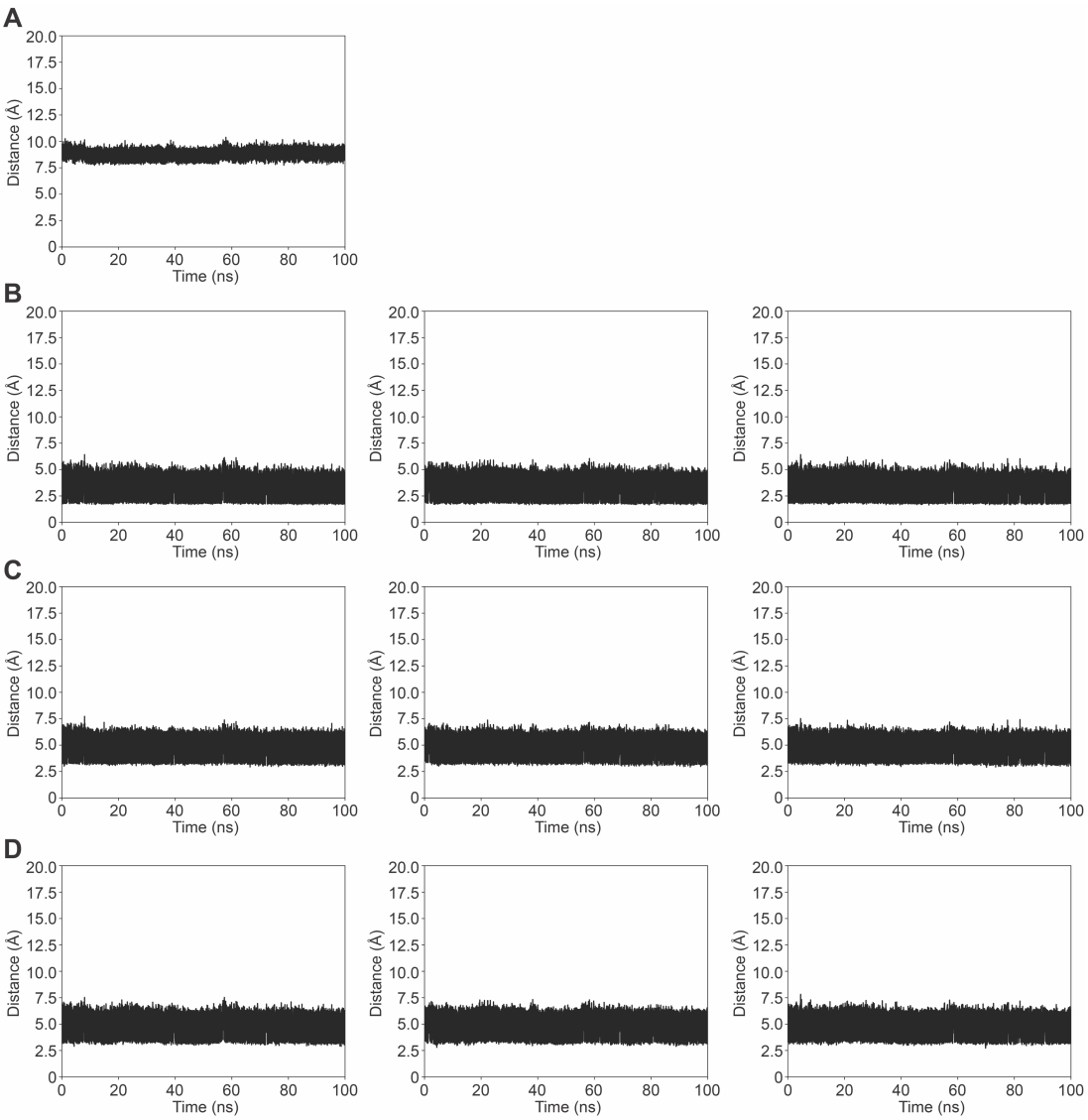

Figure S23. Analysis of the distances between residue K269 and nitrate as a function of time for MD simulation run 2. Distances are shown for (A) Cα of K269 and N of nitrate, (B) NH1 of K269 and O1, O2, O3 of nitrate (left to right), (C) NH2 of K269 and O1, O2, O3 of nitrate (left to right), and (D) NH3 of K269 and O1, O2, O3 of nitrate (left to right).

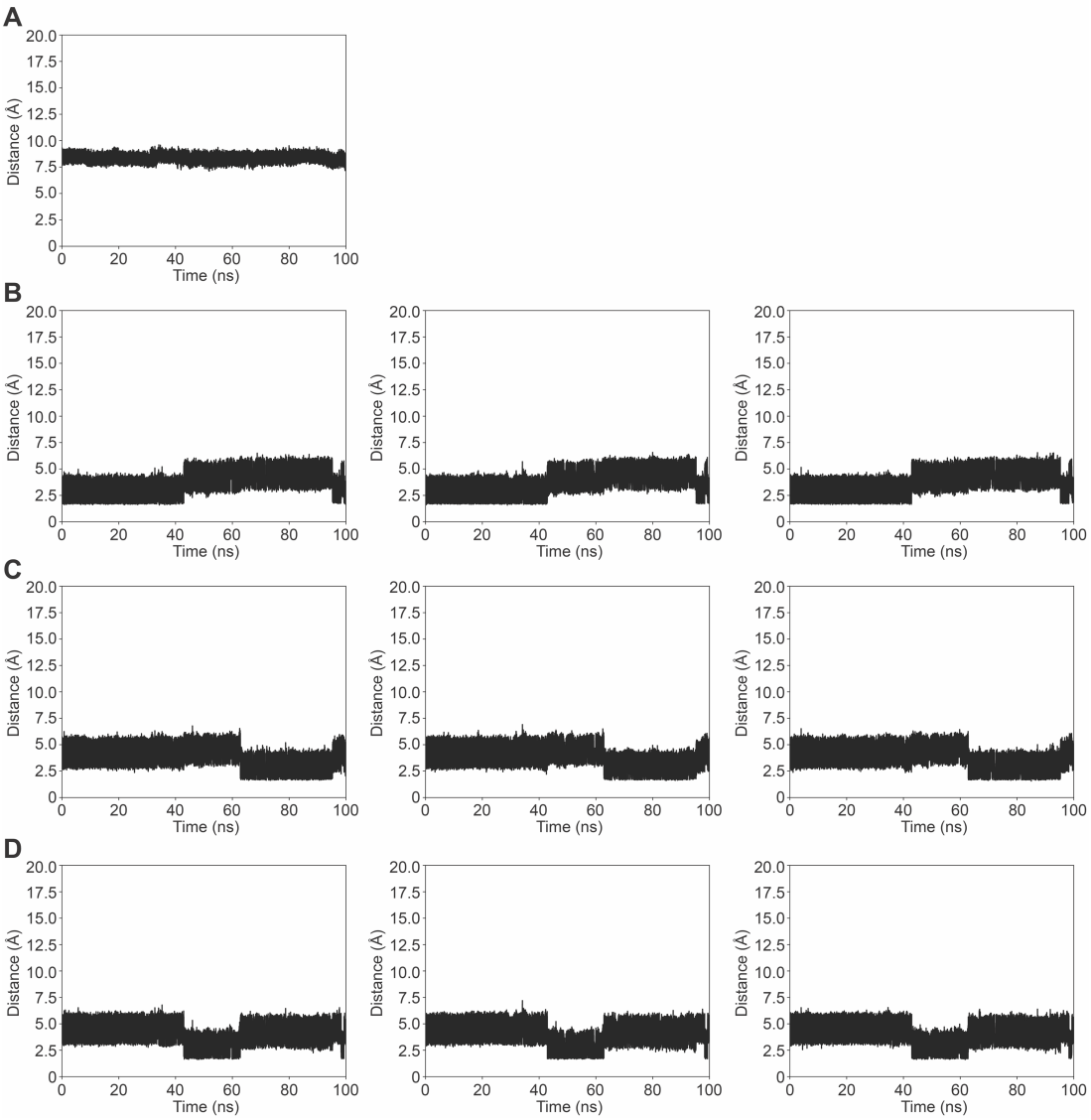

Figure S24. Analysis of the distances between residue K269 and nitrate as a function of time for MD simulation run 3. Distances are shown for (A) Cα of K269 and N of nitrate, (B) NH1 of K269 and O1, O2, O3 of nitrate (left to right), (C) NH2 of K269 and O1, O2, O3 of nitrate (left to right), and (D) NH3 of K269 and O1, O2, O3 of nitrate (left to right).

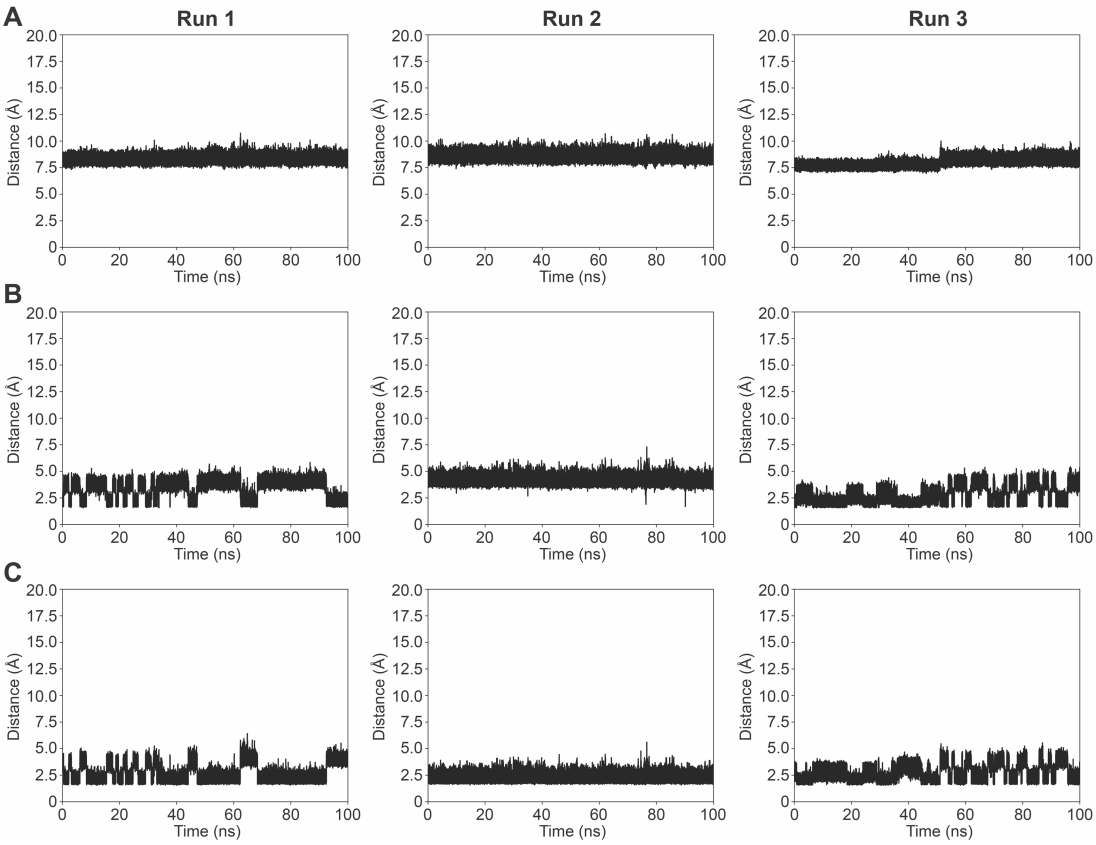

Figure S25. Analysis of the distances between residue W102 and nitrite as a function of time for three independent MD simulations. Distances are shown for (A) Cα of W102 and N of nitrite, (B) NH of W102 and O1 of nitrite, and (C) NH of W102 and O2 of nitrite.

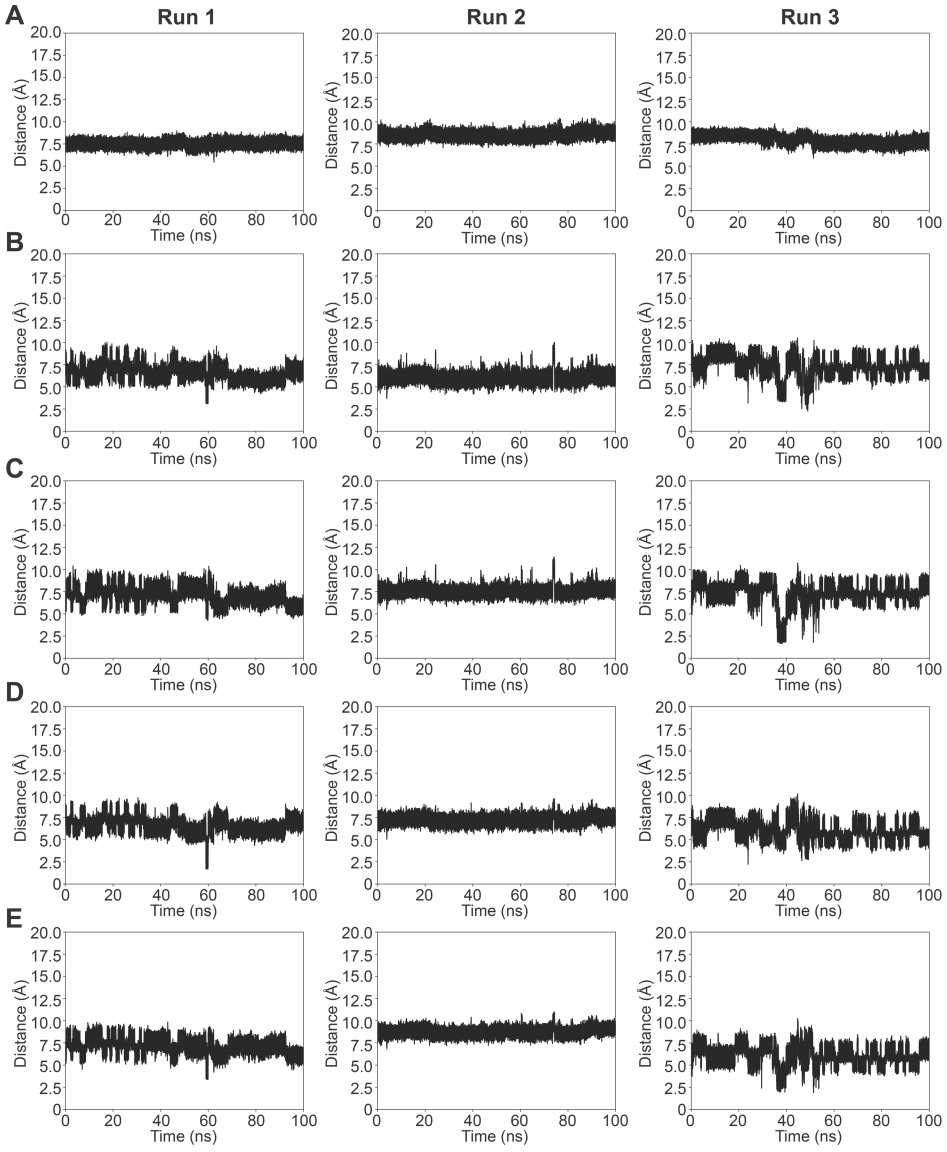

Figure S26. Analysis of the distances between residue Q155 and nitrite as a function of time for three independent MD simulations. Distances are shown for (A) Cα of Q155 and N of nitrite, (B) NH1 of Q155 and O1 of nitrite, (C) NH1 of Q155 and O2 of nitrite, (D) NH2 of Q155 and O1 of nitrite, and (E) NH2 of Q155 and O2 of nitrite.

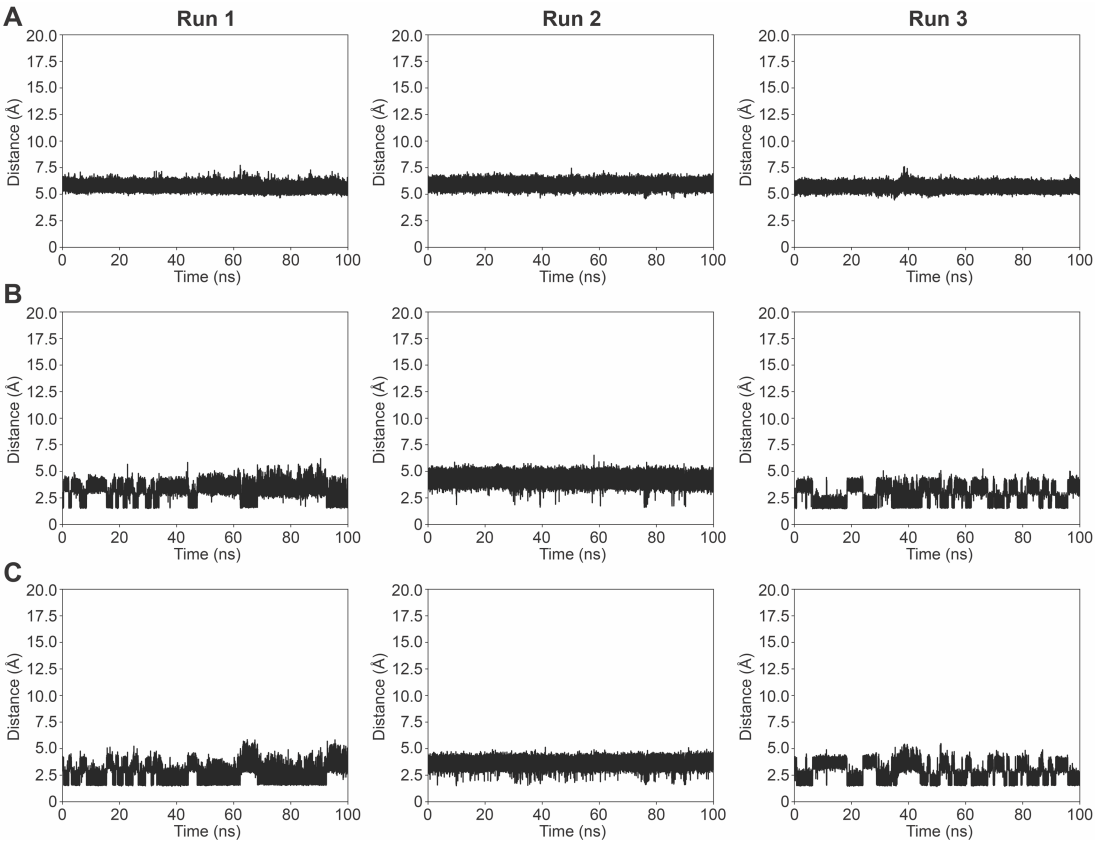

Figure S27. Analysis of the distances between residue T190 and nitrite as a function of time for three independent MD simulations. Distances are shown for (A) Cα of T190 and N of nitrite, (B) OH of T190 and O1 of nitrite, and (C) OH of T190 and O2 of nitrite.

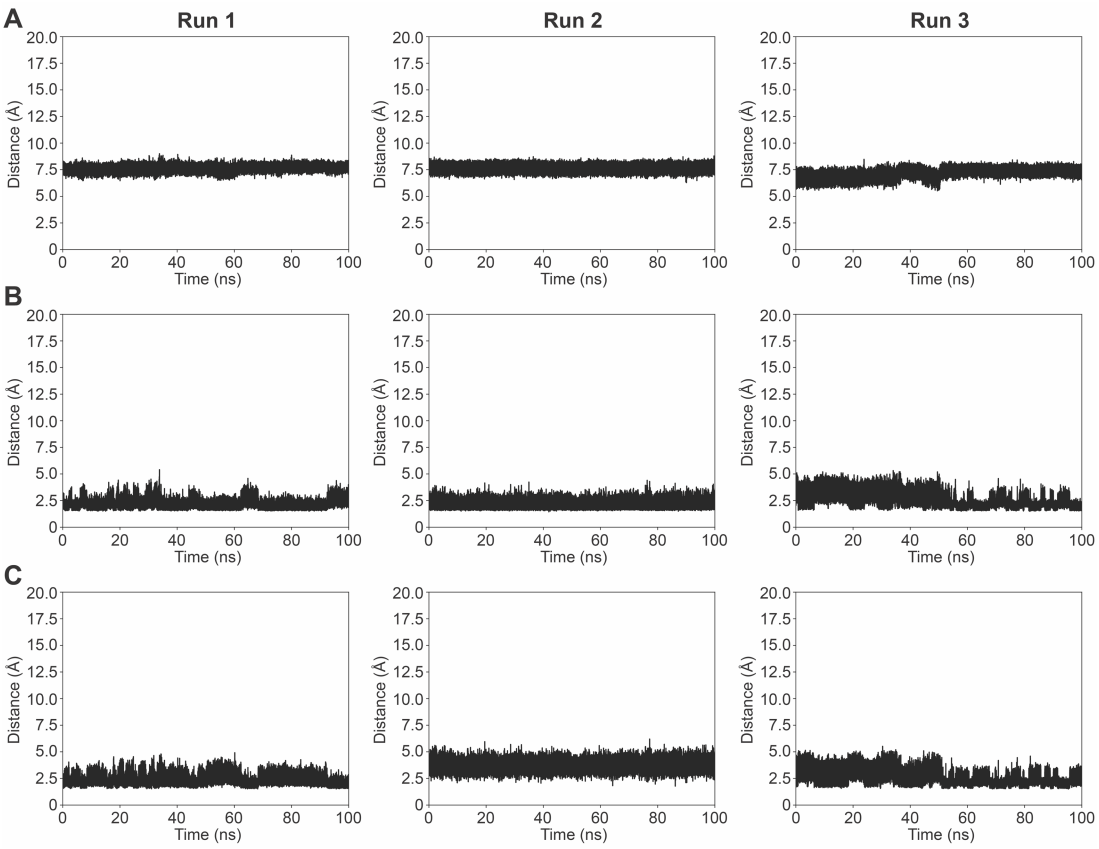

Figure S28. Analysis of the distances between residue H196 and nitrite as a function of time for three independent MD simulations. Distances are shown for (A) Cα of H196 and N of nitrite, (B) NH of H196 and O1 of nitrite, and (C) NH of H196 and O2 of nitrite.

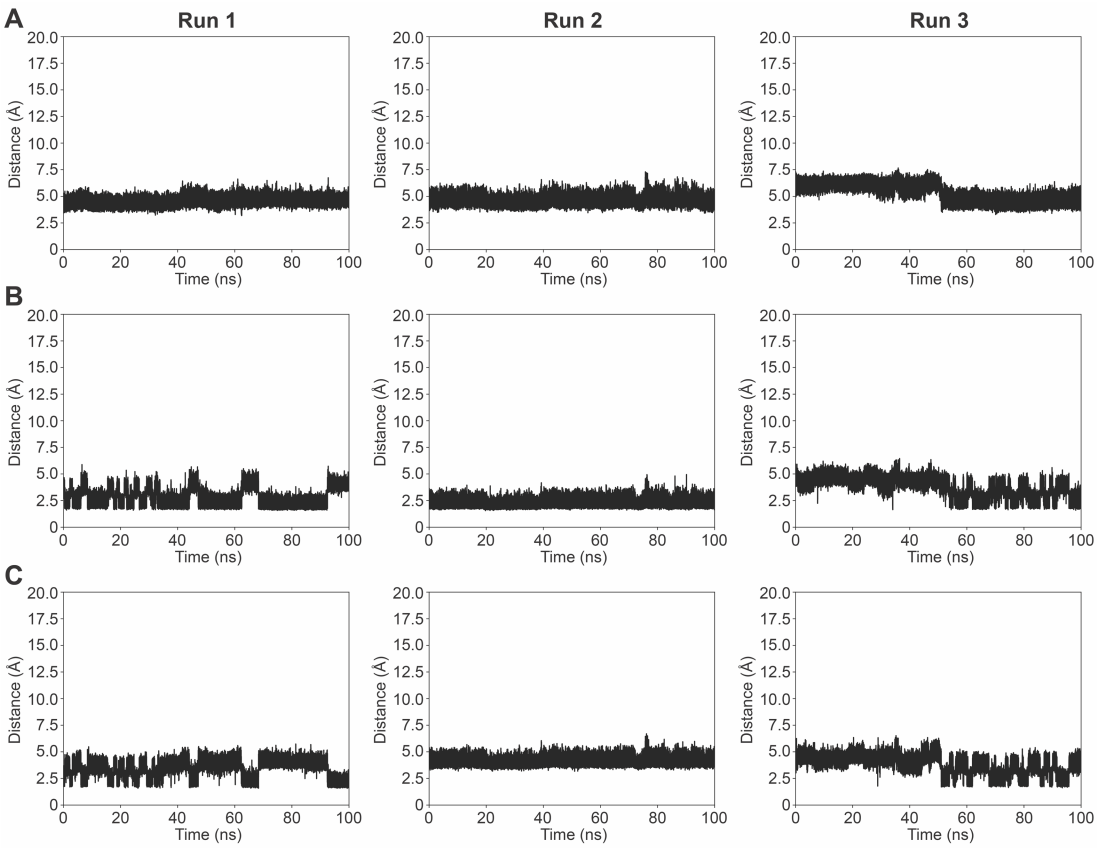

Figure S29. Analysis of the distances between residue G240 and nitrite as a function of time for three independent MD simulations. Distances are shown for (A) Cα of G240 and N of nitrite, (B) NH of G240 and O1 of nitrite, and (C) NH of G240 and O2 of nitrite.

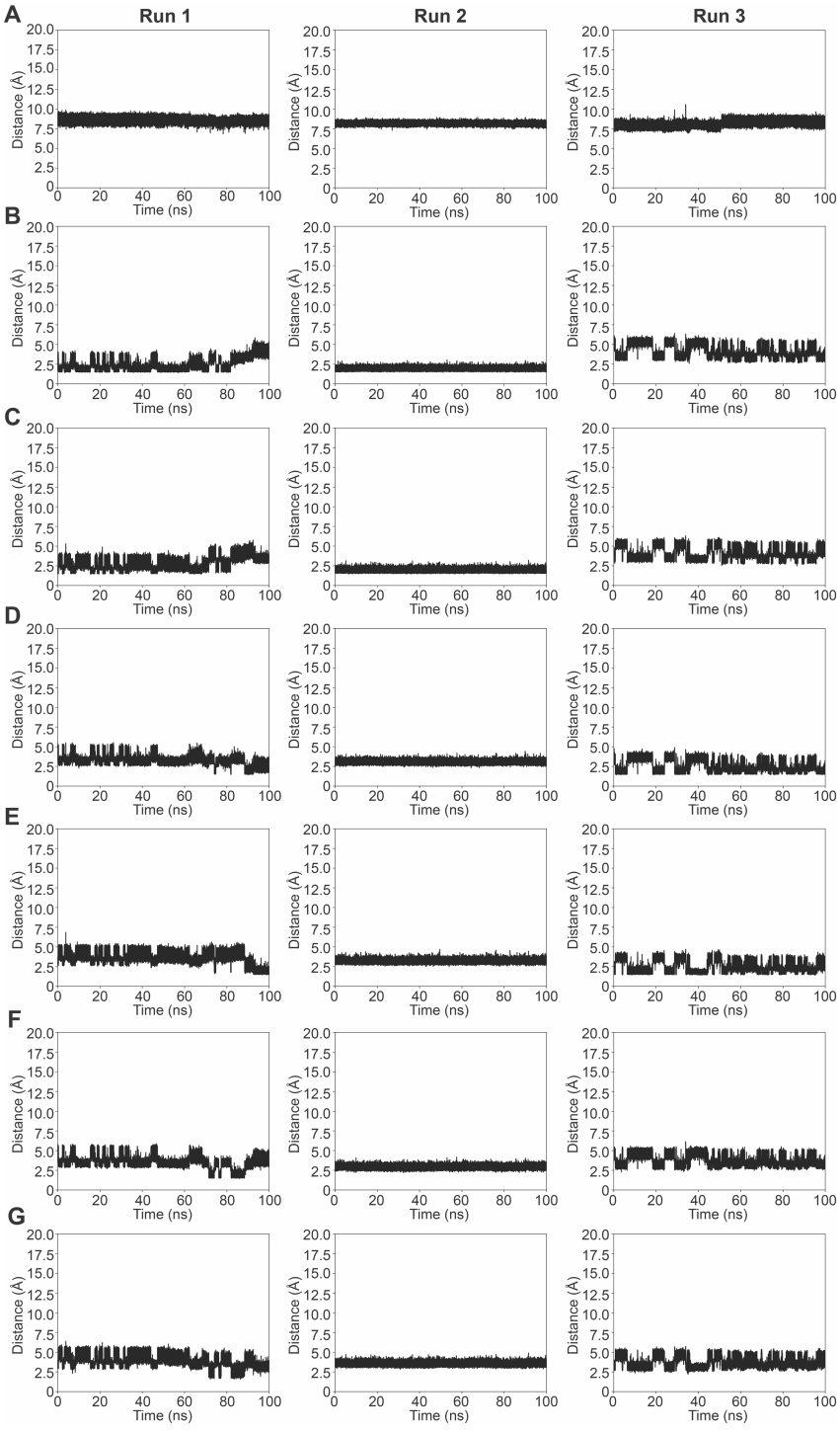

Figure S30. Analysis of the distances between residue K269 and nitrite as a function of time for three independent MD simulations. Distances are shown for (A) Cα of K269 and N of nitrite, (B) NH1 of K269 and O1 of nitrite, (C) NH1 of K269 and O2 of nitrite, (D) NH2 of K269 and O1 of nitrite, (E) NH2 of K269 and O2 of nitrite, (F) NH3 of K269 and O1 of nitrite, and (G) NH3 of K269 and O2 of nitrite.

Table S6. Distances between nitrate atoms and binding residues. All distances were measured by PyMOL according to the crystal structure of NrtA bound to nitrate (PDB ID: 2G29).

Residue		Distance (Å)			
		N	O1	O2	O3
W102	Cα	8.7			
	NH	3.2	3.6	4.3	1.9
Q155	Cα	7.0			
	NH1	4.6	3.7	2.3	4.5
	NH2	4.4	3.7	2.7	4.6
T190	Cα	6.3			
	OH	5.7	5.6	6.7	4.9
H196	Cα	6.9			
	NH	3.6	2.6	2.6	2.4
G240	Cα	4.0			
	NH	2.4	2.1	2.3	3.9
K269	Cα	8.6			
	NH1	4.4	3.1	2.3	3.1
	NH2	5.9	4.6	3.6	4.7
	NH3	4.9	3.7	2.5	3.9

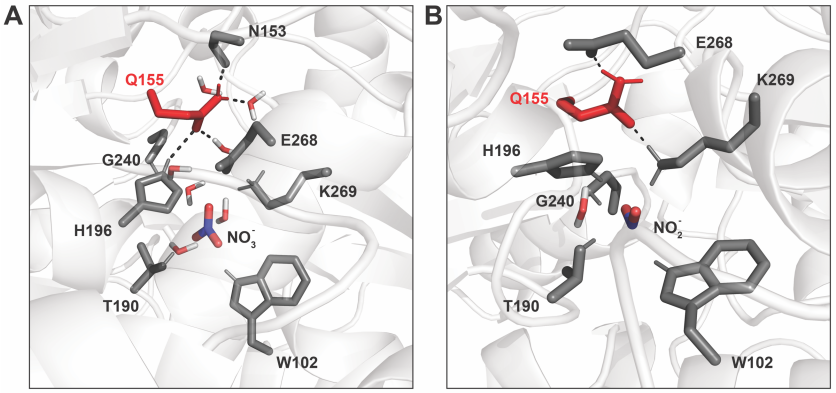

Figure S31. Representative snapshot of (A) nitrate and (B) nitrite-bound forms reveal that Q155 can interact with other residues in NrtA. Polar residues as defined in Figure 1 are shown as sticks in grey or red. Water, nitrate, and nitrite molecules are shown as sticks with all oxygen atoms in red, nitrogen atoms in blue, and hydrogen atoms in white. All hydrogen bonding or electrostatic interactions with Q155 are shown with dashed lines. Each residue is labelled with the single letter amino acid abbreviation and corresponding sequence number.

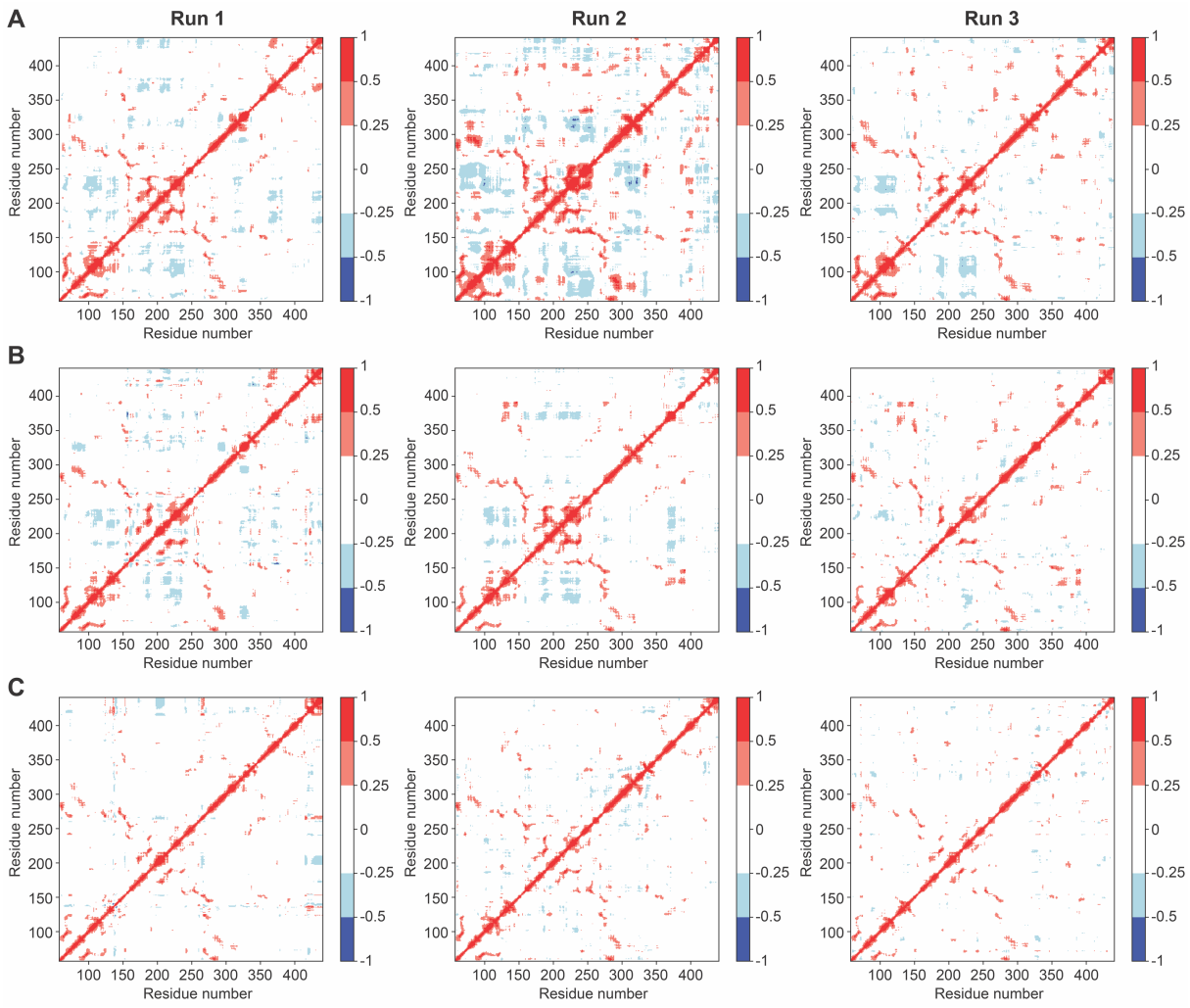

Figure S32. Dynamic cross-correlation matrix analysis (DCCM) of the Cα atoms from each MD simulation for (A) apo, (B) nitrate-bound and (C) nitrite-bound NrtA. Cross-correlation coefficients ranging from −1, -0.5, -0.25, 0, 0.25, 0.5, 1 are shown in each color scale bar.

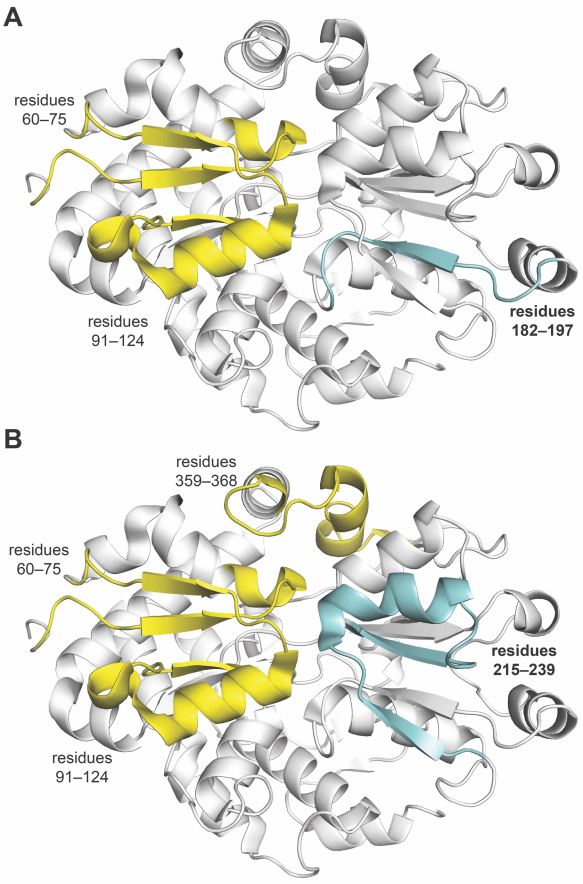

Figure S33. Two major anticorrelated regions (A) residues 182–197 and (B) residues 215–239 revealed in Figure S32 are coloured in cyan. The corresponding anticorrelated residues are labelled and coloured in yellow.

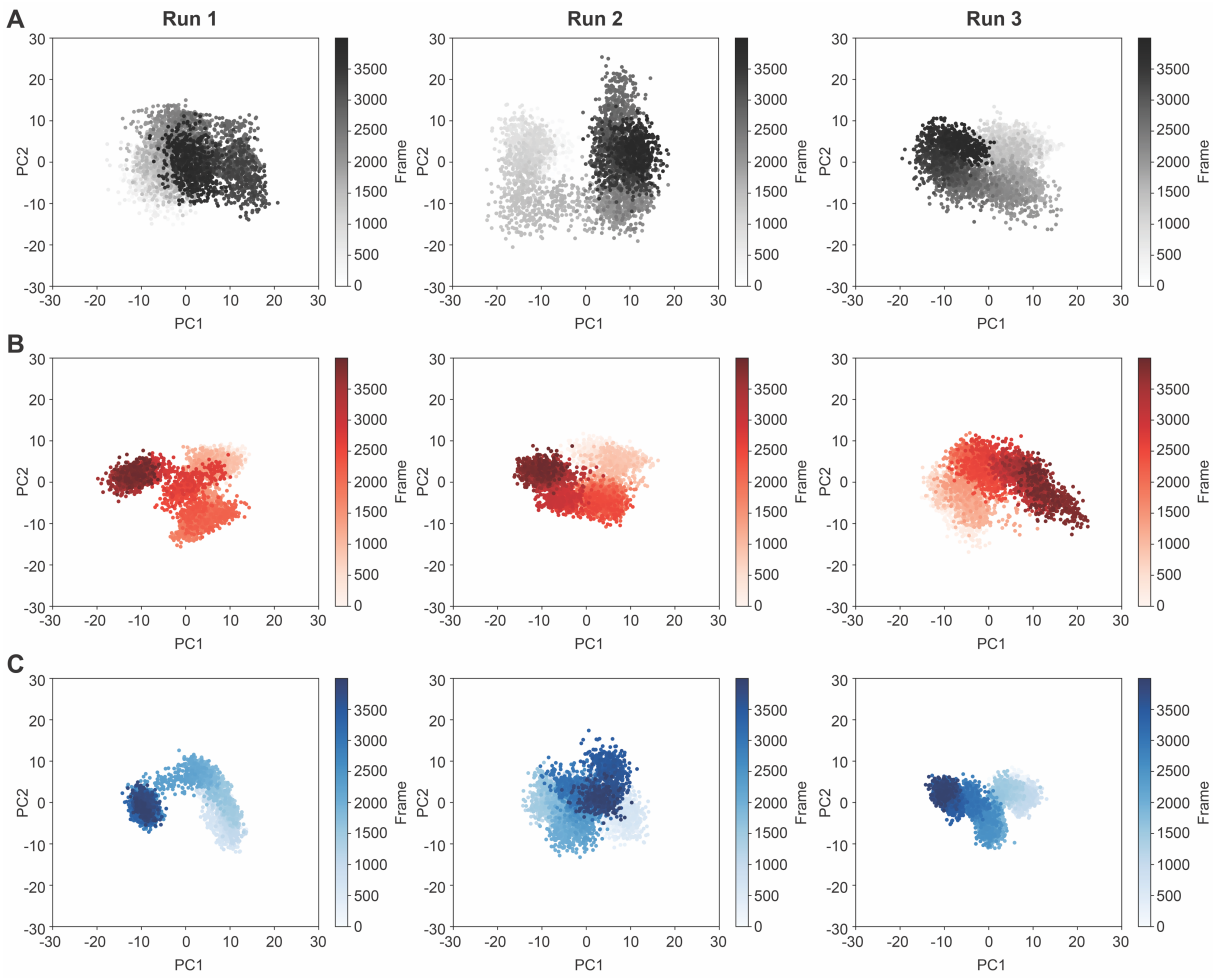

Figure S34. Conformations sampled during the MD trajectories plotted along the first and second principal components (PC1 & PC2) for (A) apo, (B) nitrate-bound, and (C) nitrite-bound NrtA.

Table S7. The time-averaged SASAs of apo, nitrate-bound, and nitrite-bound NrtA calculated from MD simulations.

	SASA (Å^2)			
	Run 1	**Run 2**	**Run 3**	**Average**
Apo	17421.0	17908.9	17412.2	17580.7
Nitrate	17014.4	16976.5	17297.0	17096.0
Nitrite	17003.8	16765.3	16611.9	16793.7

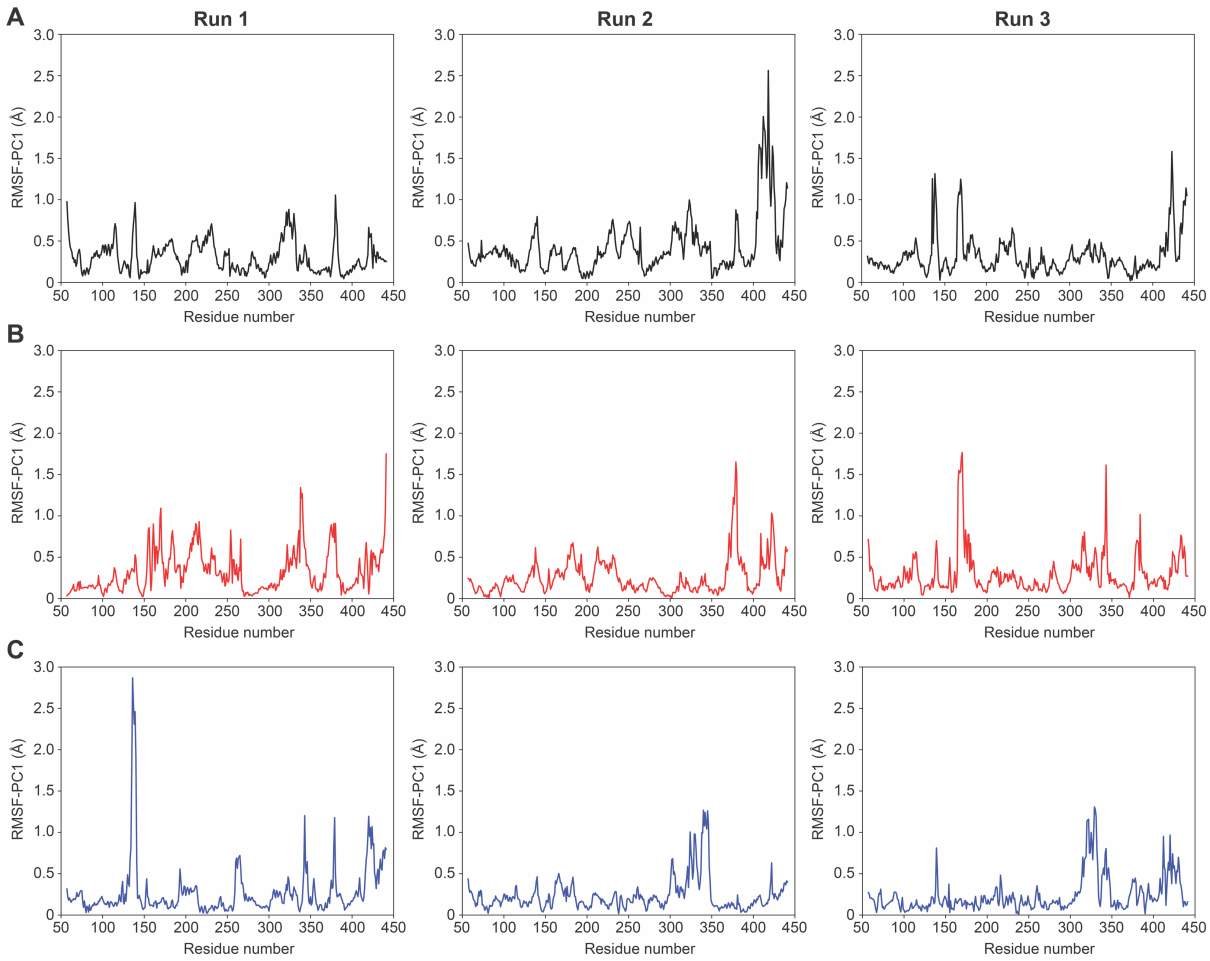

Figure S35. Root mean square fluctuation (RMSF) plots of the Cα backbone along the PC-1 direction from the MD simulations for (A), apo, (B), nitrate-bound, and (C) nitrite-bound NrtA.

Table S8. Residues that have Cα RMSF values greater than 0.4 Å along the PC1 direction in each MD simulation for apo NrtA. Residues that are common to all three MD simulations are highlighted in red.

Residue number											
Run 1				Run 2					Run 3		
57	117	221	328	57	225	304	339	421	114	220	421
58	136	222	329	73	226	305	340	422	115	221	422
59	137	223	330	88	227	306	341	423	116	224	423
60	138	224	331	89	228	307	342	424	135	227	424
61	139	225	332	90	229	308	343	425	136	228	425
100	140	226	343	92	230	309	345	426	137	229	426
101	161	227	344	100	231	310	347	427	138	230	433
103	167	228	378	103	232	311	348	429	139	231	434
104	173	229	379	104	233	312	378	430	140	232	435
107	174	230	380	132	234	313	379	433	141	233	436
108	176	231	381	133	235	314	380	434	165	234	437
110	177	232	382	134	241	315	381	435	166	252	438
111	178	233	420	135	242	317	382	436	167	266	439
112	179	234	421	136	243	318	403	437	168	302	440
113	180	235	422	137	244	319	404	438	169	303	441
114	181	252	423	138	245	320	405	439	170	319	
115	182	313	426	139	246	321	406	440	171	321	
116	183	314		140	247	322	407	441	177	323	
117	184	315		141	248	323	408		178	324	
136	185	316		142	249	324	409		180	330	
137	207	317		158	250	325	410		181	331	
138	208	318		159	251	326	411		182	338	
139	209	319		160	252	327	412		183	339	
140	210	320		161	253	328	413		184	340	
110	211	321		169	254	329	414		185	409	
111	212	322		183	255	330	415		186	412	
112	213	323		184	256	331	416		191	413	
113	214	324		186	264	332	417		216	416	
114	215	325		213	301	333	418		217	417	
115	216	326		223	302	334	419		218	419	
116	220	327		224	303	335	420		219	420	

Table S9. Residues that have Cα RMSF values greater than 0.4 Å along the PC1 direction in each MD simulation for NrtA bound to nitrate. Residues that are common to all three MD simulations are highlighted in red.

Residue number									
Run 1				Run 2			Run 3		
139	204	258	375	138	231	420	57	307	426
140	205	261	376	139	232	421	58	310	430
154	206	266	377	140	233	422	111	314	431
155	207	321	378	177	234	423	112	315	432
156	208	322	379	178	368	424	113	316	433
157	209	323	380	179	370	425	114	317	434
160	210	325	381	180	371	426	115	318	435
161	211	326	409	181	372	438	138	319	436
162	212	327	416	182	373	439	139	326	437
164	213	328	417	183	374	440	140	327	438
165	214	329	418	184	375	441	155	329	
166	215	330	422	185	376		165	330	
167	216	331	423	186	377		166	331	
168	217	332	425	190	378		167	332	
169	218	334	426	193	379		168	338	
170	219	335	427	210	380		169	339	
171	220	336	428	211	381		170	340	
172	221	337	430	212	382		171	341	
173	224	338	431	213	383		172	342	
180	227	339	433	214	384		173	343	
181	228	340	434	215	385		174	344	
182	230	341	435	216	386		175	378	
183	231	342	436	217	389		176	379	
184	232	343	437	218	409		177	380	
185	233	344	438	219	410		178	381	
186	234	346	439	220	411		180	382	
191	235	370	440	221	412		183	384	
193	253	371	441	224	414		184	385	
198	254	372		227	416		280	423	
202	255	373		228	417		303	424	
203	257	374		230	418		306	425	

Table S10. Residues that have Cα RMSF values greater than 0.4 Å along the PC1 direction in each MD simulation for NrtA bound to nitrite. Residues that are common to all three MD simulations are highlighted in red.

Residue number					
Run 1		Run 2		Run 3	
124	379	57	340	139	346
130	380	139	341	140	347
133	409	140	342	216	375
134	416	162	343	314	376
135	417	164	344	315	378
136	418	165	345	316	408
137	419	166	346	317	411
138	420	167	347	318	412
139	421	183	348	319	413
140	422	301	421	320	414
141	423	302	422	321	417
153	424	303	440	322	418
193	425	304		323	419
194	426	305		324	420
260	427	306		325	421
261	428	320		326	423
262	430	321		327	424
263	431	323		328	425
264	433	324		329	426
265	434	325		330	428
266	435	326		331	429
323	436	327		332	430
324	437	328		333	431
342	438	329		338	
343	439	330		339	
344	440	331		340	
345	441	332		341	
346		333		342	
354		337		343	
375		338		344	
378		339		345	

Table S11. Residues that are common to all three MD simulations for apo, nitrate-bound, and nitrite-bound NrtA from Tables S8–S10.

Apo	Nitrate-bound	Nitrite-bound
V136	G139	G139
T137	K140	K140
D138	A180	E323
G139	T183	D324
K140	D184	F342
T183	P378	E343
D184	A379	D344
Q224	S380	Q345
A227	T381	E346
N228	D423	F421
V229	I425	
K230	T426	
V231	S438	
N232		
A233		
M234		
V252		
K319		
P321		
E323		
D324		
K330		
G331		
T420		
F421		
F422		
D423		
T426		

References

1. H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley and R. Lopez, *Nucleic Acids Res.*, 2013, **41**, 597–600.
2. M. R. Wilkins, E. Gasteiger, A. Bairoch, J.-C. Sanchez, K. L. Williams, R. D. Appel and D. F. Hochstrasser, in *Methods Mol. Biol.*, Humana Press, NJ, 1999, vol. **112**, pp. 531–552.
3. J. N. Tutol, W. Peng and S. C. Dodani, *Biochemistry*, 2019, **58**, 31–35.
4. F. H. Niesen, H. Berglund and M. Vedadi, *Nat. Protoc.*, 2007, **2**, 2212–2221.
5. E. Freire, O. L. Mayorga and M. Straume, *Anal. Chem.*, 1990, **62**, 950A–959A.
6. R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKerell, *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.
7. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
8. H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westbrook and C. Zardecki, *Acta Crystallogr. D. Biol. Crystallogr.*, 2002, **58**, 899–907.
9. G. Papoyan, K. Gu, J. Wiorkiewicz-Kuczera, K. Kuczera and K. Bowman-James, *J. Am. Chem. Soc.*, 1996, **118**, 1354–1364.
10. C. G. Duan, W. N. Mei, R. W. Smith, J. Liu, M. M. Ossowski and J. R. Hardy, *Phys. Rev. B - Condens. Matter Mater. Phys.*, 2001, **63**, 1441051–1441056.
11. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.
12. N. R. Latorraca, M. Masureel, S. A. Hollingsworth, F. M. Heydenreich, C. M. Suomivuori, C. Brinton, R. J. L. Townshend, M. Bouvier, B. K. Kobilka and R. O. Dror, *Cell*, 2020, **183**, 1813–1825.e18.
13. J. C. Phillips, D. J. Hardy, J. D. C. Maia, J. E. Stone, J. V. Ribeiro, R. C. Bernardi, R. Buch, G. Fiorin, J. Hénin, W. Jiang, R. McGreevy, M. C. R. Melo, B. K. Radak, R. D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V. Kalé, K. Schulten, C. Chipot and E. Tajkhorshid, *J. Chem. Phys.*, 2020, **153**, 044130.
14. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf and O. Beckstein, *J. Comput. Chem.*, 2011, **32**, 2319–2327.
15. G. Fiorin, M. L. Klein and J. Hénin, *Mol. Phys.*, 2013, **111**, 3345–3362.
16. P. H. Hünenberger, A. E. Mark and W. F. van Gunsteren, *J. Mol. Biol.*, 1995, **252**, 492–503.
17. R. N. Goldberg, N. Kishore and R. M. Lennen, *J. Phys. Chem. Ref. Data*, 2002, **31**, 231–370.