Supporting Information

Enzyme Substrate Complex of the H200C Variant of Homoprotocatechuate 2,3-Dioxygenase: Mössbauer and Computational Studies

Katlyn K. Meier,[§] Melanie S. Rogers,[†] Elena G. Kovaleva,[#] John D. Lipscomb,[†] Emile L. Bominaar,[§] Eckard Münck,[§*]*

[§] Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

[†]Department of Biochemistry, Molecular Biology and Biophysics and Center for Metals in Biocatalysis, University of Minnesota, Minneapolis, Minnesota 55455, United States

[#] Stanford Synchrotron Radiation Lightsource, 2575 Sand Hill Road, Menlo Park, California 94025, United States

S1. Strategy for determining the spin Hamiltonian parameters of eqs 1 and 2.

In this section we describe how the range of most of the spin Hamiltonian parameters was narrowed by focusing on particular spectra within the data set.

Consider the 7.5 T/150 K spectrum of Figure 4D. At 150 K the electronic spin of H200C-HPCA relaxes fast compared to the nuclear precession frequencies, and under this condition the magnetic hyperfine field acting at the ^{57}Fe nucleus is given by $\boldsymbol{B}_{int} = -\langle\hat{\boldsymbol{S}}\rangle_{th}\cdot\mathbf{A}/g_n\beta_n$, where $\langle\hat{\boldsymbol{S}}\rangle_{th}$ is the thermally averaged expectation value of the electronic spin.[1] Plots of $\langle\hat{S}_{x,y,z}\rangle_{th}$ *vs.* temperature are shown in Figure S4. At 150 K the effects of the ZFS can be neglected and the system is in the temperature region where the Curie law applies, for which $\langle\hat{S}_i\rangle_{th} = -g_i\beta BS(S + 1)/3kT$ for i = x,y,z; for $D = 0$ and $B = 7.5$ T, $\langle\hat{S}_{x,y}\rangle_{th} \approx -0.134$ and $\langle\hat{S}_z\rangle_{th} \approx -0.144$ for $g_x = g_y = 2.00$ and $g_z = 2.10$ (we justify this value below). These values are modified by $\approx \pm 0.01$ for $|D| = 8$ cm^{-1}. Simulations of the 150 K spectrum (see comments in the caption of Figure 4) using eq 1 showed that (i) $\Delta E_Q > 0$, (ii) η is confined to $0 \le \eta \le 0.5$, (iii) the A-tensor has two large and one small component, (iv) the direction of small component, $A_{z''} \approx 7.5$ T, is within 20° along z′, the major axis of the EFG tensor, and (v) the components of the A-tensor perpendicular to $A_{z''}$ average to $(1/2)(A_{x''} + A_{y''})/g_n\beta_n \approx -23$ T. The properties just found indicate that the ground state of ES is essentially a d_{xy} orbital (see eq 10 of ref. 2).

The ZFS tensor, in a major way, determines the character of the 4.2 K spectra shown in Figure 3 of the main text. Its orientation relative to the EFG- and A-tensors can be obtained from an analysis of the variable field spectra of Figure 3 A-D. In the $B = 0.5$ T spectrum of Figure 3A the high- and low-energy lines of the quadrupole doublet of Figure 2B are split by paramagnetic hyperfine structure. The low-field ($B < 1$ T) spectral pattern suggests that the *lowest* spin level of the $S = 2$ multiplet, designated $2^{s'}$, produces an internal magnetic field, $\boldsymbol{B}_{int} = -\langle\hat{\boldsymbol{S}}\rangle_{th}\cdot\mathbf{A}/g_n\beta_n$, that is nearly uniaxial. Such situations commonly occur for $D < 0$ or for $D > 0$ when E/D is near the rhombic limit, $E/D = 1/3$. For $D < 0$ and $E/D = 0$ the two lowest spin levels, for small B, have $M_S = \pm 2$. For $E/D \neq 0$ the two levels more appropriately are designated as $|2^a\rangle$ and $|2^{s'}\rangle$, where $|2^a\rangle = (|+2\rangle - |-2\rangle)/\sqrt{2}$ and $|2^{s'}\rangle = a^+(|+2\rangle + |-2\rangle)/\sqrt{2} + a^-|0\rangle$, with $a^{\pm} = [(1 \pm D/(D^2 + 3E^2)^{1/2})/2]^{1/2}$ (these expressions are from eqs 4 of Hendrich and Debrunner.[3] The rise of B_{int} for $B < 2.0$ T suggests that the energy gap between the two spin levels is $\Delta \approx 3$ cm^{-1}. The splitting of the low field ($B < 1$ T) spectra depend on $\Delta \approx 3|D|(E/D)^2$. Choosing $E/D \approx 1/3$ yields $D \approx -9$ cm^{-1}; this

estimate is close to the value D = –8 cm^{-1} obtained from simulating the whole data set. Substantially smaller E/D values would yield unreasonably large D values; for instance, E/D = 0.2 would yield D = –25 cm^{-1}, a value not compatible with the spectra of Figure 3A-C. Thus, E/D of HPCD-HPCA is near the rhombic limit 1/3. In the rhombic limit the sign of D depends on the coordinate system chosen. If we take $D < 0$ (this is an arbitrary choice of axes for x,y,z), the lowest spin level $2^{s'}$, for B < 1 T, has an 'unique' axis along z, i. e. $|\langle\hat{S}_z\rangle| >> |\langle\hat{S}_{x,y}\rangle|$ and therefore $|B_{int,z}| >> |B_{int,x,y}|$. Figure S3 shows a plot of spin expectation values for our final parameters D = –8 cm^{-1} and E/D = 0.37. In a 'proper' coordinate system, D = +8 cm^{-1} and E/D = 0.29 (see footnote of Table 1). The important result is that the ZFS, within the uncertainties, is near the rhombic limit.

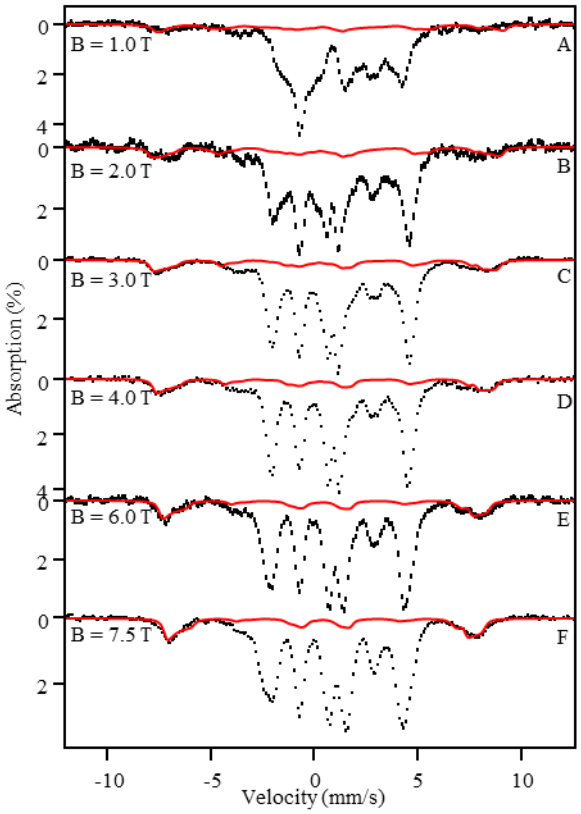

Figure S1. 4.2 K variable field spectra of H200C-HPCA. The black hash-mark curves are the raw experimental data. The solid red curves are simulations for the high-spin ferric contaminants (at least two species) drawn to represent a total spectral area of 15%. The red simulated curves were prepared using the parameters: D = –0.23 cm^{-1}, E/D = 0.33, $\Delta E_{Q,1}$ = 1.24 mm/s, $\Delta E_{Q,2}$ = 0.53 mm/s, $\delta_1 = \delta_2$ = 0.40 mm/s, and A_{iso} = –21.5 T.

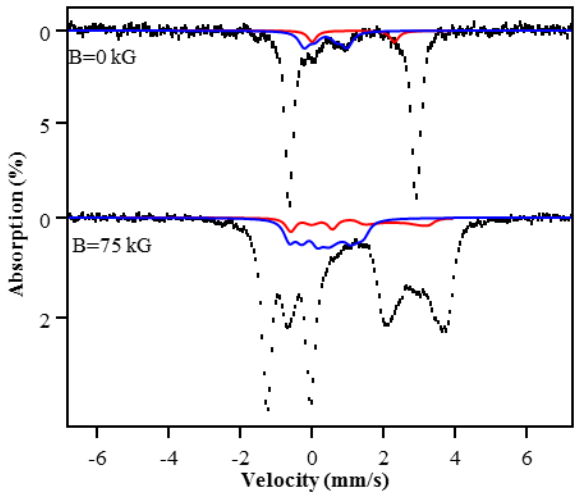

Figure S2. 150 K spectra of complex ES; same data as in Figure 2C and 4D. The blue lines show the (approximate) contributions of the two Fe^{III} contaminants (simulated in the fast relaxation limit at T = 150 K). For the 5% minority Fe^{II} species (red) we took ΔE_Q = +2.33 mm/s and η = 0 and δ = 1.20 mm/s. Note that the high-energy feature of the 7.5 T spectrum is essentially unaffected by the three contaminants, as are the positions of the major bands at 0 mm/s and ≈ –1.2 mm/s. These unaffected features determine the components of $\boldsymbol{B}_{int}$ at 150 K.

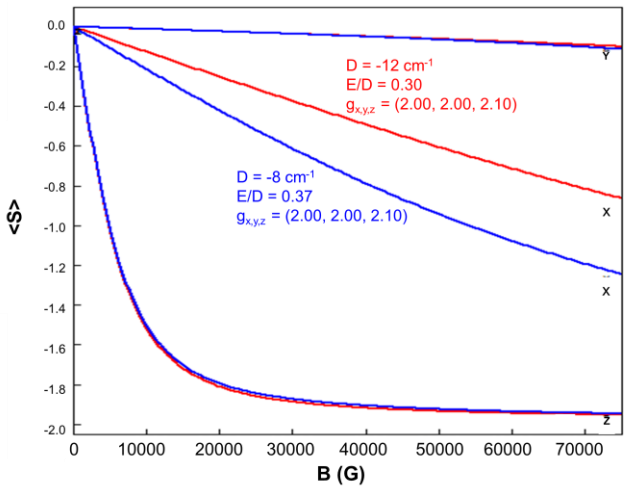

Figure S3. Plots of expectation values $<\hat{S}_{x,y,z}>$ for the lowest state, $|2^{s'}>$, of an S = 2 system for D = –8 cm^{-1}, E/D = 0.37 and D = –12 cm^{-1}, E/D = 0.30, and $g_{x,y,z}$ = (2.00, 2.00, 2.10). These parameters yield Δ_g = 3.27 cm^{-1}. The plot reveals that z is the 'unique' axis for small B.

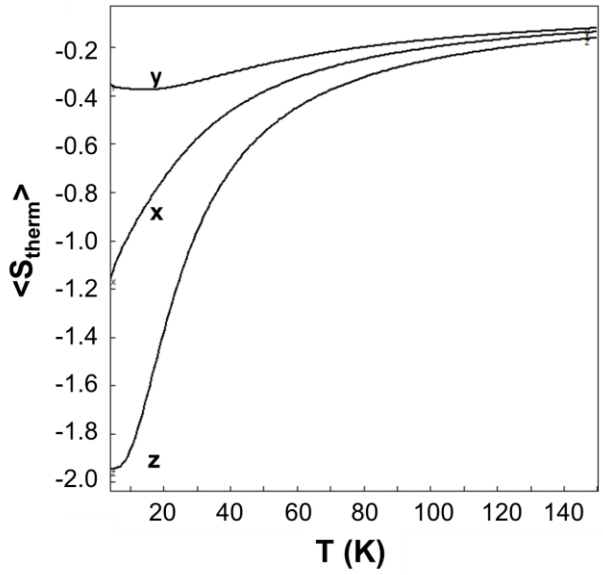

Figure S4. Plot of thermal expectation values $<\hat{S}_{x,y,z}>_{th}$ for $D = -8$ cm^{-1}, $E/D = 0.37$, $g_x = g_y = 2.00$, $g_z = 2.1$, and $B = 7.5$ T. $<\hat{S}_{x,y,z}>_{th}$ was obtained for B along x, y or z. Note that $<\hat{S}_{x,y,z}>_{th}$ is nearly independent of D at 150 K.

S2. Mössbauer spectra of the resting H200C-HPCD.

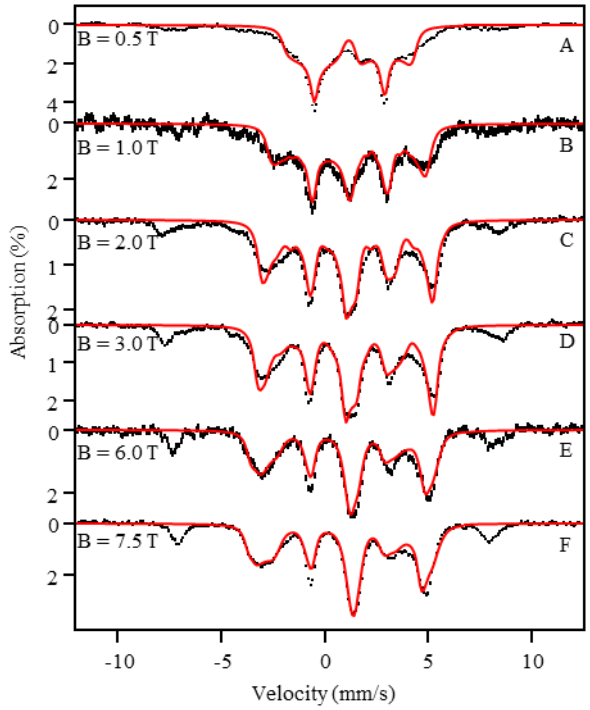

Figure S5. 4.2K variable field spectra of H200C resting enzyme. The black hash-mark curves are the raw experimental data. The solid red curves are simulations to the experimental spectra for the majority, resting enzyme. The red simulated curves were prepared using the parameters of Table 1, namely $D = -7$ cm^{-1}, $E/D = 0.38$, $\Delta E_Q = 3.28$ mm/s, $\eta = 0.2$, $\delta = 1.22$ mm/s, $A_x = -19.9$ T, $A_y = -25.7$ T, $A_z = -8.0$ T.

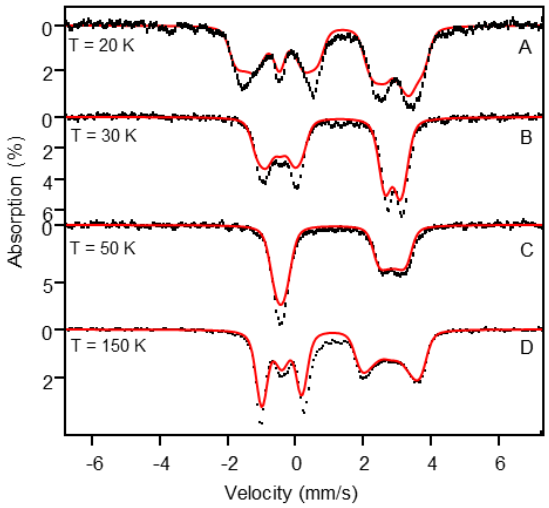

Figure S6. Variable temperature Mössbauer spectra of resting H200C recorded for B = 7.5 T at temperatures indicated. The high-spin ferric contaminants (≈ 15 %) and the minor (5%) ferrous species have *not* been removed from the data.

S3. TD-DFT analysis of ES complex.

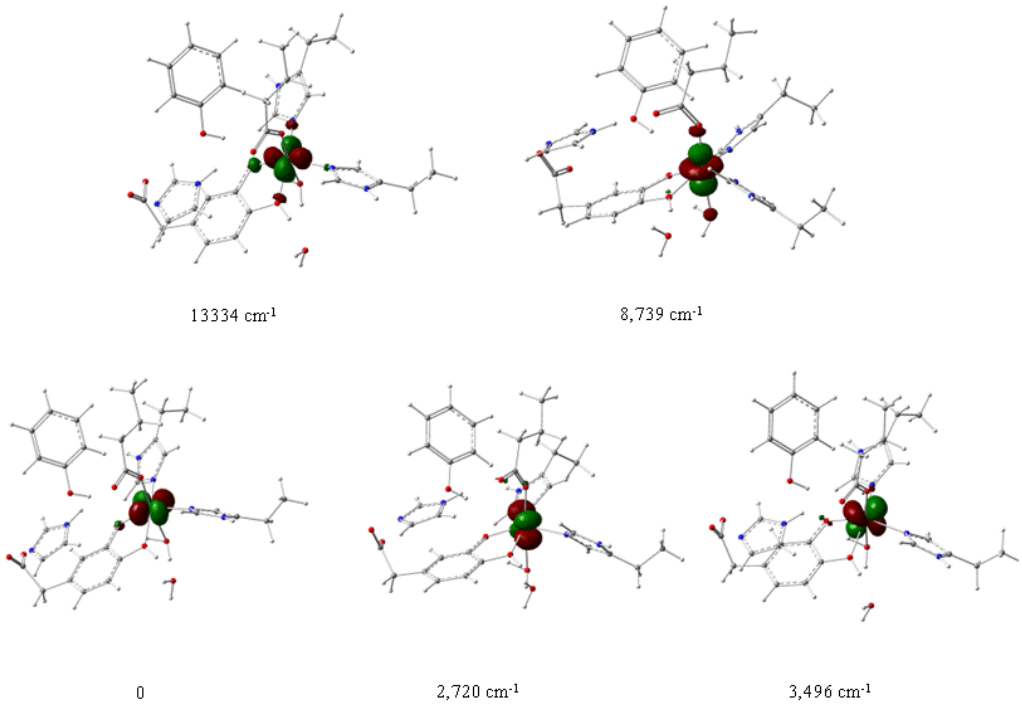

Figure S7. Contour plots of the 3*d* orbital derived from TD-DFT calculations for H200C-HPCA. The numbers are TD-DFT excitation energies.

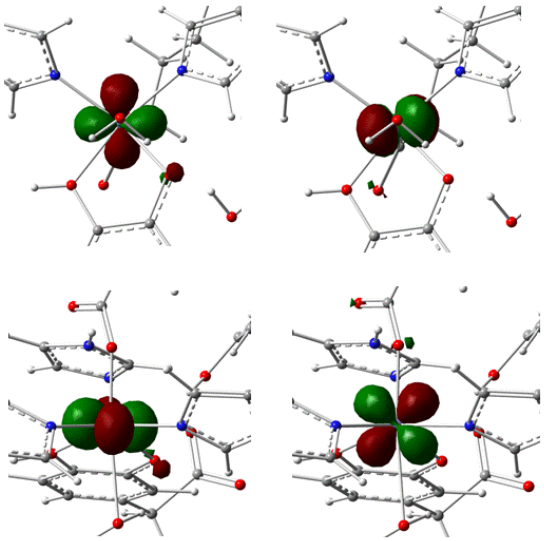

Figure S8. Enlarged contour plots of the two lowest orbitals of H200C-HPCA. Left column: two views of ground state. Right column: two views of first excited state.

S4. Analysis of CF model for obtaining $D < 0$ along z′ axis of largest component of the EFG.

The only way to obtain $D_{zz} < 0$ with the CF model of Figure 6 is by invoking the $d_{xy} \rightarrow d_{x2\text{-}y2}$ excitation. The viability of this mechanism for explaining the large negative D_{zz} for the ES complex can easily be tested. For given values of D and E (with D as defined in eq 1) the excitation energies for which these values are obtained can be expressed as

$$\Delta_{xz} = \lambda^2/q \tag{S1a}$$

$$\Delta_{yz} = \lambda^2/(q + 2E) \tag{S1b}$$

$$\Delta_{x2-y2} = 4\lambda^2/(q + E - D) \tag{S1c}$$

The solutions for the Δ_{di} are not unique but depend on an energy variable q fulfilling the conditions $q \geq \text{Max}\{0, -2E, D - E\}$ in the case $\Delta_{di} > 0$ considered here. The experimental values for D and E imply that $q \geq 6$ cm^{-1}. At the lower limit, $q = 6$ cm^{-1}, one obtains the unrealistic solution $\Delta_{yz} = \infty$, $\Delta_{xz} = 1{,}066$ cm^{-1}, and $\Delta_{x2\text{-}y2} = 2{,}327$ cm^{-1}. When q is raised, the values for Δ_{yz} and $\Delta_{x2\text{-}y2}$ must drop to keep D and E fixed. For example, for $q = 12$ cm^{-1}, one obtains $\Delta_{yz} = 1{,}066$ cm^{-1}, $\Delta_{xz} = 533$ cm^{-1}, and $\Delta_{x2\text{-}y2} = 1{,}506$ cm^{-1}. This analysis provides the condition $\Delta_{x2\text{-}y2} \leq 2{,}327$ cm^{-1}.

The model of Figure 6 can obviously be improved by taking into account that the Fe^{II} site in the ES complex is distorted from ideal octahedral symmetry. These distortions can lead to the

admixture of the e_g orbitals into the t_{2g} orbitals. For example, admixture of d_{x2-y2} into d_{xz} would result in a negative contribution to D_{zz} as desired. These admixtures may change the orbital shape in contour plots of the type shown in Figure 5 or may simply rotate the contour plot without affecting the shape. For example, a rotation of the d_{xz} orbital by 35° around y redirects the unique axis (initially along x) into a direction that makes a polar angle β = 20° with the z-axis.

The DFT contour plots in Figure S8 reveal the familiar appearance of t_2-like 3*d* orbitals. However, a closer consideration reveals some degree of orbital mixing by the action of low symmetry ligand field components as is evident from the shapes of the first and second excited state orbitals (the ones mainly contributing to *D* and *E*). Roughly, these orbitals can be labeled as $d_{x'z}$ and $d_{y'z}$ as they differ from d_{xz} and d_{yz} by an α ≈ 30° rotation around z. This rotation, however, does not produce a negative D_{zz} in the context of the model of Figure 6.

S5. Effective 1-electron SOC operator.

In the main text it was shown that the ZFS tensor of the ES complex calculated by ORCA/CASSCF was in excellent agreement with experiment whereas the tensors predicted by TD-DFT assisted crystal-field theory (CFT) and ORCA/DFT were not, raising questions as to the cause of the difference in the performances of these methods. The fundamental operator for calculating the SOC contribution to the ZFS is the Breit–Pauli (BP) operator (eq S2) which consists of 1- and 2-electron terms:

$$\hat{H}_{\mathrm{SOC}} = \hat{H}^{(1)}_{\mathrm{SOC}} + \hat{H}^{(2)}_{\mathrm{SOC}} \qquad \text{(S2)}$$

The 1e-term represents the contributions due to the nuclear potentials and the 2e-term those due to the electronic potentials. The 1e-term simplifies to the expression in eq S3 (where the symbols have their conventional meanings; *cf.* page 492 of the ORCA 3.0.1 manual),

$$\hat{H}^{(1)}_{\mathrm{SOC}} \approx \frac{\alpha^2}{2}\, Z_{\mathrm{Fe}} \sum_i \frac{\hat{\boldsymbol{l}}_i \cdot \hat{\boldsymbol{s}}_i}{{r_i}^3} \ , \qquad \text{(S3)}$$

provided SOC terms for nuclei that are lighter than iron can be neglected (r_i is the distance of electron *i* to the iron nucleus at the origin). The definition of the 2e-term is given on page 492 of the ORCA manual (ref. 4). Due to the 2e-terms in the BP operator the electrons of the iron core shield the nuclear charge (Z_{Fe}), an effect that can be mimicked by reducing the nuclear charge, so

that for the purpose of calculating the ZFS tensor one could use the effective 1e-operator in eq S4 with effective nuclear charge $Z_{\text{eff}} < Z_{\text{Fe}} = 26$ (see eqs 3 and 4 of ref. 4),

$$\hat{H}_{\text{eff},1} = \frac{\alpha^2}{2} Z_{\text{eff}} \sum_i \frac{\hat{\boldsymbol{l}}_i \cdot \hat{\boldsymbol{s}}_i}{{r_i}^3} \tag{S4}$$

To investigate the accuracy of using $\hat{H}_{\text{eff},1}$ instead of the full BP operator $\hat{H}_{\text{SOC}}$ for calculating the ZFS parameters, the ORCA/DFT results for *D* and *E* in the ES complex as obtained with $\hat{H}_{\text{SOC}}$ and $\hat{H}_{\text{eff},1}$ are listed in Table S1, using for Z_{eff} the value as defined in the SOC module of ORCA. Table S1 shows that the results obtained with $\hat{H}_{\text{SOC}}$ (eq S2) and $\hat{H}_{\text{eff},1}$ (eq S4) are in excellent agreement for both the individual excitations (including states with $S = 1$, 2, and 3) and the net totals for *D* and *E*. As the ZFS tensor is given by a 2$^{\text{nd}}$-order perturbation expression, the scaling of the 1e-operator in eq S3 to yield the effective operator $\hat{H}_{\text{eff},1}$ implies that the values for *D* and *E* obtained with the operator of eq S3 can be scaled to the *D* and *E* values obtained with $\hat{H}_{\text{eff},1}$ (by using as scaling factor the square of the scaling factor for Z_{Fe}). As expected, the values for *D* and *E* obtained with the Z_{Fe}-based 1e-operator $\hat{H}^{(1)}_{\text{SOC}}$ in eq S3 (also been listed in Table S1) are larger by a factor ≈ 3 than those obtained with the Z_{eff}-based operator $\hat{H}_{\text{eff},1}$. The scaling factors (*f*) to be applied to *D*, *E* and *E*/*D* for passing from eq S3 to eq S4, from eq S4 to eq S2, and from eq S3 to eq S2 are given in Table S2. The accuracy of using $\hat{H}_{\text{eff},1}$ as a substitute for $\hat{H}_{\text{SOC}}$ in the calculation of these parameters is reflected by scaling factors near unity, $f(Z_{\text{eff}} \rightarrow \text{BP})$ ≈ 1. The *f* values for *E*/*D* are ≈ 1 because the *f* factors for *D* and *E* are approximately equal. The factor $f(Z_{\text{Fe}} \rightarrow \text{BP}) = 0.30$ (≈ (0.295 + 0.310)/2) accurately scales the results for *D* and *E* as obtained with the Z_{Fe}-based 1e-operator in eq S3 to the values calculated with the full BP operator (eq S2).

Table S1. ORCA/DFT results for *D* and *E* obtained from SOC operators in eq S2, S3, and S4.

BP86		BP (eq S2)		Z_{eff} (eq S4)		Z_{Fe} (eq S3)		rank[a]
def2-tzvp(-f)	*S*	*D*	*E*	*D*	*E*	*D*	*E*	
Spin–Orbit Coupling								
SOMO→VMO	2	0.066	0.012	0.039	0.010	0.134	0.035	4
DOMO→SOMO	2	–0.868	–0.108	–0.823	–0.103	–2.910	–0.362	2
SOMO→SOMO	1	–2.474	–0.266	–2.351	–0.243	–8.232	–0.869	1
DOMO→VMO	3	0.090	–0.009	0.060	0.007	0.188	–0.004	3
1 + 2 [b]		–3.342	–0.374	–3.174	–0.346	–11.251	–1.231	
1 + 2 + 3 + 4 [b]		–3.187	–0.372	–3.075	–0.330	–10.819	–1.199	
Spin–Spin Coupling								
		–0.432	–0.033	–0.432	–0.033	–0.432	–0.033	
SOC + SSC								
		–3.619	–0.405	–3.507	–0.363	–11.251	–1.232	

[a] Ranking according to magnitude of contribution to *D*. The three operators have identical rankings. [b] Sum of terms specified by their ranking.

Table S2. Scaling factors for *D*, *E*, and *E*/*D* relating the results for these quantities as obtained with eq S2 (BP), eq S3 (Z_{Fe}) and eq S4 (Z_{eff}).[a]

	f(X→Y)		
X→Y	*D*-scaling	*E*-scaling	(*E*/*D*)-scaling
$Z_{\text{Fe}} \rightarrow Z_{\text{eff}}$	0.284	0.275	0.97
$Z_{\text{eff}} \rightarrow$ BP	1.036	1.127	1.09
$Z_{\text{Fe}} \rightarrow$ BP	0.295	0.310	1.05

[a] The factors are related, e.g., $f(Z_{\text{Fe}} \rightarrow Z_{\text{eff}}) \times f(Z_{\text{eff}} \rightarrow \text{BP}) = f(Z_{\text{Fe}} \rightarrow \text{BP})$.

As ORCA/CASSCF, to our knowledge, does not offer a module for calculating the SOC contribution to the ZFS tensor on the basis of the Z_{eff}-based $\widehat{H}_{\text{eff},1}$, we have followed an indirect way of testing the utility of this effective Hamiltonian for the purpose of calculating the ZFS parameters: we have multiplied the values for *D* and *E*, obtained with the Z_{Fe}-based 1e-operator in eq S3 (3rd column of Table S3), with the factor 0.30 (4th column of Table S3) deduced from the DFT calculations (see Table S2). The 4th column of Table S3 shows that the 0.3 scaling of the *D* and *E* values obtained with $\widehat{H}_{\text{SOC}}^{(1)}$ in eq S3 (3th column) gives an excellent agreement with the target values obtained with the full BP operator (2nd column). This result implies that, like in the case of DFT, the Z_{eff}-based Hamiltonian $\widehat{H}_{\text{eff},1}$ can also be used for the evaluation of the SOC contribution to the ZFS tensor of the ES complex in the context of CASSCF calculations. Thus,

the unexpected directional properties of the ZFS tensor in the ES complex are independent of the specifics of the 2e-interactions in the BP operator and can be accurately described by means of the effective 1e-operator in eq S4.

Table S3. ZFS parameter D and E/D evaluated with the operators in eq S2 and eq S3, and from scaling (eq S3).

CASSCF	BP (eq S2)	Z_{Fe} (eq S3)	Scaled [a]
$D_{S=2}$ (cm^{-1})[b]	–7.73	–25.43	–7.63
$D_{S=1}$ (cm^{-1})[b]	–1.99	–6.63	–1.99
D (cm^{-1})	–9.72	–32.06	–9.62
E/D	0.245	0.242	0.242 [b]
α (°)	–99.2	–98.9	–98.9 [b]
β (°)	+29.1	+26.9	+26.9 [b]
γ (°)	–69.0	–69.3	–69.3 [b]

[a] D and E from eq 2 multiplied by scaling factor 0.3 (*cf.* Table S2 and discussion). [b] Not affected by the scaling of the ZFS tensor. [b] Contributions to D due to SOC to excited states with the indicated spin.

S6. Wave functions and effective SOC operators used in CFT.

In crystal-field theory Eq S4 is often adopted as the starting point for calculating the ZFS of spin multiplets. In CFT, such as applied to iron complexes, the open shell orbitals are assumed to be *pure* 3*d* orbitals with identical radial functions. Under this assumption, the spatial integration in the evaluation of the matrix elements between $3d^6$ configurations of the operator in eq S4 can be performed by independent integrations over radial and angular coordinates. After radial integration, the operator in eq S4 can be represented by the effective Hamiltonian in eq S5, which acts on the non-integrated angular and spin variables.

$$\widehat{H}_{\text{eff},2} = \frac{\alpha^2}{2} Z_{\text{eff}} \langle r^{-3} \rangle_{3d} \sum_i \hat{\boldsymbol{l}}_i \cdot \hat{\boldsymbol{s}}_i \equiv \zeta \sum_i \hat{\boldsymbol{l}}_i \cdot \hat{\boldsymbol{s}}_i \qquad \text{(S5)}$$

However, in the general case it is questionable whether one can approximate 3*d*-type MOs, such as the active space orbitals of our CASSCF calculations for the ES complex, which are linear combinations of 3*d* (and 4*p*, 4*s*, …) orbitals of iron and ligand based orbitals, by pure 3*d* orbitals. In particular, there is the possibility that by truncating ligand-admixed active space orbitals to pure 3*d* orbitals some of the non-vanishing matrix elements of $\widehat{H}_{\text{eff},1}$ between these active space orbitals are eliminated. The Hamiltonian in eq S5 has non-zero matrix elements between states with different spin (selection rule $\Delta S = 0, \pm 1$) and thus gives non-vanishing contributions to the

ZFS in $3d^6$ $S = 2$ complexes, such as the ES complex, originating from SOC of the ground state with excited $3d^6$ $S = 1$ states. The dismissal of $S = 1$ excitations as a significant source of the ZFS in the ES complex (see main text) allows one to use, instead of $\widehat{H}_{\text{eff},2}$ in eq S5, the effective operator $\widehat{H}_{\text{eff},3}$ in eq S6 for the sake of calculating ZFS parameters (selection rule $\Delta S = 0$). $\widehat{H}_{\text{eff},3}$ has only meaning when acting inside the 25-dimensional space, $(2L + 1) \times (2S + 1) = 25$ for $L = S = 2$, which contains the ground state of the ES complex: It is this Hamiltonian that was adopted in our CF treatment of the ZFS in the ES complex.

$$\widehat{H}_{\text{eff},3} = -\frac{\zeta}{4}\widehat{\boldsymbol{L}} \cdot \widehat{\boldsymbol{S}} \equiv \lambda \widehat{\boldsymbol{L}} \cdot \widehat{\boldsymbol{S}} \tag{S6}$$

Concerning the question as to why CFT performs so poorly in the ES complex, it appears that the difference between the *ab initio* CASSCF and CFT predictions for the axis associated with the negative D value in the ES complex may have arisen from covalent admixtures of non-3*d* orbitals into the active space 3*d* orbitals of iron. These admixtures violate the key premise of CFT (i.e., of having pure 3*d* orbitals) and invalidate the Hamiltonians in eqs S5 and S6. Alternatively, the admixture of 3*d e* orbitals into the t_2 active space orbitals of the CASSCF wave function may be more substantial than suggested by our TD-DFT calculations for the ES complex. Future exploration of these differences may deepen our insight into the metal–ligand bonding interactions in ES and related complexes.

REFERENCES

(1) Münck, E.; Surerus, K.K.; Hendrich, M. *Methods Enzym.* **1993**, *227*, 463-479.

(2) Fox, B. G.; Hendrich, M. P.; Surerus, K. K.; Andersson, K. K.; Froland, W. A.; Lipscomb, J. D.; Münck, E. *J. Am. Chem. Soc.* **1993,** *115*, 3688-3701.

(3) Hendrich, M. P.; Debrunner, P. G. *Biophys. J.* **1989,** *56,* 489-506.

(4) Ganyushin, D.; Neese, F. 2006 *J. Chem. Phys*. **2006**, *125*, 024103-024111.