SUPPORTING INFORMATION

EXAFS reveals two Mo environments in the nitrogenase iron-molybdenum cofactor biosynthetic protein NifQ

Simon J. George,[†a] Jose A. Hernandez,[†b] Emilio Jimenez-Vicente,[c] Carlos Echavarri-Erasun,[c] and Luis M. Rubio[*c]

[a] Simon Scientific Inc., Berkeley, California 94701

[b] Department of Biochemistry, Midwestern University, Glendale, Arizona 85308

[c] Centro de Biotecnología y Genómica de Plantas, Universidad Politénica de Madrid (UPM) – Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA). Campus Montegancedo UPM, 28223-Pozuelo de Alarcón. Madrid (Spain). E-mail: lm.rubio@upm.es

[†] These authors contributed equally

[*] Corresponding author

TABLE OF CONTENTS

MATERIALS AND METHODS

All chemicals used were analytical grade and were used as received from the chemical manufacturer, unless otherwise indicated. In particular, $CuCl_2$ and α, α'-bipyridyl were purchased from Sigma-Aldrich and used as received. The high quality ultrapure water used in all experiments was from Milli-Q grade (Millipore).

Purification of NifQ:

The His-tagged NifQ protein was purified from cell-free extracts of *Azotobacter vinelandii* strain UW300 (P_{nifH}-His_9-*nifQ*)[1] by anaerobic Co^{2+}-affinity chromatography. Cell-free extracts of *A. vinelandii* were prepared by osmotic shock as described,[2] followed by centrifugation at 30,000 x *g* for 1 h to remove cell debris. Cell-free extracts were loaded onto a 25-ml Co^{2+}-affinity column equilibrated in buffer A (10 mM sodium phosphate, 1.8 mM potassium phosphate, 140 mM NaCl, 2.7 mM KCl, 10% glycerol, pH 7.3). The column was washed with 200 ml of buffer B (50 mM Tris-HCl, 500 mM NaCl, 25 mM imidazole, pH 7.9), and the NifQ protein was eluted from the column by applying 40 ml of buffer C (50 mM Tris-HCl, 150 mM NaCl, 300 mM imidazole, pH 7.9). Eluted NifQ was concentrated by ultrafiltration through a YM10 membrane in an Amicon cell under a N_2 atmosphere, and then subjected to a cycle of dilution-concentration in buffer D (50 mM Tris-HCl, 150 mM NaCl, pH 7.9) to remove residual imidazole. NifQ preparations were free of any protein contaminant based on SDS-PAGE analysis. The resulting NifQ protein was concentrated by ultrafiltration using an Amicon cell equipped with YM10 membrane and an Ultrafree-0.5 device equipped with a PBGC Biomax 10 kDa membrane (Millipore) to 8.4 mM NifQ for Fe-EXAFS or 5 mM NifQ for Mo-EXAFS analyses.

NifQ metal-chelating treatments:

A preparation of purified NifQ was divided into three samples for EXAFS analysis: *(i)* NifQ treated with assay buffer; *(ii)* NifQ treated with $CuCl_2$ (2-fold molar excess over the concentration of NifQ); *(iii)* NifQ treated with α, α'-bipyridyl (10-fold molar excess over the concentration of NifQ). All samples were incubated with the corresponding reagent for 15 minutes. After treatment, the reagent was removed by five steps of concentration in a Centricon Centrifugal Filter Unit device equipped with an Ultracel YM10 membrane (Millipore) followed by dilution in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10% glycerol. All manipulations of the protein and sample preparation were carried out under strict anaerobic conditions. Final protein concentration was 1 mM in all cases.

***In vitro* FeMo-co synthesis assays with purified components:**

The *in vitro* FeMo-co synthesis reactions were performed in 9-ml serum vials sealed with rubber stoppers. All vials were acid-washed overnight in 4 M HCl and thoroughly rinsed with Milli-Q water (Millipore) before use. The vials were repeatedly evacuated and flushed with argon gas and finally rinsed with 0.3 ml of anaerobic 50 mM Tris-HCl (pH 7.5), 1 mM DTH. The complete reactions contained 50 mM Tris-HCl (pH 7.5), 0.182 mM homocitrate, 2.92 mM DTH, 3.3% glycerol, an ATP-regenerating mixture (1.32 mM ATP, 18.7 mM phosphocreatine, 2.27 mM MgCl2, 45.4 µg/ml creatine phosphokinase), 150 µg of NifH (2.4 nmol), 100 µg of apo-NifDK (0.43 nmol), 100 µg of NifEN (0.48 nmol) as a source of FeMo-co precursor (purified NifEN protein carries the VK-cluster),[3] and 0.1 mM Na_2MoO_4. When indicated, 330 µg of NifQ (14 nmol) previously treated with (*i*) assay buffer, (*ii*) $CuCl_2$, or (*iii*) α, α'-bipyridyl were used as

Mo source in the reaction. The reactions (total volume of 350µl) were incubated at 30°C for 35 min to allow for the FeMo-co synthesis and insertion reactions. The resulting activation of apo-NifDK was analyzed by the acetylene-reduction assay after addition of 0.2 mg of NifH and 0.8 ml of ATP-regenerating mixture.

XAS sample preparation:

Purified NifQ preparations were transferred to custom Lucite cuvettes of 4 × 2 × 10 mm (width × height × depth) for Mo measurements, and 20 × 3 × 1 mm (width × height × depth) for Fe measurements, respectively. When used for Mo measurements, NifQ preparations contained 10% glycerol. For Fe measurements, glycerol concentration was increased to 20% to minimize glassing effects. All protein manipulations were performed under strict anaerobic conditions using crimp-sealed vials and gas-tight syringes or inside an anaerobic (<5 ppm O_2) glove box (Coy) with a 95% N_2/5% H_2 environment.

XAS data collection:

Mo and Fe K-edge X-ray absorption data were measured at Stanford Synchrotron Radiation Laboratory beamline 9-3, using a Si 220 double-crystal monochromator and two rhodium-coated mirrors: one flat premonochromator mirror for harmonic rejection and vertical collimation and one toroidal postmonochromator mirror for focusing. Fluorescent X-rays were measured using a 30-element Ge fluorescence detector (Canberra Industries), fitted with Soller slits and a filter to minimize the relative contribution of scattered radiation. Mo measurements used a Zr filter while Fe measurements employed a Mn filter. An Oxford Instruments CF1208 liquid helium cooled sample cryostat was used to maintain the sample temperature at 9K. The X-ray energy was calibrated using the first major inflection point of a standard Mo or Fe foil set as 20003.9 eV or 7112.0 eV respectively. This was measured at the same time as the sample spectrum using two ion chambers positioned downstream of the cryostat. In order to minimize the effect of X-ray photoreduction, the sample was moved after each scan so that the beam irradiated a different spot on the sample for each scan – typically a total of 3 – 4 spots were available for each sample and 5 – 11 scans were collected requiring reuse of each spot. The near-edge structure and position were monitored to ensure no photochemistry was occurring. 10 mM sodium molybdate or 25 mM $K_3Fe(CN)_6$ solutions were used as standard samples.

EXAFS data analysis:

EXAFS data were analyzed using the EXAFSPAK software suite.[4] Although the samples contained the glassing agent glycerol, scans from the individual detector elements were nevertheless rigorously screened in order to eliminate any diffraction artifacts. Curve fitting used the EXAFSPAK program OPT, with single-scattering phase and amplitude functions calculated using FEFF 7.0.[5] The fit quality F was defined as $\sqrt{[\Sigma(\chi_o - \chi_c)^2 k^6/\Sigma\chi_o^2 k^6]}$ where χ_o = the observed EXAFS and χ_c = the calculated EXAFS. The NafY:FeMo-co complex, prepared as previously described,[6] was used as a reference sample to provide initial, experimentally determined estimates of Debye-Waller factors. No smoothing, Fourier filtering, or related manipulations were performed upon the data as part of the fitting process.

THE EFFECT OF INCLUDING AN FE-MO INTERACTION IN THE FE K-EDGE EXAFS FIT

A key question is whether including a Fe-Mo interaction in the Fe K-edge EXAFS fit in Fig. 2a would provide any insight into, for example $MoFe_3S_4$ cluster occupancy via Fe-Mo number (N) or the details of Fe-Fe interactions.

The effect of including a Fe-Mo interaction is illustrated in Table S1. The Fe-Mo distance (R) and Debye-Waller factor (σ^2) are constrained by the values obtained by Mo K-edge EXAFS (Table 1). The Fe-Mo number N is expected to be substantially lower than the 1.0 expected for the proposed $MoFe_3S_4$ cluster since this cluster will likely have relatively low Mo occupancy with most of the Fe in the protein comprising Mo-free Fe_3S_4 clusters. We can estimate reasonable values for N from the overall Mo occupancy, around 0.3 Mo per NifQ compared to 3.1 Fe per NifQ[6], combined with the two approximately equal Mo environments proposed in this paper. It follows reasonable values for the number of Fe-Mo interactions are between 0.1 and 0.2 with the limiting extremes being 0.0 and 0.3.

Hence, in the example fit presented in Table S1, we include as non-floating parameters 0.2 Fe-Mo interactions with the Fe-Mo distance (2.713 Å) and Debye-Waller factor (0.0034 $Å^2$) taken from the values from Mo-Fe EXAFS curve fitting (Table 1). Subsequent fixing of the Fe-Fe parameters and floating these Fe-Mo parameters causes only a small further change in the fit, indicating the fit quality is at a minimum.

This fit results in a significant, 8%, Fe-Mo contribution to the EXAFS, which causes a reduction in the Fe-Fe Debye-Waller interaction compared to the fit with no included Fe-Mo component, but importantly, does **not** change the Fe-Fe distance, the Fe-S distance or Fe-S Debye-Waller factor. In addition, it is clear from Table S1 that including the Fe-Mo component only produces a slight improvement in fit quality, which, since we have included extra parameters, is probably not statistically significant.

As stated in the main text, this lack of effects on the other parameters or fit quality is a consequence of the Fe-Mo interaction being essentially out of phase with the Fe-Fe interaction. Essentially, the presence of Fe-Mo intensity can be almost completely compensated by increasing the Fe-Fe intensity. This effect reaches a limit when the Fe-Fe Debye-Waller factor drops below the reasonable minimum limit, around 0.0032 Å for an Fe-S cluster, which correlates with Fe-Mo $N < 0.55$. It follows that accurate quantitation of the Fe-Mo number (N) is not possible from the Fe EXAFS spectrum, which can be reasonably fitted to any value from 0 to 0.55.

Hence our decision to not include this Fe-Mo component in our *a priori* analysis of this spectrum in Fig 2a and Table 1. While the Mo EXAFS and other evidence demonstrates that it should be present, the negative correlation with Fe-Fe means that it is not obviously apparent in the Fe EXAFS. Moreover, including it does not generate any new useful information.

DESCRIPTION OF ALTERNATIVE Mo K-EDGE EXAFS FITS

As stated in the main text, inspection of the Mo K-edge EXAFS (Fig. 2a and b) showed at least 4 interactions that were apparent in the Fourier transform (Fig. 2b), making NifQ a complex Mo system. It was apparent that a MoFeS cluster was present with Mo-S interactions at 2.34 Å and Mo-Fe interactions at 2.71 Å. Interestingly, a short, 1.73 Å, Mo-O interaction was also present. In analyzing these data, several models were trialed. These trials are summarized below and detailed in Fig. S1 and Table S2.

1) NifQ contains a single $MoFe_3S_4$ cluster with a novel short Mo=O bond. This model fails because the Mo-Fe and Mo-S Debye-Waller factors are too high at > 0.007 $Å^2$. These are about double the acceptable values (Fig. S1a).

2) NifQ contains a $MoFe_3S_4$ cluster and additional bound (or free) MoO_4^{2-}. This model arose from the hypothesis that the pre-edge feature at 20.008 keV was entirely due to MoO_4^{2-}. Integration of this feature would then indicate that NifQ contained 27% MoO_4^{2-} with the balance of 73% Mo presumably being contained in the $MoFe_3S_4$ cluster (fitting the NifQ Mo K-edge spectrum with a mixture of the NafY:FeMo-co and MoO_4^{2-} spectra in Fig. 1 gives a similar quantitation).

This model fitted the data much better than model 1 (Fig. S1b). However, it fails because (a) the Mo-S and Mo-Fe Debye-Waller factors are too high and (b), as stated in the main text, the observed Mo=O bond length of 1.73 Å is significantly shorter than both the 1.79 Å of MoO_4^{2-} free in solution and the 1.75-1.77 Å distance found in MoO_4^{2-} bound to molbindin proteins.

3) NifQ contains two Mo sites (each containing about half the Mo): a $MoFe_3S_4$ cluster and a Mo=O containing site. This model gave a reasonable fit, provided additional Mo-O interactions were included at ~ 2.1 Å (Fig. S1c). The quality of the fit was greater than in models 1 or 2. However, the Mo-S Debye-Waller factor is still high, and it is clear from the Fourier transform that this fit did not reproduce the split "Mo-S" peak at ~ 2.3 Å.

4) NifQ contains two Mo sites (each containing about half the Mo): a $MoFe_3S_4$ cluster and a Mo=O containing site that includes a short Mo-S bond. This gave the best fit (Fig. S1d) and is presented in the main text (Fig. 2). This model is similar to model 3 but it includes an additional short Mo-S distance at 2.23 Å. This achieves a very reasonable Debye-Waller factor for the 2.34 Å Mo-S interaction of ~ 0.0040 $Å^2$. Including longer Mo-O bonds at around 2.12 Å also consistently improved the fit, although in this case these contribute only about 3% intensity to the overall EXAFS. For this fit, the split "Mo-S" peak at ~ 2.3 Å in the Fourier transform was simulated nicely.

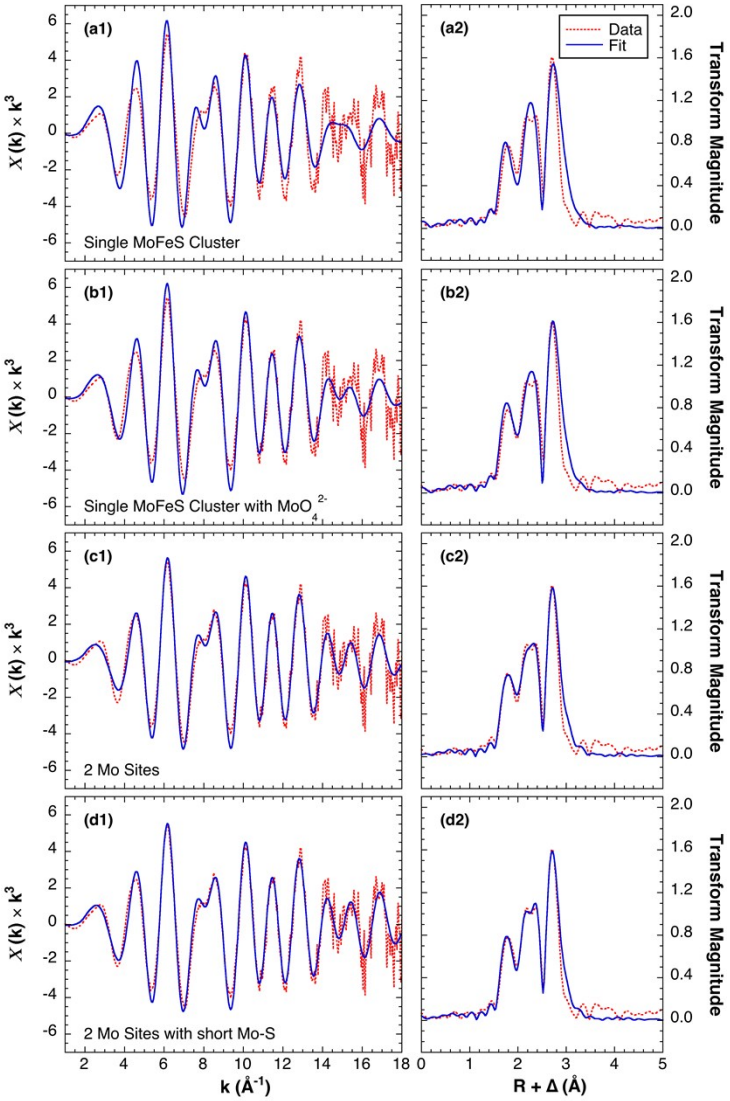

Fig. S1 Mo K-edge EXAFS recorded from *A. vinelandii* NifQ showing fits from the various models. In each case, the k^3 weighted EXAFS spectrum of NifQ is the broken line, while the fit is the solid line. The left panels (a1, b1, c1, d1) are the EXAFS data and fits while the right panels (a2, b2, c2, d2) are the corresponding Mo-S phase corrected Fourier transforms. The fitting models are:

(a1, a2) One single Mo site, as a novel O=$MoFe_3S_4$ cluster.

(b1, b2) One Mo site as $MoFe_3S_4$ cluster with additional bound or free MoO_4^{2-}. The MoO_4^{2-} was quantitated to 27% total Mo by comparing the pre-edge Mo-O feature observed in the NifQ near-edge spectrum (Fig. 1a) and the corresponding spectrum from MoO_4^{2-} solution (Fig. 1e).

(c1, c2) Two Mo sites, in approximately 50:50 distribution, one a $MoFe_3S_4$ cluster, the other an isolated bound Mo=O.

(d1, d2) Two Mo sites, in approximately 50:50 distribution as in (c1, c2), but including a short Mo-S bond. This is the fit presented in Fig. 2b.

Fitting parameters are presented in Table S2.

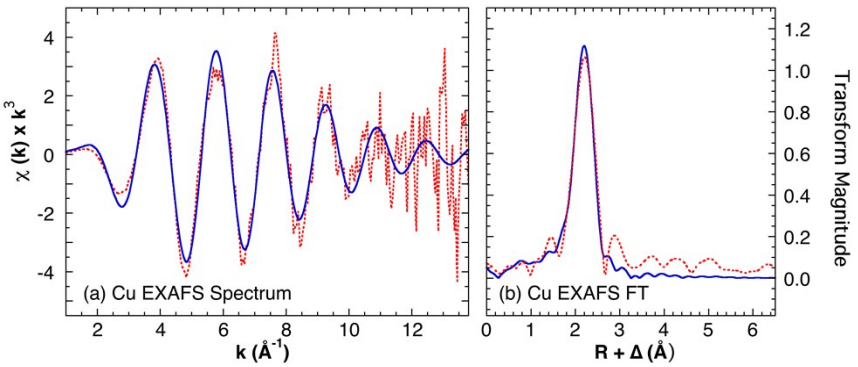

Fig. S2 Cu K-edge EXAFS spectrum of NifQ after Cu treatment. Data and fits are presented as broken and solid lines, respectively. Left panel, k^3 weighted Cu EXAFS spectrum. Right panel, Cu-S phase corrected Fourier transform from left panel.

Fitting parameters are presented in Table S4.

While this EXAFS spectrum shows that Cu has been retained, it also shows that Cu-Mo complexes with a characteristic 2.7 Å distance are absent.

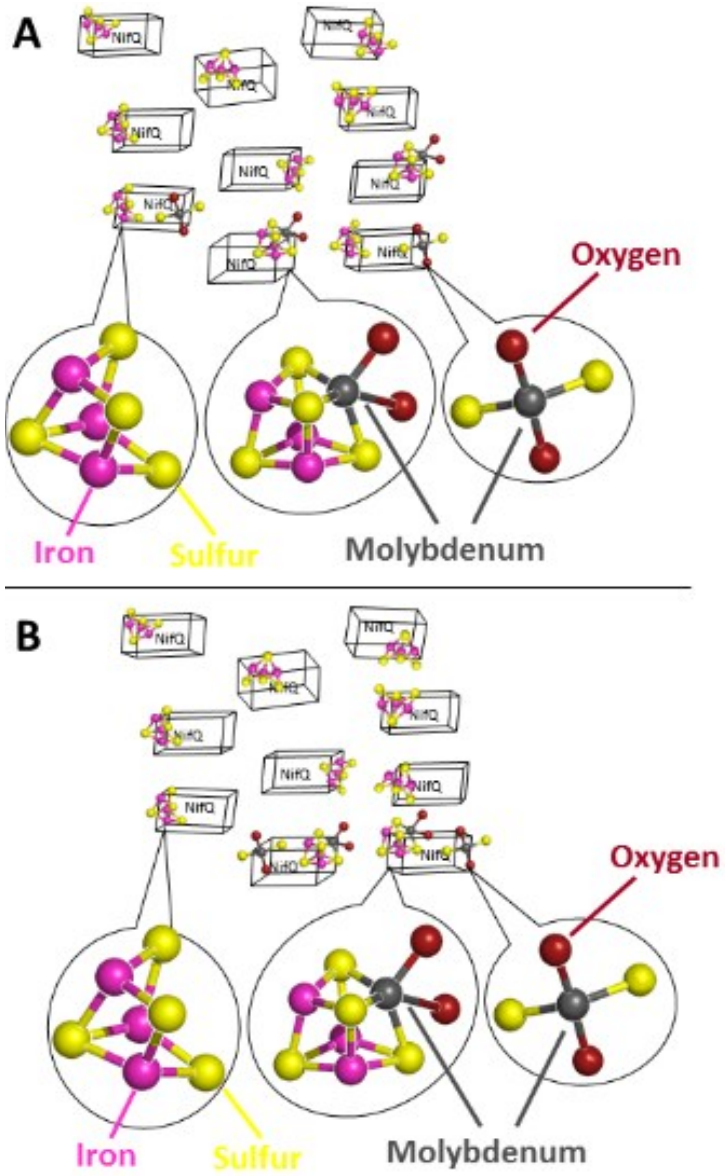

Fig. S3 Schematic representation of Fe-S clusters and Mo environments in NifQ. The purified preparations of NifQ would contain two Mo sites, which could either be found in different NifQ molecules (panel A) or coexist in the same NifQ molecule (panel B).

Table S1. Alternative fits for the NifQ Fe K-edge EXAFS spectrum (figure 2a) with and without Fe-Mo interactions.

	Without Fe-Mo			With Fe-Mo		
	N	**R (Å)**	**σ^2 (Å^2)**	**N**	**R (Å)**	**σ^2 (Å^2)**
Fe-S	4.0	2.269 (1)	0.0037 (1)	4.0	2.269 (1)	0.0037 (1)
Fe-Fe	2.0	2.707 (2)	0.0048 (2)	2.0	2.708 (2)	0.0040 (2)
Fe-Mo				0.2	2.713	0.0034
ΔE_0 (eV)		-13.3 (4)			-13.2 (3)	
Scale Factor		0.85			0.85	
F		199			194	

N = number of backscattering atoms. R = distance. σ^2 = Debye-Waller factor. ΔE_0 = threshold energy. F = fit quality defined as $\sqrt{[\Sigma(\chi_o - \chi_c)^2 k^6/\Sigma\chi_o^2 k^6]}$ (χ_o = observed EXAFS; χ_c = calculated EXAFS). In general, the backscatter atom numbers (N) were fixed to the values expected for the models while R, σ^2 and ΔE_0 were floated. Threshold energies (ΔE_0) were constrained to be the same for all components. Phase and amplitude functions were calculated using FEFF 7.0.[5]

The floated parameters are indicated by uncertainties in parentheses to the indicated precision. The Fe-Mo distance (R) and Debye-Waller (σ^2) interactions were constrained to those fitted to the Mo EXAFS (Tables 1 and S2).

Table S2. Fitting parameters for Mo K-edge EXAFS models as illustrated in Fig. S1

	Single MoFeS Cluster (a1, a2)			Single Cluster + MoO_4^{2-} (b1, b2)			2 Site Model (c1, c2)			2 Sites & short Mo-S (d1, d2)		
	N	**R (Å)**	**σ^2 (Å^2)**	**N**	**R (Å)**	**σ^2 (Å^2)**	**N**	**R (Å)**	**σ^2 (Å^2)**	**N**	**R (Å)**	**σ^2 (Å^2)**
Mo-S (long)	3.0	2.290	0.00795	2.18	2.302	0.00641	1.5	2.315	0.00505	1.5	2.336	0.00400
Mo-Fe	3.0	2.713	0.00724	2.18	2.714	0.00517	1.5	2.717	0.00345	1.5	2.713	0.00343
Mo-O (short)	1.0	1.733	0.00227	1.10	1.734	0.00252	0.81	1.737	0.00153	0.88	1.734	0.00186
Mo-O (long)				0.73	2.069	0.00487	1.27	2.084	0.00454	1.05	2.124	0.00500
Mo-S (short)										0.88	2.227	0.00377
ΔE_0 (eV)	-10.55			-9.94			-8.23			-9.70		
Scale Factor	1.05			1.05			1.05			1.05		
F	317			208			131			118		

N = number of backscattering atoms. R = distance. σ^2 = Debye-Waller factor. ΔE_0 = threshold energy. F = fit quality defined as $\sqrt{[\Sigma(\chi_o - \chi_c)^2 k^6/\Sigma\chi_o^2 k^6]}$ (χ_o = observed EXAFS; χ_c = calculated EXAFS). In general, the backscatter atom numbers (N) were fixed to the values expected for the models while R, σ^2 and ΔE_0 were floated. Threshold energies (ΔE_0) were constrained to be the same for all components. The estimated uncertainties in N, R, σ^2 and ΔE_0 were in general less than ± 0.25, ± 0.006 Å, ± 0.0008 Å^2, and ± 0.6 eV, respectively. Phase and amplitude functions were calculated using FEFF 7.0.[5]

Table S3. EXAFS curve fitting parameters for the spectra of NifQ samples treated with reagents to selectively eliminate each Mo environment as illustrated in Fig. 4.

	as-isolated			after buffer-exchange treatment			after α, α'-bipyridyl treatment			after Cu(II) treatment		
	N	**R (Å)**	**σ^2 (Å^2)**	**N**	**R (Å)**	**σ^2 (Å^2)**	**N**	**R (Å)**	**σ^2 (Å^2)**	**N**	**R (Å)**	**σ^2 (Å^2)**
Mo-S	1.5	2.336	0.00400	1.39	2.325	0.00454	1.53	2.250	0.00454	2.89	2.263	0.00454
Mo-Fe	1.5	2.713	0.00343	1.44	2.713	0.00345	0.95	2.701	0.00345	2.07	2.690	0.00345
Mo-O (short)	0.88	1.734	0.00186	0.76	1.734	0.00200	0.57	1.736	0.00200	0.60	1.730	0.00200
Mo-O (long)	1.05	2.124	0.00500	1.52	2.095	0.00514	1.24	2.088	0.00514	2.43	2.186	0.00514
ΔE_0 (eV)	-9.70			-9.29			-9.23			-7.92		
Scale Factor	1.05			1.05			1.05			1.05		
F	124			170			212			412		

N = number of backscattering atoms. R = distance. σ^2 = Debye-Waller factor. ΔE_0 = threshold energy. F = fit quality defined as $\sqrt{[\Sigma(\chi_o - \chi_c)^2 k^6/\Sigma\chi_o^2 k^6]}$ (χ_o = observed EXAFS; χ_c = calculated EXAFS). Threshold energies (ΔE_0) were constrained to be the same for all components. The estimated uncertainties in N, R, σ^2 and ΔE_0 were in general less than ± 0.25, ± 0.006 Å, ± 0.0008 Å^2, and ± 0.6 eV, respectively. Phase and amplitude functions were calculated using FEFF 7.0.[5]

The Debye-Waller factors were constrained to the same values for the buffer-exchange, α, α'-bipyridyl and Cu(II) treated samples.

Table S4. Fitting parameters for Cu EXAFS model as illustrated in Fig. S2

Interaction	N	R (Å)	σ^2 (Å^2)	ΔE_0 (keV)	F
Cu-S	2.0	2.277 (0.007)	0.0072 (0.0006)	-6.5 (1.2)	173
Cu-O	2.0	2.064 (0.015)	0.0100 (0.0019)		

EXAFS curve fitting parameters for the spectra of Cu-treated NifQ. The corresponding spectrum is illustrated in Fig. S2. N = number of backscattering atoms. R = distance. σ^2 = Debye-Waller factor. ΔE_0 = threshold energy, F = fit quality defined as $\sqrt{[\Sigma(\chi_o - \chi_c)^2 k^6/\Sigma\chi_o^2 k^6]}$ (χ_o = observed EXAFS; χ_c = calculated EXAFS). Threshold energies (ΔE_0) were constrained to be the same for all components. The values of N were fixed. The estimated uncertainties in R, σ^2 and ΔE_0 are given in parentheses. Phase and amplitude functions were calculated using FEFF 7.0.[5]

REFERENCES

1 J. A. Hernandez, L. Curatti, C. P. Aznar, Z. Perova, R. D. Britt and L. M. Rubio, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 11679-11684.

2 V. K. Shah, L. C. Davis and W. J. Brill, *Biochim. Biophys. Acta*, 1972, **256**, 498-511.

3 J. A. Hernandez, R. Y. Igarashi, B. Soboh, L. Curatti, D. R. Dean, P. W. Ludden and L. M. Rubio, *Mol. Microbiol.*, 2007, **63**, 177-192.

4 G. N. George, S. J. George and I. J. Pickering, *Stanford Synchrotron Radiation Laboratory*, 1998.

5 J. J. Rehr and R. C. Albers, *Rev. Mod. Phys.*, 2000, **72**, 621-654.

6 L. M. Rubio, S. W. Singer and P. W. Ludden, *J. Biol. Chem.*, 2004, **279**, 19739-19746.