Supplementary Material for:
Laue Crystal Structure of *Shewanella oneidensis* Cytochrome *c* Nitrite Reductase from a High-yield Expression System

Matthew Youngblut, Evan T. Judd, Vukica Srajer, Bilal Sayyed, Tyler Goelzer, Sean Elliott, Marius Schmidt, A. Andrew Pacheco

S1. Comparison between the wild-type ccNiR gene and the optimized gene used for over-expression.

Optimized gene
Wild type gene
Restriction sites (opt. gene)
Ribosome binding site and spacer (opt. gene)
N-terminal signal sequence for periplasmic translocation

GAATTCGGAGGATACAATT
ATGTCGAAGAAACTACTAAGTGTCCTATTTGGTGCGTCGCTGGCTGCTCTGGCCCTAAGTCCTACCGCCTTT
ATGATGAAGAAGATGACAGGTAAGACTTTTGCATTAAGTGCATTGGTTGCCGCCAGCTT TATGGCTGCGGGGGCTATG
GCGTCGGATAAGACCGAACCACGTAATGAGGTGTACAAGGATAAGTTTAAGAACCAATACAACAGTTGGCATGATACC
GCGAGTGATAAAACTGAGCCTCGCAACGAAGTTTATAAAGATAAATTTAAAAATCAATACAATAGCTGGCATGACACC
GCTAAATCGGAAGAGCTAGTCGATGCGCTGGAACAGGATCCCAACATGGTGATTTTGTGGGCCGGCTATGCTTTTGCG
GCAAAGAGTGAAGAGCTTGTTGATGCGCTAGAGCAGGACCCTAATATGGTTATCCTGTGGGCAGGCTATGCATTTGCC
AAAGATTACAAGGCACCACGTGGTCATATGTATGCTGTGACCGATGTTCGTAATACCCTACGCACCGGCGCACCCAAA
AAAGATTATAAAGCCCCTCGTGGCCACATGTATGCCGTGACTGACGTACGTAATACCTTGCGTACTGGCGCGCCAAAA
AACGCTGAAGATGGTCCCCTGCCAATGGCCTGTTGGAGTTGCAAATCGCCTGATGTCCCCCGCTTAATTGAAGAGCAA
AATGCGGAAGATGGTCCATTACCTATGGCTTGCTGGAGCTGTAAAAGTCCTGACGTGCCTCGCTTAATTGAAGAGCAA
GGCGAAGATGGTTATTTTAAGGGCAAATGGCAAAAGGCGGTCCAGAAGTGACCAATACCATCGGCTGTAGCGATTGC
GGTGAAGACGGTTACTTCAAAGGTAAGTGGGCGAAAGGTGGCCCAGAAGTGACTAACACTATTGGCTGTAGCGATTGC
CATGAAAAGGGTTCGCCTAAATTACGTATTAGCCGCCCCTACGTCGATCGCGCACTTGATGCTATCGGTACCCCTTTT
CACGAAAAAGGTTCACCAAAATTACGTATTTCTCGTCCTTATGTTGACCGCGCATTAGATGCAATTGGTACTCCATTT
AGCAAAGCTAGTAAGCAAGATAAAGAAAGTATGGTGTGTGCGCAGTGCCATGTTGAATACTACTTTGAAAAGAAAGAA
AGCAAAGCTTCTAAGCAAGACAAAGAGTCAATGGTATGCGCCCAGTGTCACGTTGAATACTACTTTGAGAAAAAAGAA
GATAAGAAGGGCTTTGTTAAGTTTCCCTGGGATATGGGTGTCACCGTGGATCAAATGGAAGTGTATTACGATGGCATT
GATAAGAAAGGTTTCGTCAAATTCCCTTGGGATATGGGCGTAACTGTAGATCAAATGGAAGTTTACTATGATGGCATT
GAGTTTAGCGATTGGACCCATGCCCTTAGTAAGACCCCAATGCTTAAAGCACAGCATCCTGAATACGAGACCTGGAAG
GAGTTCTCTGATTGGACTCACGCACTGTCTAAAACGCCAATGTTAAAAGCCCAGCATCCCGAGTATGAAACGTGGAAA
ATGGGCATCCATGGTAAAAATAACGTCTCGTGTGTGGATTGCCATATGCCCAAAGTGACCAGCCCAGAAGGCAAAAAG
ATGGGTATCCATGGTAAAAATAATGTCAGTTGCGTTGACTGTCATATGCCTAAAGTGACTAGCCCTGAAGGTAAGAAG
TTTACCGATCATAAGGTTGGTAACCCATTTGATCGTTTTGAAGAGACCTGTGCTACCTGCCATTCGCAAACCAAAGAA
TTTACTGACCATAAAGTGGGTAACCCATTTGATCGTTTCGAAGAAACCTGTGCAACTTGCCACAGCCAAACCAAAGAG
TTTTTAGTTGGCGTCACCAACGAACGCAAGGCAAAAGTTAAGGAGATGAAACTCAAGGCGGAAGAGCAGTTGGTCAAG
TTCTTAGTTGGCGTCACTAATGAGCGCAAAGCTAAAGTGAAAGAAATGAAACTCAAAGCGGAAGAGCAATTAGTGAAA
GCCCATTTTGAAGCCGCAAAAGCATGGGAGCTCGGTGCGACCGAAGCCGAGATGAAACCTATTTTGACCGATATCCGT
GCGCACTTCGAAGCGGCGAAAGCATGGGAATTAGGTGCCACTGAAGCCGAAATGAAGCCAATTCTGACTGATATCCGC
CATGCCCAGTGGCGCTGGGATCTTGCCATCGCAAGTCATGGTGGCTGCACATGCACCTGAAGAGGCCCTCCGTGTT
CATGCTCAATGGCGCTGGGATTTAGCCATTGCCTCACATGGCGTTGCAGCTCATGCTCCAGAAGAAGCACTACGCGTG
TTGGGTACCTCGGTCAATAAAGCCGCAGATGCCCGCGTCAAACTAGCACAATTACTTGCTAAAAAGGGCCTGACCGAT
TTGGGCACCTCTGTTAATAAAGCGGCCGATGCCCGTGTTAAGTTAGCACAGTTGTTAGCGAAAAAAGGCTTAACAGAC
CCAGTCGCAATTCCTGATATCAGTACCAAAGCTAAGGCGCAGGCCGTGCTCGGTATGGATATGGAAAAGATGAACGCG
CCTGTTGCCATCCCTGATATTTCGACTAAAGCTAAGGCTCAAGCAGTACTCGGAATGGACATGGAAAAGATGAATGCA
GAAAAAGAGGCCTTTAAGAAAGATATGCTGCCAAAGTGGGATGCGGAGGCGAAGAAGCGTGAGGCGACCTACAAGTGA
GAGAAAGAAGCATTCAAGAAAGACATGTTACCTAAGTGGGATGCTGAAGCGAAAAAACGCGAAGCGACTTACAAGTAA
TCTAGA

S2. Spectropotentiometric analysis of ccNiR

S2.1. Derivation of Equation 5. Using the terminology presented in Scheme 2, we can write the Nernst equations that relate two given ccNiR species C_n and C_{n-1}, as shown in Eqs. S1-S5.

$$\varepsilon_{app} = \varepsilon_1^o - \frac{RT}{nF} \ln \frac{C_1}{Ox} \tag{S1}$$

$$\varepsilon_{app} = \varepsilon_2^o - \frac{RT}{nF} \ln \frac{C_2}{C_1} \tag{S2}$$

$$\varepsilon_{app} = \varepsilon_3^o - \frac{RT}{nF} \ln \frac{C_3}{C_2} \tag{S3}$$

$$\varepsilon_{app} = \varepsilon_4^o - \frac{RT}{nF} \ln \frac{C_4}{C_3} \tag{S4}$$

$$\varepsilon_{app} = \varepsilon_5^o - \frac{RT}{nF} \ln \frac{C_5}{C_4} \tag{S5}$$

In exponential form these equations can be re-written in the forms shown in Eqs. S6 and S7:

$$C_1 = Ox \times E_1 \tag{S6}$$

$$C_n = C_{n-1} \times E_n \tag{S7}$$

where E_n is defined by Eq. 5c from the main text:

$$E_n = \exp\left[\frac{nF}{RT} (\varepsilon_n^o - \varepsilon_{app}) \right] \tag{5c}$$

Now the total concentration of ccNiR, C_T, will be the given by the sum of the various species present at any given applied potential (Eq. S8). All of the species except Ox can be eliminated from Eq. S8 by substituting each C_n value with the expressions S6, S7. The result is shown in

S3

$$C_T = Ox + \sum_{n=1}^{5} C_n \tag{S8}$$

$$C_T = Ox\left\{1 + E_1\left[1 + E_2\left[1 + E_3\left(1 + E_4(1 + E_5)\right)\right]\right]\right\} \tag{S9}$$

Eq. S9. Solving for Ox gives Eq. S10, where the large term in the denominator has been abbreviated as "$denom$" (Eq. 5b from main text). Finally, by doing successive substitutions into Eqs. S6 and S7, all of the C_n species can be rewritten in terms of C_T, as shown in Eq. 5a of the main text.

$$Ox = \frac{C_T}{denom} \tag{S10}$$

$$denom = 1 + E_1\left\{1 + E_2\left[1 + E_3\left(1 + E_4(1 + E_5)\right)\right]\right\} \tag{5b}$$

$$C_n = \frac{C_T \times \prod_1^n E_n}{denom} \tag{5a}$$

S2.2. Beer's law for difference spectra. We begin by eliminating Ox from the Beer's law expression (Eq. S11) by first solving for Ox in Eq. S8 (which yields Eq. S12), and then

$$A_\lambda = \left(\epsilon_{ox\lambda}Ox + \sum_{n=1}^{5}\epsilon_{n\lambda}C_n\right) \times l \tag{S11}$$

substituting Ox using the Eq. S12 expression to give Eq. S13. In these equations A_λ is the absorbance at a given wavelength λ, $\epsilon_{ox\lambda}$ and $\epsilon_{n\lambda}$ are the extinction coefficients of the oxidized

$$Ox = C_T - \sum_{n=1}^{5} C_n \tag{S12}$$

$$A_\lambda = \left\{\epsilon_{ox\lambda}\left[C_T - \left(\sum_{n=1}^{5}C_n\right)\right] + \sum_{n=1}^{5}\epsilon_{n\lambda}C_T\right\} \times l \tag{S13}$$

and n^{th} reduced species at that wavelength, and l is the cell pathlength. In the absence of applied potential, all of the ccNiR is in the oxidized form, and Eq. S13 simplifies to Eq. S14, where $A_{0\lambda}$ is the absorbance of fully oxidized ccNiR at λ. If we use Eq. S14 to substitute for $\epsilon_{ox\lambda}C_T$ in Eq. S13 we get, after rearrangement, Eq. S15. Straightforward rearrangement of the summations then yields the desired Eq. S16, in which $\Delta A_\lambda = A_\lambda - A_{0\lambda}$ and $\Delta\epsilon_\lambda = \epsilon_{n\lambda} - \epsilon_{ox\lambda}$. Equation 6 is a matrix

$$A_{0\lambda} = \epsilon_{ox\lambda} C_T l \tag{S14}$$

$$A_{\lambda} = \left(A_{0\lambda} - \epsilon_{ox\lambda} \sum_{n=1}^{5} C_n + \sum_{n=1}^{5} \epsilon_{n\lambda} C_T \right) \times l \tag{S15}$$

$$\Delta A_{\lambda} = l \times \sum_{n=1}^{n} \Delta\epsilon_{n\lambda} C_n \tag{S16}$$
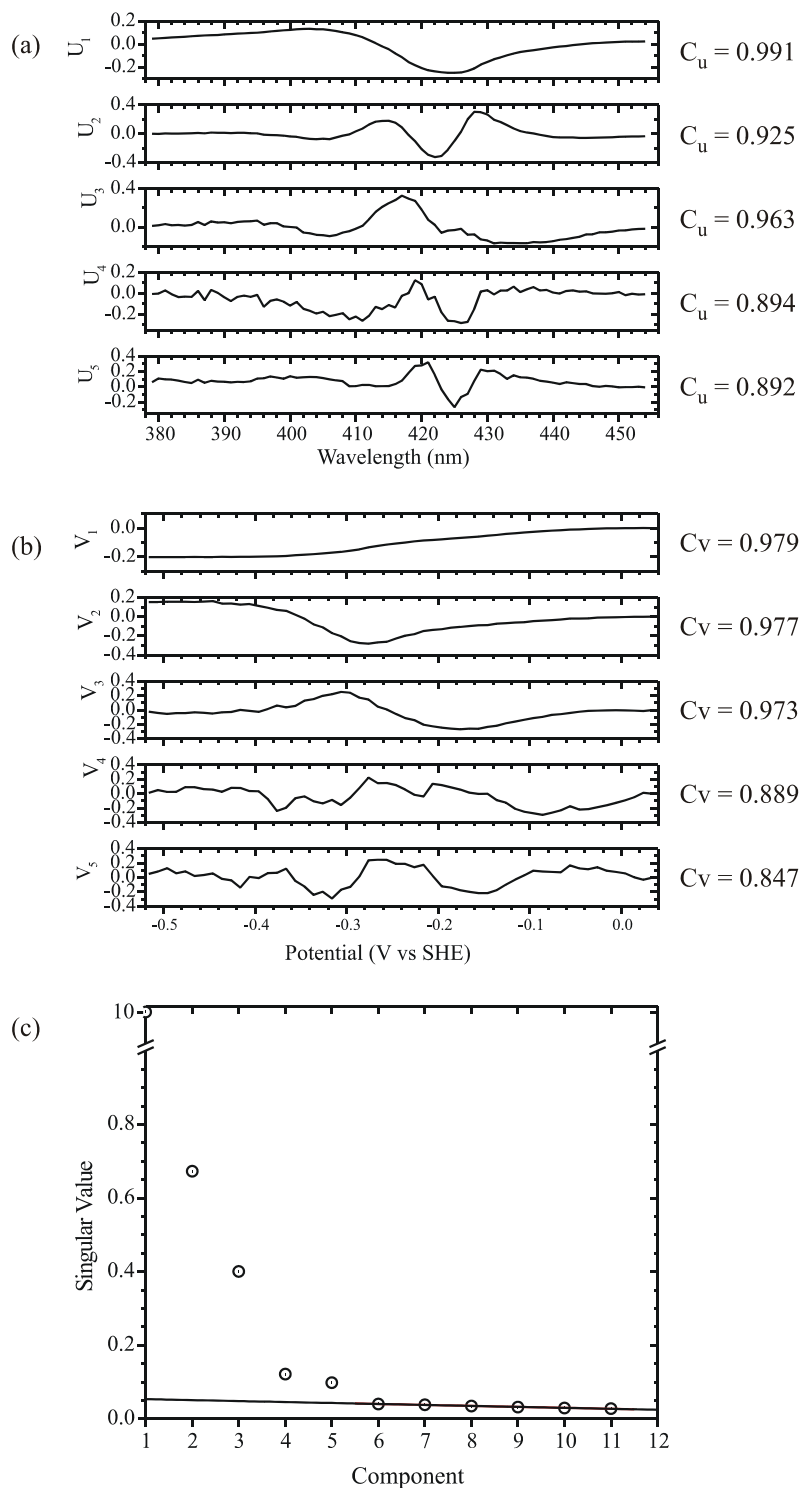
form of Eq. S16, rearranged to give the extinction coefficient spectra as a function of the concentrations, absorbances and cell pathlength. Because the concentration matrix is not square and thus has no inverse, in Eq. 6 one must construct the pseudoinverse, as explained in ref. [52].

S2.3. Complete description of the global analysis procedure. Initially, the absorbance difference data collected between 379 nm and 454 nm, for all potentials between +34 mV and –516 mV vs SHE, were subjected to singular value decomposition (SVD) analysis. This analysis revealed five components with U and V autocorrelation values greater than 0.8, and singular values above the noise background, as shown in Section S2.4. All subsequent U and V components had autocorrelation values well below 0.8, which corresponds roughly to $s/n = 1$ [51].

Selected ΔA_{λ} vs. ε_{app} traces of the SVD-processed data were next fitted to Eqs. 5 and S16 using the Levenberg-Marquardt routine supplied by Origin 6.0 [52]. At some wavelengths it was possible to fit the ΔA_{λ} vs. ε_{app} slices with less than 5 potentials, because some of the $\Delta\epsilon_{n\lambda}$ were small at those wavelengths. For example, 420 nm and 430 nm data could be fitted with 4 and 3 potentials, respectively. Taken together, these initial analyses of individual ΔA_{λ} vs. ε_{app} slices provided rough guesses of the $\varepsilon^o{}_n$ values.

In the final step the entire SVD-processed data set was fitted using Eqs. 5 and the matrix Eq. 6, using a program created with Mathcad 13. Briefly, this global analysis program first allows the user to manually enter trial values of the five midpoint potentials. For a given set of midpoint potential values the program then uses Eq. 5 to calculate the concentration of each reduced species at a given applied potential, and stores the concentrations in $\mathbf{C_{red}}$. $\mathbf{C_{red}}$, $\mathbf{\Delta A}$ and l are then used to calculate the matrix of extinction coefficient spectra $\mathbf{\Delta\epsilon}$ using Eq. 6. Finally, the program uses l and the $\mathbf{\Delta\epsilon}$ and $\mathbf{C_{red}}$ matrices to generate an absorbance matrix $\mathbf{\Delta A_{calc}}$, and the sum of squares for $\mathbf{\Delta A} - \mathbf{\Delta A_{calc}}$ is computed. By varying the trial values of the five ccNiR midpoint potentials, this sum of squares can be minimized. The midpoint potentials obtained by fitting the single wavelength ΔA_{λ} vs. ε_{app} slices to Eqs. 5 and S16 were used as initial trial values in the global fit. These trial values proved to be excellent initial guesses, and only minor modifications to the $\varepsilon^o{}_n$ values were needed in the subsequent sum of squares minimization procedure.

S2.4. Results of singular value decomposition analysis of the Soret region data.

(a)

$C_u = 0.991$

$C_u = 0.925$

$C_u = 0.963$

$C_u = 0.894$

$C_u = 0.892$

(b)

$Cv = 0.979$

$Cv = 0.977$

$Cv = 0.973$

$Cv = 0.889$

$Cv = 0.847$

(c)

Figure S1. (a) The 5 U components corresponding to singular values above the noise level, and their associated autocorrelation values C_U. (b) The 5 V components corresponding to singular values above the noise level, and their associated autocorrelation values C_V. (c) Singular values of the first 11 components.

S2.5. Results of singular value decomposition analysis of the α,β region data. The spectral region from 500 nm – 650 nm, where characteristic bands are observed for reduced low-spin hemes (the α, β bands), was investigated in a manner analogous to that described above and in the main paper for the Soret region. The absorbance difference data collected between 500 nm and 650 nm, for all potentials between +34 mV and -516 mV vs SHE, were first subjected to SVD analysis. For this region, where the signals are less intense than in the Soret region, inspection of the U and V autocorrelation values and singular values unambiguously showed only four spectral components above the noise background (Figs. S2a and S2b).

Initially, the new SVD-processed data set was fitted directly to a 4-component version of the matrix Eq. 6, using a 4-component version of the Mathcad global analysis program described above and in the main paper, and using the optimized midpoint potential values obtained by fitting the 379 nm – 454 nm data as guiding initial guesses. The optimized midpoint potential values thus obtained are listed in Table S1; the values obtained from the Soret region are also given for comparison. Notice that $\mathscr{E}^o_1$, $\mathscr{E}^o_2$ and $\mathscr{E}^o_4$ correlate very well with the Soret values $\mathscr{E}^o_1$, $\mathscr{E}^o_2$ and $\mathscr{E}^o_5$, whereas the $\mathscr{E}^o_3$ value sits between the Soret $\mathscr{E}^o_3$ and $\mathscr{E}^o_4$ values. We interpret this to mean that the spectral changes seen in the α, β region are tracking the same electrochemical processes as those seen in the Soret region, but that the extinction coefficient difference spectra of the various reduced species are too similar to each other to give rise to 5 distinct components in the SVD. To explore this hypothesis we fitted the α, β region data using 5 potentials, but assuming that the extinction coefficient difference spectra $\Delta\epsilon_3$ and $\Delta\epsilon_4$ are identical. The results are listed in Table S1, representative fits are provided in Fig. S3, and the calculated extinction coefficient difference spectra are given in Fig. S4. In support of the model, the fits are slightly better than those obtained for the 4-potential 4-$\Delta\epsilon$ model, and excellent across the spectral region.

Table S1. Midpoint potentials of ccNiR hemes (in V vs. SHE) obtained by spectropotentiometry. Comparison between values obtained by fitting the Soret region and those obtained by fitting the 500 - 650 nm region.

| Soret region analysis	500nm - 650 nm region analysis		
5 spectra, 5 potentials	4 spectra, 4 potentials[1]	4 spectra, 5 potentials	4 spectra, 4 potentials[2]
-0.063±0.008	-0.048	-0.045	-0.063
-0.148±0.006	-0.140	-0.139	☐
-0.232±0.013	-0.269	-0.271	-0.232
-0.282±0.002		-0.301	-0.282
-0.338±0.002	-0.338	-0.336	-0.338
Sum of squares:	1.78E-03	1.73E-03	7.72E-03

[1.] No restrictions on midpoint potential values

[2.] Second reduction event is assumed to produce no absorbance change
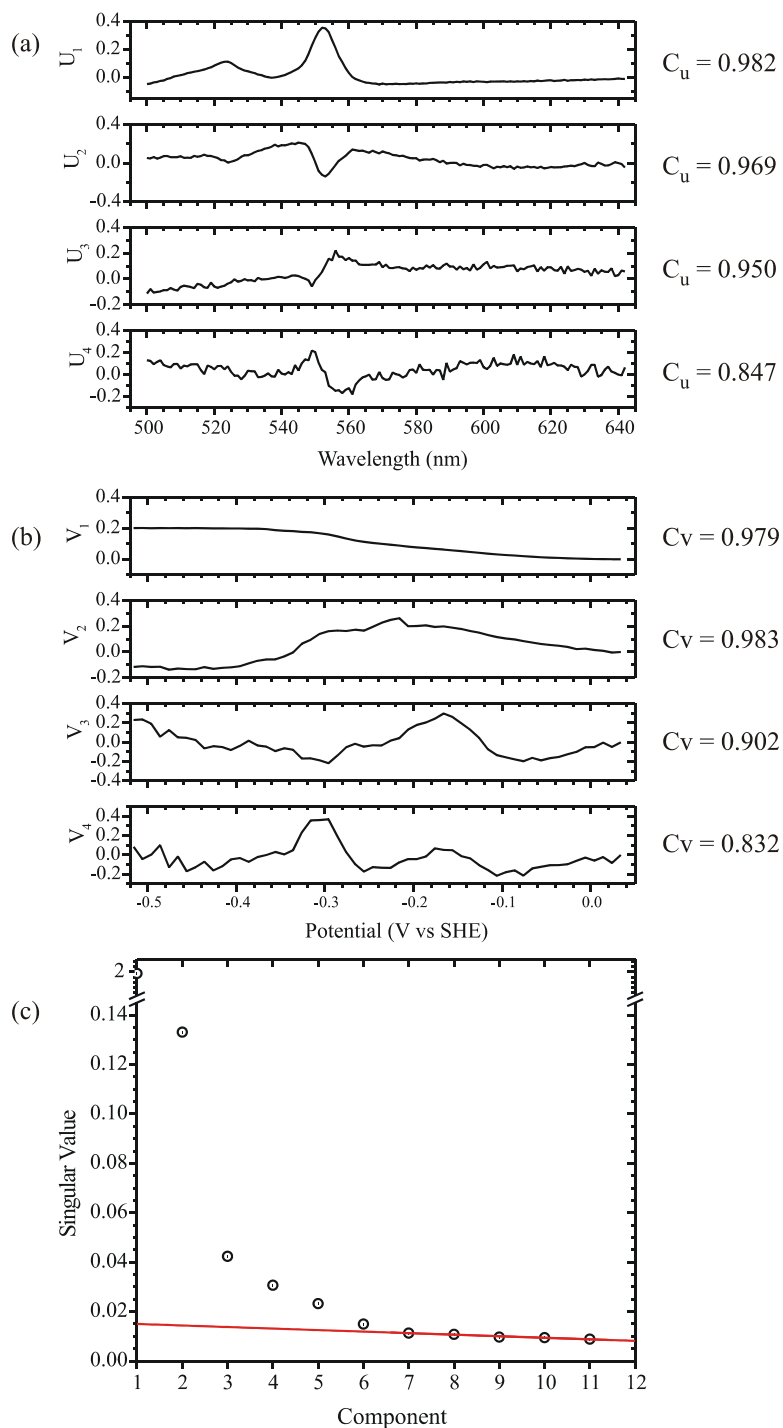
Figure S2. Analysis of the α,β region. (a) The 5 U components corresponding to singular values above the noise level, and their associated autocorrelation values C_U. (b) The 5 V components corresponding to singular values above the noise level, and their associated autocorrelation values C_V. (c) Singular values of the first 11 components.

(a)

Experimental
Calculated

(b)

Aexper
Acalc
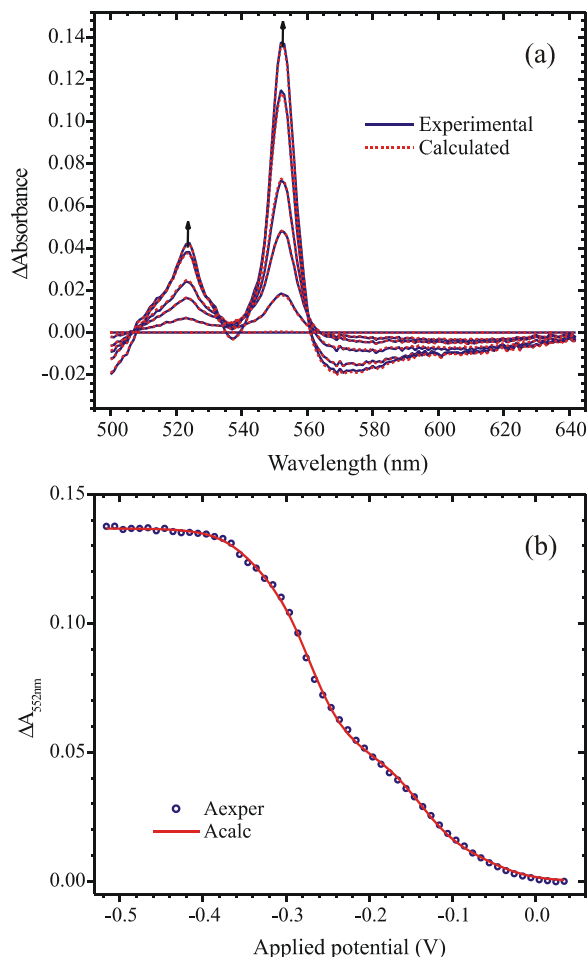
Figure S3. Best least-squares fit of the α,β spectral region data to a model with 5 midpoint potentials but only 4 extinction coefficient difference spectra, where C_3 and C_4 are assumed to have identical spectra.

C_1
C_2
C_3, C_4
C_5
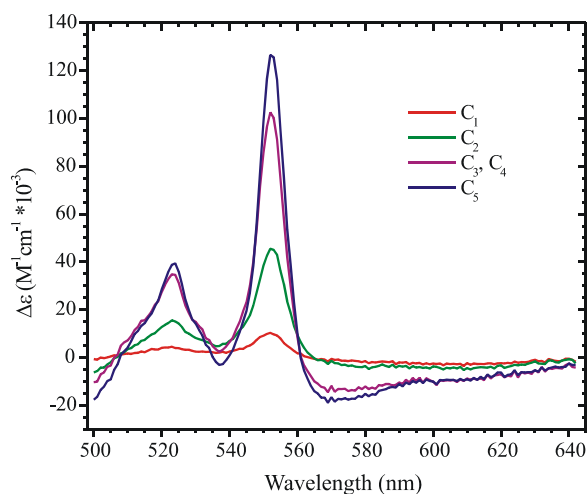
Figure S4. Extinction coefficient difference spectra calculated for the α,β spectral region using the model with 5 midpoint potentials and 4 extinction coefficients.

An alternative explanation for there being only 4 SVD components in the α, β region is that reduction of the high-spin active site (heme 1) does not result in significant spectral changes in this region. Certainly the sharp features seen at 552 nm and 524 nm are characteristic of (isolated) low-spin ferrous hemes and not high-spin ones (ref. [65]). We tried fitting the 500 nm – 650 nm data to 4 potentials using the Soret $\mathscr{E}^o_1$ and $\mathscr{E}^o_3 - \mathscr{E}^o_5$ values as initial guesses, and assuming that the second reduction event did not produce a significant spectral change in this region (this assumes that the second reduction event corresponds to reduction of the high-spin active site heme, as concluded in studies of the *E. coli* protein). However, the fits thus obtained were substantially worse than those obtained with 4 potentials when the $\mathscr{E}^o_3$ is allowed to sit between the Soret $\mathscr{E}^o_3$ and $\mathscr{E}^o_4$ values, and the other midpoint potentials correlate 1:1 with the corresponding Soret values (Table S1). We conclude that every reduction event in *S. oneidensis* ccNiR results in α, β region spectral changes of the type normally associated with reduction of low-spin ferri-hemes to the low-spin ferro form.

S3. Molecular packing in orthorhombic and monoclinic crystal forms
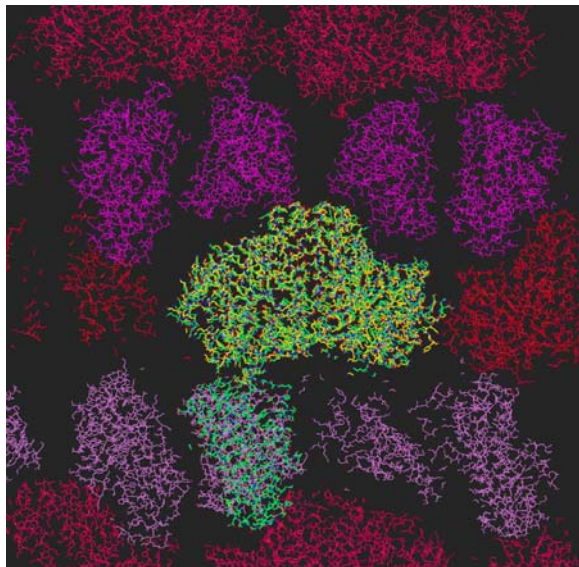
Figure S5. The content of the asymmetric unit in the orthorhombic crystals is shown in yellow, and dimers related by crystallographic symmetry are shown in red and purple tones. One of the two dimers (green) in the asymmetric unit of the monoclinic crystals is superimposed by least squares fitting on the yellow dimer. The non-crystallographically related second dimer is also shown in green. One can easily identify the rotation that breaks the symmetry. If the green dimer were to remain at the same position as the purple dimer, crystals would remain orthorhombic.