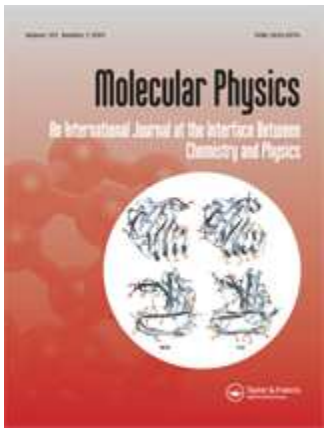
Wide zero field interaction distributions in the high spin EPR of metalloproteins

Wilfred R Hagen



Wide zero field interaction distributions in the high spin EPR of metalloproteins

Journal:	*Molecular Physics*
Manuscript ID:	TMPH-2007-0158.R1
Manuscript Type:	Full Paper
Date Submitted by the Author:	01-Jul-2007
Complete List of Authors:	Hagen, Wilfred; Delft University of Technology, Department of Biotechnology
Keywords:	EPR, strain, zero-field interaction, metalloprotein, high spin

Wide zero field interaction distributions in the high-spin EPR of metalloproteins

WILFRED R HAGEN

Department of Biotechnology, Delft University of Technology, Julianalaan 67,

2628 BC Delft, The Netherlands

Correspndence:

Prof dr Wilfred R Hagen

Department of Biotechnology

Delft University of Technology

Julianalaan 67

2628 BC Delft

Tel: +31 15 2785051

Fax: +31 15 2782355

e-mail: w.r.hagen@tudelft.nl

Abstract

The EPR spin Hamiltonian parameters of transition ion active centres in frozen aqueous solutions of proteins are distributed as a reflection of distributions in spacial conformations. This phenomenon is generally referred to with the generic term ‘g-strain’, however, its manifestation is not limited to a distribution in the g-values. The equivalent name ‘D-strain’ applies to the situation common for biological half-integer high-spin systems whose powder EPR shape is predominantly modified through the distributive effects in the second order terms of the spin-spin interaction Hamiltonian. A simple, one-parameter model is developed to account for D-strain, and this forms the basis for an efficient and practical numerical analysis procedure for D-strained spectra in the weak field limit. Analysis of S = 5/2, 7/2, and 9/2 protein examples is used to illustrate the drastic modification of relative intensities and widths and the occurrence of extra turning points in these spectra as a consequence of D-strain.

1. Introduction

Biomacromolecules have a very high dimensional conformational space, which exhibits rather shallow absolute and relative minima reflecting the high structural flexibility that is presumably required for proper action in key biological events, e.g., catalysis, signal transduction, and regulation of gene expression [1,2]. A frozen-in distribution of conformations is apparently retained in crystallized proteins [3] but also in frozen dilute aqueous solutions [4] which is the common sample form in biomolecular EPR spectroscopy of metalloproteins. In its turn this conformational distribution leads to a distribution in spin-Hamiltonian parameters also known as 'g-strain' [4,5]. Although the ultimate cause of g-strain appears to be describable in terms of a simple, one-dimensional hydrostatic pressure, possibly related to the average size of ice microcrystals in the frozen dilute solution [6], its translated effect on paramagnetic sites through a stress-strain relation via the complex (namely: symmetry lacking) 3D structure of the protein, results in g-strain to be a tensorial quantity not colinear with the g-tensor itself [7]. This implies the existence of two independent interactions (g-strain and the electronic Zeeman interaction) that are linear in the magnetic field, and high quality multi-frequency data are required to separate these from field-independent terms, e.g., hyperfine interactions, for accurate analysis of g-strain [8,9].

It has been realized early on in the development of g-strain analysis that no a priori reason would prevent any other parameter than the g-value in the spin Hamiltonian to also be subject to distribution. However, although some resulting spectral effects have

been identified (e.g., the variation of line width over a set of hyperfine lines [10]), attempts at quantitative simulation of the powder EPR pattern have been rare [11]. This is particularly true for 'D-strain', the name given to a distribution in the axial second order term that frequently dominates the zero-field Hamiltonian [5]. Since the magnitude of |D| in metalloproteins is typically of the order of a wavenumber, a relevant data set for quantitative D-strain analysis should, in addition to X-band (9-10 Hz) spectra, also include spectra taken at significantly higher frequencies and fields. At present, and presumably also in the forseeable future, this poses a technical problem of sensitivity: as a consequence of their high molecular mass (typically 10^5 Da) the maximum concentration of (paramagnetic sites in) metalloproteins is usually limited to values near or below circa 1 mM, and in combination with the very wide field range covered by high-spin spectra when the zero-field interaction is comparable in magnitude to the electronic Zeeman interaction (i.e. at microwave frequencies of the order of 100 GHz) extant high-field/high-frequency EPR spectrometers simply do not have the concentration sensitivity required for meaningful data collection on these systems [12,13]. Indeed, the vast majority of high-field EPR data published on high-spin systems concerns powders of pure (i.e. undiluted) non-biological coordination compounds of limited molecular mass (i.e. typically 1 M or higher in concentration) [14].

In principle, the classical concept of the spin Hamiltonian could provide a framework for the development of a numerical tool for the description of distributed high-spin EPR powder patterns, equivalent to the statistical theory of g-stain for distributed S=1/2 systems. We worked out this premise in 1984 [15] in the form of a computationally intensive powder spectral simulator that would (i) employ a spin

Hamiltonian with second and higher order zero field terms, (ii) diagonalize the energy matrix for every discrete value of the magnetic field scan, (iii) search for pairs of energy levels in the spin manifold that would fit the resonance condition, and (iv) calculate the exact transition probability from the new state vectors. The resulting stick spectrum would then (v) be convoluted with an inhomogeneously broadended line in frequency space, and, finally, (vi) a distribution in the zero field Hamiltonian parameters, or 'D strain', would be implemented under a 'reasonable' model. The question of what in fact is a 'reasonable' model for inhomogeneously broadened high-spin EPR was more recently explored in some detail on multi-frequency (specifically, high-frequency) data sets from non-biological model compounds [16-18], however, this work has yet to lead to a general proposal for D strain in metalloproteins. In the intervening decades we have collected extensive data sets of high-spin powder EPR spectra from a broad variety of metalloproteins, and we have been consistently confronted with the observation that, although spectral peak positions would appear to approximately fit those predicted by appropriate spin Hamiltonians, the actual shape of the powder spectrum, in particularly the relative peak intensities and widths, and sometimes the occurrence of extra features, would remain enigmatic (cf, e.g., the data below). Other workers have also developed computationally intensive simulators (cf [19] and numerous references quoted therein), however, none seem to have been particularly aimed at dealing with the problem that we address here, and which is summarized as follows: Metalloproteins can typically not be prepared in concentrations above 1 mM and, consequently, high-spin EPR data of sufficient quality for spectral analysis, especially for systems of pronounced rhombicity, can frequently be obtained only in X-band. An accurate and unique determination of their

zero field splitting parameters is perhaps not a priority goal of biomolecular spectroscopy where multi-frequency data sets are usually not available, and where a meaningful interpretation of the zero field splitting parameters is yet to be developed [20]. It would, however, be quite useful for the biologist to know what spectral features to group in a single spectral component in order to count the total number of detectable spectral components from multi-centre metalloproteins, and to be able to determine, to reasonable approximation, their relative and absolute stoichiometry, because the signals can then be used as quantitative flags of active centres as a function of biologically relevant parameters, e.g., solution redox potential, pH, substrate concentration, etcetera. It is against this background that we have developed the minimal hypothesis of D-strain in the weak-field limit and its derived simple and rapid simulator VisualRHOMBO for powder EPR of distributed high-spin systems described below.

2. Materials and methods

The proteins and their EPR spectra analyzed here as examples of wide D-strain distributions, have been described: the S = 5/2 systems of *Desulfovibrio vulgaris* (strain Hildenborough) desulfoferrodoxin and its spectral model Fe(III) DTPA (diethylenetriaminepentaacetate) [21]; the S = 7/2 system of *Azotobacter vinelandii* FeMo-nitrogenase [22]; the S = 9/2 system of *Methanothrix soehngenii* (proposed to be renamed as: *Methanosaeta concilii* [23]) carbonmonoxide dehydrogenase [24,25].

The code for the wide D-strain simulator has been written in FORTRAN 90/95 with a graphical user interface for the Windows operating system (Windows XP and later). The program was compiled with the Intel Visual FORTRAN 9.1 compiler integrated in the Microsoft Visual Studio 2005 developer environment. The PC program VisualRHOMBO can be downloaded free of charge as a stand-alone application (single-file executable) from the Departmental website: www.bt.tudelft.nl > Research > Download centre.

3. Results and discussion

3.1. *Theory of D-strain and its numerical analysis*

The minimal spin Hamiltonian for biological half-integer high-spin systems is

$$H = \mathrm{S}\cdot\mathbf{D}\cdot\mathrm{S} + \beta\mathrm{B}\cdot\mathbf{g}\cdot\mathrm{S} \qquad (1)$$

in which **D** and **g** are tensorial quantities. At X-band frequencies ($h\upsilon \approx 0.3$ cm^{-1}) the first term, the second-order zero field interaction, is generally dominant (typically by an order of magnitude or more) over the second term, the electronic Zeeman interaction (six-coordinate Mn^{2+} is the exception [26]). Hyperfine interaction is ignored here (but should be included as a perturbation for, e.g., high-spin Co^{2+} proteins). Superhyperfine interactions are usually not resolved in the type of spectra discussed here. Higher-order terms in the zero-field interactions can, and will, occur for $S \geq 2$, however, it is usually not possible in half-integer high-spin powder spectra in the weak-field limit at a single frequency to separate their contribution from the second-order terms (for integer spins they may not be ignored, cf [15,27]). Iron is the most ubiquitous metal in biological high-spin systems, and all the examples worked out below, are from iron proteins. This means that deviations of the principal elements of the g-matrix from the free electron value, $g_e = 2.00$, will be insignificant for Fe^{3+}, and will be relatively small for clusters containing Fe^{3+} (i.e. all examples, below), and so will be the g-anisotropy, and therefore, the unstrained spin Hamiltonian reduces to

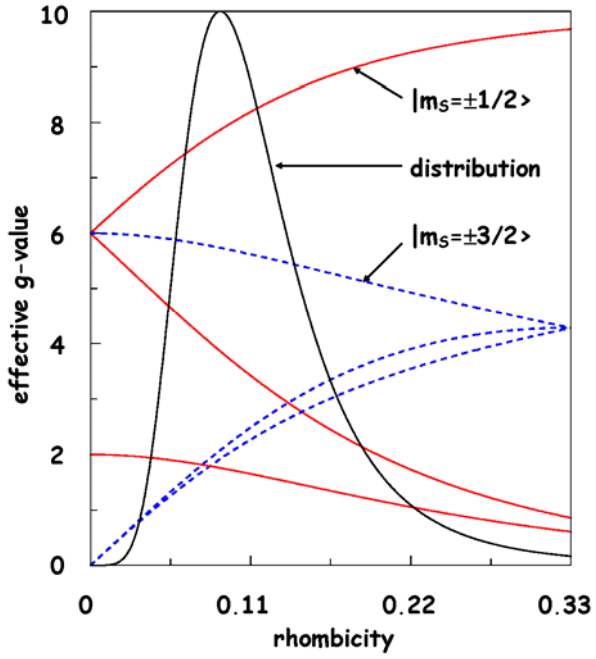
$$H = D[S_z^2 - S(S+1)/3] + E(S_x^2 - S_y^2) + g_{iso}\beta BS \qquad (2)$$

In which the zero field interaction term has been written in its two-parameter form implied by the traceless nature of the D-tensor, and thus all coefficients, D, E, g, are now scalars. The spin manifold in zero field factors into Kramers doublets (i.e. degenerate level pairs).

In the weak-field limit ($D >> 0.3$ cm^{-1}) only intra-doublet transitions are possible in such systems, and it has been realized more than four decades ago, that the resulting (S+1/2) spectra are highly *in*sensitive to the absolute value of D, but very sensitive to the

ratio E/D, also known as the rhombicity η [28]. For high-spin Fe^{3+} ($g \approx g_e$) this means that the complete S = 5/2 sumspectrum is determined by one single parameter only, which, however, does not necessarily imply that the spectra are simple in appearance: each intra-doublet transition is describable by an *effective* spin 1/2 with *three* effective principal g-values. In principle, this affords for S = 5/2 nine, or in general 3(S+1/2) different detectable features (turning points or effective g-values) in the overall powder spectrum. On the other hand, the analysis of these spectra is straightforward with the help of a vertical ruler moved over plots of effective g-values versus rhombicity (also known as rhombograms, cf [29] and in Figure 1).

1



The fact that each individual intra-doublet transition is describable as an effective spin one-half system, i.e. that for each transition the effective Hamiltonian

$$H = \beta B \cdot g_{eff} \cdot S_{eff} \quad (3)$$

applies, means a very considerable simplification of the numerical spectral analysis because the Hamiltonian in eq (2) has to be diagonalized only three times, namely along each of the prinicipal (molecular) axes, and the resonance condition $h\upsilon = g_{eff}\beta B$ for all intermediate orientations, and therefore the complete powder pattern, can be solved analytically as

$$g_{eff} = (\sum l_i^2 g_{i\,eff}^2)^{0.5} \quad (4)$$

with transition probability

$$I = g^{-1}_{eff} \sum [g_{i\,eff}^2 - (l_i^2 g_{i\,eff}^4 / g^2_{eff})] \quad (5)$$

in which the summation is over the molecular axes x, y, z, and l_i are the direction cosines between the magnetic field vector B and the molecular axes (cf [27,29]).

With the implementation of a few earlier developed numerical 'tricks', notably, tabulation of the line-shape function in a large array to avoid massively repetitive calculations of exponentials and of floating point interpolations [30], smart coding (do-loop unrolling), and the use of an efficient interpolation scheme to scan the unit sphere (the 'Igloo' method due to R.L. Belford's group, cf paragraph 5.3.2 in [31]), a simulator for *un*distributed half-integer spin spectra can be set up that typically generates a complete powder pattern for up to S = 9/2 in less than one second on an off-the-shelf PC. Real-time fitting of D-strain spectra now becomes practical: a distribution in rhombicity is set up as follows. Both zero field interaction parameters, D and E, are subject to a normal distribution with standard deviation Δ_D and Δ_E expressed as percentage of the

average values of D and E. A full positive correlation between these Δ's would lead to no broadening at all in the weak-field EPR spectra as they depend only on the rhombicity, the ratio E/D. On the other hand, the most pronounced broadening would occur in case of full negative correlation, $\Delta_D = -\Delta_E$, and this definition reduces the fitting parameters in the strain model to one. Below, it is shown that this model affords semi-quantitative fits to the powder EPR spectra of several high-spin metalloproteins. Note also that in previous extensive multi-frequency analyses of g-strain in metalloprotein EPR full negative correlation of strain parameters was found to be the predominant, if not the only case of relevance [6,7,9,30].

In practice, the average of D is set at a high dummy value, $D >> 0.3$ cm^{-1}, for example $|D| \equiv 10$ cm^{-1}, the average of E is defined through the rhombicity, $E = \eta D$, amd the sign of E is taken to be opposite to that of D. The fully negatively correlated normal distributions in D and E now lead to an asymmetric distribution on the rhombicity η, defined by a single parameter Δ ($= |\Delta_D| = |\Delta_E|$). An example is given in figure 1 for $\eta = 0.09$ and $\Delta = 15.6\%$, and with the individual actual values of D and E undetermined.

To obtain the strained EPR powder spectrum, for each set of (D,E) values eq (2) is diagonalized, and eqs (4) and (5) are solved. All resulting spectra are then summed with proper weighing according to the intensity of the rhombicity distribution to obtain the D-strained spectrum, which on a standard PC takes circa one minute or less for up to S = 9/2. Since in many practical cases the rhombicity is readily read out, or at least approximately so, from experimental peak positions, the complete D-strain simulation is de facto dependent only on a single parameter, Δ. In this approach the line width of the individual line shape turns into a 'semi-parameter': its value becomes increasingly

irrelevant with increasing D-strain, except perhaps for turning point features at low field (high effective g-value) which prove to be relatively insensitive to D-strain (see below).

In practice then, the procedure is as follows: (i) an approximate real g-value is estimated from EPR theory (e.g., $g \approx 2.00$ for hs Fe^{3+}, $S = 5/2$; $g \approx 2.0 \pm 0.1$ for hs Fe/S clusters, $S \leq 9/2$; $g \approx 2.2 \pm 0.1$ for hs Co^{2+}, $S = 3/2$, etc.); (ii) a fairly accurate estimate of the average rhombicity, $\eta = E/D$, is made from the peak positions in the experimental spectrum (in particular the low-field ones) using rhombograms; (iii) the individual inhomogeneous line width, γ, is set at a dummy value significantly less than the apparent line width, Γ, of the lowest field feature of the experimental D-strained spectrum (see below); (iv) the distribution width Δ is determined interactively by repeated powder spectrum simulation with the VisualRHOMBO program; (v) the fit may be fine tuned with minor adjustments to g_{real}, η, γ, and subsequently Δ, and perhaps with minor adjuistments (increases) of the number of sampled molecular orientations (typically 10^4-10^5) and the number of sampled values in the rhombicity distribution (typically 10^2) to eliminate mosaic artifacts. The whole interactive fitting procedure should converge in well less than an hour; otherwise a mistake is indicated in the assignment of the system spin.

Note that there is an ambiguity in the 'units' in which Δ can be expressed. Since in the weak-field limit the actual value of $|D|$ is undetermined, and $\eta = E/D$ is dimensionless, Δ can be conveniently expressed as a percentage of (D,E). However, when the system is of (near) axial symmetry, i.e. $E \approx 0$, and thus $\eta \approx 0$, then Δ becomes the (very large) percentage deviation of a very small number. For these systems Δ may perhaps preferably be expressed in energy units (e.g., reciprocal centimeters), i.e. in

absolute standard deviations of D and E. The range of possible physically meaningful values for the rhombicity is theoretically limited to $0 \leq \eta \leq 1/3$ [28]. In all the examples worked out below the average rhombicity is $\eta > 0.05$, and Δ is reported as percentage of (D,E).

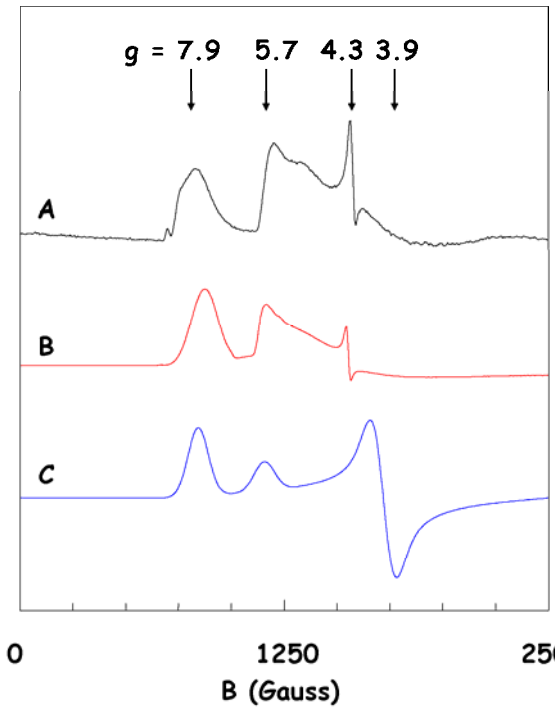
The key relevance of the above D-strain analysis is to provide insight in the shape of powder EPR spectra from metalloproteins and their models, in particular as to which spectral features combined make up the spectrum of a single high-spin system, and what are their relative amplitudes. In its turn this information affords assignment of spectra to active centers, to quantitate their concentrations (i.e. spin counting) and to monitor amplitudes of assigned peaks as function of biochemically relevant external parameters, e.g., redox potentials.

3.2. *An S = 5/2 example: multiplicity of D-strained spectra*

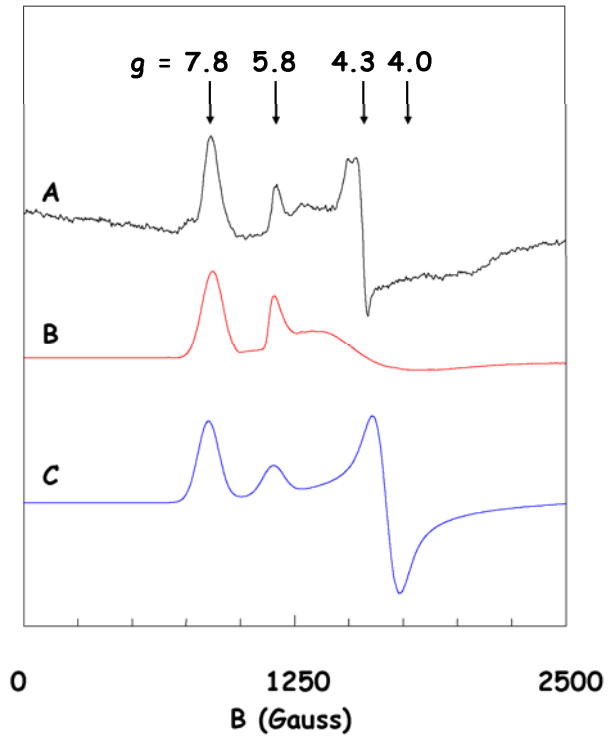
Sulfate reducing micororganism contain a small (2x14 kDa homodimer) protein called desulfoferrodoxin [32], which functions as a superoxide reductase [33]. The enzyme carries two iron ions per subunit: center I with coordination by four cysteine ligands, $Fe(S^{\gamma}_{Cys})_4$, and center II in which the Fe is coordinated by one cysteine and four histidines, $FeS^{\gamma}_{Cys}N^{\delta}_{His}(N^{\varepsilon}_{His})_3$ [34]. The iron of center II, the site of superoxide activation, is in the ferrous Fe(II) state before reaction with the substrate, but becomes Fe(III)-hydroxide at high pH [35]. The iron is also slowly oxidized by air during aerobic protein purification.

For the ferric Center II the low-field peak positions have been assigned to an S = 5/2 system with rhombicity $\eta = E/D \approx 0.08$ [32]. The spectrum was later found to be remarkably similar to that of Fe(III) DTPA (diethylenetriaminepentaacetic acid) [21], which implies $\eta \approx 0.08$ also for this spectroscopic model, however, confirmation of these assignments by powder spectral simulation was not attempted. In point of fact, as can be seen in figures 2 and 3, traces C,

2



3



a simulation based on this rhombicity value may approximately reproduce the field positions of the first two low-field lines, but the relative width and, for FeDTPA, the intensity of the second line is poorly reproduced. Furthermore, a strong, derivative shaped feature at higher field (g_{eff} = 4.1), generated in the simulations, is absent in the experimental spectrum of the protein and the model.

The introduction of a Gaussian distribution in (D,E) around their central moments, corresponding to the rhombicity value originally obtained from rhombogram analysis of low-field peaks, drastically changes the overall simulated spectral shape and affords a semi-quantitative reproduction of the experimental spectra (figures 2 & 3, traces

B; table 1). In particular, a sharp feature at $g_{eff} = 4.3$ in the FeDTPA spectrum is reproduced attesting to the fact that this line is not from a contaminant of high rhombicity, but rather reflects the wide nature of the distribution that essentially samples *all* possible rhombicities (cf. figure 1). The occurrence of a g = 4.3 line as an intrinsic part of the distributed spectra of high-spin S = 5/2 model compounds has been recognized before [36], however, notice that the line is *absent* in the simulated spectrum of the desulfoferrodoxin protein, although the experimental spectrum exhibits a relatively strong feature at this effective g-value. This is consistent with our previous observation that the second iron, center I, is partially oxidized in the as isolated protein and has a 'rhombic' EPR spectrum centered around g = 4.3 [21]. Apparently, D-strain simulation is a necessity to be able to discriminate between a contaminating and an intrinsic g = 4.3 line.

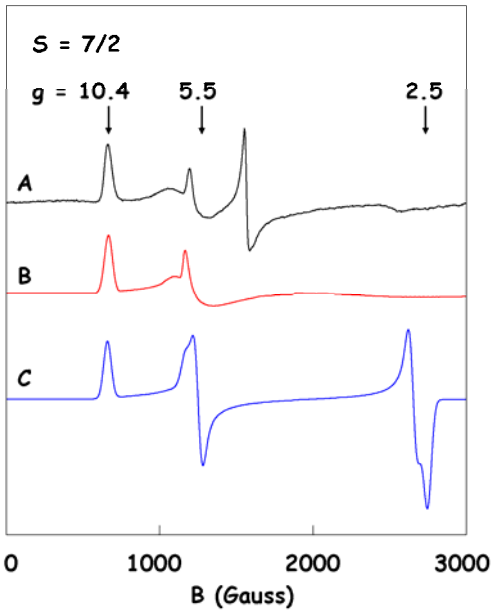
3.3. *An S = 7/2 example: quantification of D-strained spectra*

Two decades ago the P-cluster prosthetic group in nitrogenase was audaciously proposed to be a 'supercluster' containing eight iron ions. The proposal was based on temperature-dependent, quantitative EPR spectroscopy of thionine-oxidized enzyme, in which the cluster exhibits S = 7/2 paramagnetism with D = -3.7 cm^{-1} [22]. The assignment of S = 7/2 was based on a rhombogram fit to the peak positions at low field (g_{eff} = 10.4, 5.8, 5.5), and spin counting was done by singly integrating the lowest-field absorption shaped peak at g = 10.4 using the single-peak integration method propsed by Aasa and Vänngård [37] with reference to the EPR spectrum of an external standard of known concentration.

This afforded approximately two P-clusters per nitrogenase, and with circa 16 Fe availabe based on analytical chemistry (in addition to the iron in the FeMo-cofactor), the first identification of a biological 8Fe cluster was straightforward.

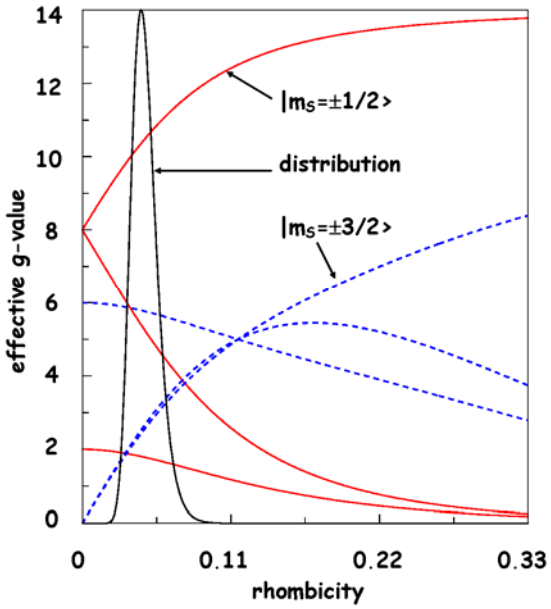
However, the conclusion would perhaps have been less straightforward if one would have tried to actually reproduce the powder spectrum by simulation as shown in figure 4, trace C:

4



low field peak positions do approximately fit, but the shape around g = 5.5 is poorly reproduced. Furthermore, the simulation shows a very pronounced feature at g = 2.5 where the experimental spectrum in trace A only exhibits perhaps a minor declination. Introduction of a widely distributed rhombicity in the simulation (trace B) comes to the rescue of the spin assignment, where the spectrum is well reproduced on basis of an S = 7/2 with rhombicity $\eta = 0.044$ and a standard deviation of $\Delta(D,E) = 8\%$ (cf figure 5).

5



Furthermore, comparison of the distributed (trace B) and undistributed (trace C) simulation shows that the spectral effects of D-strain are more pronounced towards higher fields (lower effective g-values), and that the lowest-field peak at g = 10.4 is the only feature whose shape remains essentially unaffected, which implies that the original quantification incidentally has been carried out properly, and that the 'supercluster' proposal is corroborated. This argument has been put on a quantitative footing in the last column of Table 1, which gives the intensity ratio of the simulated lowest field peak in the presence and the absence of strain: a ratio of unity implies the Aasa-Vänngård single-peak integration method to be applicable also to the strained spectrum. Generalizing; spin

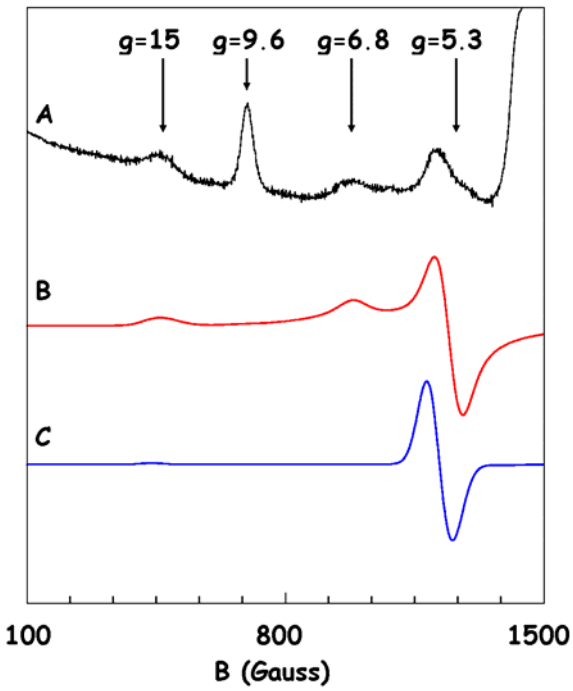
counting using D-strained spectra should be done only on the first integral of the lowest-field peak (or, alternatively, via full simulation of the D-strained powder spectrum).

3.4. *An S = 9/2 example: intensities and extra turning points in D-strained spectra*

The unusually high spin S = 9/2 has been identified repeatedly in proteins [24,38,39], however, the basis of the assignment has not always been fully convincing in these experimentally demanding systems of high structural complexity. Not only do the S = 9/2 spectra typically exhibit several features of relatively low intensity over wide field ranges, they are also partially blocked out by interference of overlapping spectra from other prosthetic groups or contaminants in the same enzyme with S < 9/2 covering a less wide field range and, therefore, usually exhibiting a higher amplitude than the S = 9/2 spectra.

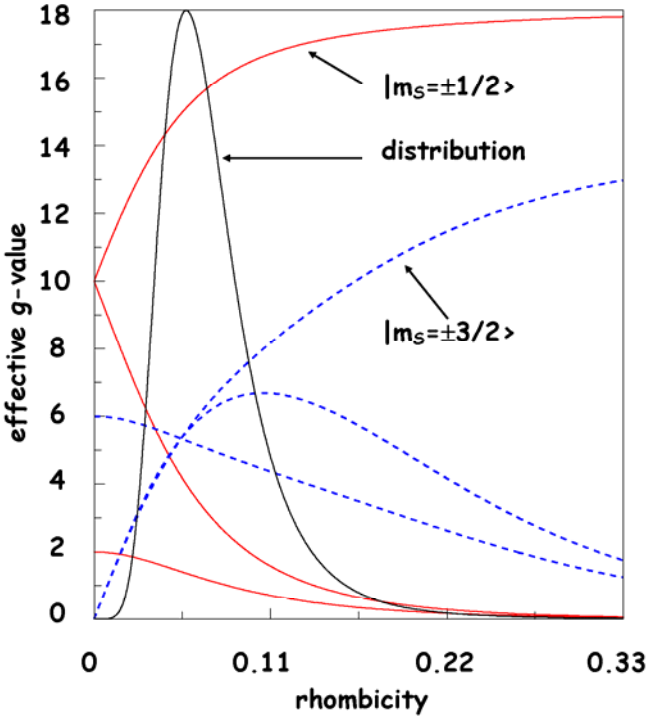
A characteristic example is the complex $\alpha_2\beta_2$ heterotetrameric enzyme CO dehydrogenase (acetyl-coenzymeA cleaving) of *M. soehngenii* [27,28]. The EPR spectrum has several low-field peaks of relatively low intensity compared to a rhombic S = 5/2 contaminant (g = 9.6, 4.3; $\eta \approx 1/3$, cf figure 1) that partially blocks the S = 9/2 spectrum and deforms a derivative feature at g ≈ 5.3 (figure 6).

6



The S = 9/2 assignment is indicated by rhombogram analysis predicting g_{eff} = 15 and 5.3 for the lowest-field peak of the |±1/2> and the |±3/2> intradoublet transition, respectively, for η = 0.047 ([27] and figure 7).

7



Remains a peak at $g_{eff} = 6.8$ which was taken at the time to indicate some form of inhomogeneity and was left unassigned [24]. The undistributed and the distributed S = 9/2 simulations based on $\eta = 0.058$ and $\Delta = 15\%$ are presented in traces C and B of figure 6. Introduction of D-strain does not only correct an improper signal intensity ratio of the g = 15 and 5.3 peaks in the undistributed simulation, but it also generates a peak at g = 6.8, eliminating the need to consider spectral inhomogeneity. The physical nature of the g = 6.8 feature in the powder spectrum can be traced back in figure 7 to originate in a turning point at intermediate rhombicity ($\eta \approx 0.11$) of one of the rhombogram traces for S = 9/2, $m_S = \pm 3/2$. Such a turning point implies the effective g-value to be invariant in

small changes in the rhombicity, and this will lead to a relatively high EPR amplitude in a broad sampling of rhombicities. In fact, the g = 6.8 line is an S = 9/2 equivalent of the g = 4.3 line that frequently occurs in distributed S = 5/2 systems of intermediate rhombicity (e.g., in figure 2).

3.5. *Biological relevance of wide strain distributions*

A biological interpretation of wide D-strain, if any, is yet to be established. The concept of strain, notably in strained protein conformation, is not unfamiliar to the biochemist, and can perhaps be traced back to the original proposal of Vallee and Williams of an 'entatic' (i.e. strained) state of active centres, based on the observation of unusual optical and EPR properties, to explain, e.g., the apparent minimal reorganization energy of the Cu(I/II) redox transition in blue copper proteins [40,41]. It has later been questioned, on the basis of x-ray data and theoretical calculations, whether blue copper centres are indeed entatic at all [42]. From the perspective of the present work one can wonder whether the width of a distribution in strain (the width of a distributed state whose average conformation may or may not be entatic) would not be of equal, or even greater biological importance than its average magnitude. In terms of biological functionality this would translate into the question whether the nature and width of the distribution is such that it can sample, with significant frequency, geometrical conformations close to the transition state. Such a hypothesis should in principle be testable by relating, e.g., EPR-

detected D-strain to properties, such as reduction potential values, hitherto related to the existence of a single entatic state.

4. Conclusions

D-strain analysis of half-integer high-spin powder EPR data from biological systems and their models, based on the minimal hypothesis of a distribution in rhombicity, derived from a joint Gaussian distribution in D and E, defined in the weak-field limit, affords numerically rapid, semi-quantiative reproduction of experimental spectra. The determination of spin concentration, or spin counting, should be based on the first integral of the lowest-field peak or on full simulation of the D-strained spectrum. D-strain analysis is a mandatory requirement for an understanding of relative intensities and widths of peaks in the powder spectra as well as the appearance of extra features due to turning points in the rhombograms. Such an understanding is required for a biologically meaningful interpretation of data in terms of stoichiometry and multiplicity of protein prosthetic groups. Finally, strain in metalloprotein EPR reflects a conformational distribution that is possibly of relevance for biological functionality namely as a route to sample transition state geometry.

References

[1] R.H. Austin, K.W. Beeston, L. Eisenstein, H. Frauenfelder, I.C. Gunsalus. *Biochemistry*, **14**, 5355 (1975).

[2] H. Frauenfelder. *Proc. Natl. Acad. Sci. USA*, **99**, 2479 (2002).

[3] H. Frauenfelder, G.A. Petsko, D. Tsernoglou. *Nature*, **280**, 558 (1979).

[4] J. Fritz, R. Anderson, J. Fee, G. Palmer, R.H. Sands, J.C.M. Tsibris, I.C. Gunsalus, W.H. Orme-Johnson, H. Beinert. *Biochim. Biophys. Acta*, **253**, 110 (1971).

[5] W.R. Hagen. *J. Magn. Reson.*, **44**, 447 (1981).

[6] W.R.Hagen. In *Advanced EPR; Applications in Biology and Biochemistry*, A.J. Hoff (Ed), p. 785, Elsevier, Amsterdam (1989).

[7] W.R. Hagen, D.O. Hearshen, R.H. Sands, W.R. Dunham. *J. Magn. Reson.*, **61**, 220 (1985).

[8] W.R. Hagen, S.P.J. Albracht. *Biochim. Biophys. Acta*, **702**, 61 (1982).

[9] D.O. Hearshen, W.R. Hagen, R.H. Sands, H.J. Grande, H.L. Crespi, I.C. Gunsalus, W.R. Dunham. *J. Magn. Reson.*, **69**, 440 (1986).

[10] W. Froncisz, J.S. Hyde. *J. Chem. Phys.*, **73**, 3123 (1980).

[11] W.R. Hagen. *EPR spectroscopy of metalloproteins with emphasis on respiratory chain complexes*. PhD thesis, University of Amsterdam. (1982).

[12] W.R. Hagen. *Coord. Chem. Rev.*, **190-192**, 209 (1999).

[13] W.R. Hagen. In *High Magnetic Fields; Science and Technology Vol. 3*, F. Herlach and N. Miura (Eds), p. 207, World Scientific, Signapore (2006).

[14] J. Krzystek, A. Ozarowski, J. Telser. *Coord. Chem. Rev.*, **250**, 2308 (2006).

[15] W.R. Hagen, W.R. Dunham, R.H. Sands, R.W. Shaw, H. Beinert. *Biochim. Biophys. Acta*, **765**, 399 (1984).

[16] P.J. van Dam, A.A.K. Klaassen, E.J. Reijerse, W.R. Hagen. *J. Magn. Reson.*, **130**, 140 (1998).

[17] A. Priem, P.J.M. van Bentum, W.R. Hagen, E.J. Reijerse. *Appl. Magn. Reson.*, **21**, 535 (2001).

[18] A. Priem. *Exprorations in high-frequency EPR*, PhD thesis, University Nijmegen, http://webdoc.ubn.kun.nl/mono/p/priem_a/explinhif.pdf (2002).

[19] G.R. Hanson, K.E. Gates, C.J. Noble, M. Griffin, A. Mitchell, S. Benson. *J. Inorg. Biochem.*, **98**, 903 (2004).

[20] F. Neese. *J. Am. Chem. Soc.*, **128**, 10213 (2006).

[21] M.F.J.M. Verhagen, W.G.B. Voorhorst, J.A. Kolkman, R.B.G. Wolbert, W.R. Hagen. *FEBS Lett.*, **336**, 13 (1993).

[22] W.R. Hagen, H. Wassink, R.R. Eady, B.E. Smith, H. Haaker. *Eur. J. Biochem.*, **169**, 457 (1987).

[23] P. de Vos, H.G. Trüper, B.J. Tindall. *Int. J. Syst. Evol. Microbiol.*, **55**, 525.

[24] M.S.M. Jetten, A.J. Pierik, W.R. Hagen. *Eur. J. Biochem.*, **202**, 1291 (1991).

[25] R.I.L. Eggen, R. van Kranenburg, , A.J.M. Vriesema, A.C.M. Geerling, M.F.J.M. Verhagen, W.R. Hagen, W.M. de Vos. *J. Biol. Chem.*, **271**, 14256 (1996).

[26] J.S. Shaffer, H.A. Farach, C.P. Poole, Jr. *Phys. Rev. B*, **13**, 1869 (1976).

[27] W.R. Hagen. *Dalton Trans.*, **2006**, 4415 (2006).

[28] G.J. Troup, D.R. Hutton. *Brit. J. Appl. Phys.*, **15**, 1493 (1964).

[29] W.R. Hagen. *Adv. Inorg. Chem.*, **38**, 165 (1992).

[30] W.R. Hagen, D.O. Hearshen, L.J. Harding, W.R. Dunham. *J. Magn. Reson.*, **61**, 233 (1985).

[31] J.R. Pilbrow. *Transition Ion Electron Paramagnetic Resonance*. Clarendon Press, Oxford (1990).

[32] I. Moura, P. Tavares, J.J.G. Moura. *J. Biol. Chem.*, **265**, 21596 (1990)

[33] C.V. Romão, M.Y. Liu, J. Le Gall, C.M. Gomes, V. Braga, I. Pacheco, A.V. Xavier, M. Teixeira. *Eur. J. Biochem.*, **261**, 438 (1999).

[34] A.V. Coelho, P.M. Matias, V. Fülöp, A. Thompson, A. Gonzalez, M.A. Carrondo, *J. Biol. Inorg. Chem.*, **2**, 680 (1997).

[35] C. Mathe, V. Niviere, T.A. Mattioli. *J. Am. Chem. Soc.*, **127**, 16436 (2005).

[36] J.T. Weisser, M.J. Nilges, M.J. Sever, J.J. Wilker. *Inorg. Chem.*, **45**, 7736 (2006).

[37] R. Aasa, T. Vänngård. *J. Magn. Reson.*, **19**, 308 (1975).

[38] A.J. Pierik, W.R. Hagen. *Eur. J. Biochem.*, **195**, 505 (1991).

[39] A.J. Pierik, W.R. Hagen, W.R. Dunham, R.H. Sands. *Eur. J. Biochem.*, **206**, 705 (1992).

[40] B.L. Vallee, R.J.P. Williams. *Proc. Natl. Acad. Sci. USA*, **59**, 498 (1968).

[41] R.J.P. Williams. *Eur. J. Biochem.*, **234**, 363 (1995).

[42] U. Ryde, M.H.M. Olsson, K. Pierloot, B.O. Roos. *J. Mol. Biol.*, **261**, 586 (1996).

Table 1: Simulation parameters of wide D-strain distributions in the powder EPR of high-spin metalloproteins

Species	rhombicity η	strain width Δ (%)	dummy width[a] γ (Gauss)	line width[b] Γ (Gauss)	intensity ratio[c]
S = 5/2					
desulfoferrodoxin	0.085	12.0	5	40	1.04
Fe-DTPA	0.090	15.6	5	40	0.84
S = 7/2					
nitrogenase P-cluster	0.044	8.0	5	20	0.96
S = 9/2					
CO dehydrogenase	0.058	15.0	25	25	0.79

Note: a) line width of distributed spectrum; b) line width of undistributed spectrum; c) ratio of intensities of the lowest lield peak in distributed over undistributed spectra.

Legends to the figures

Figure 1. Rhombogram for S = 5/2: effective g-values as a function of rhombicity, $\eta = E/D$. The solid traces are the three effective g-values for the $m_S = \pm 1/2$ doublet; the broken traces are the g-values for the $m_S = \pm 3/2$ doublet. In this and following rhombograms only g-values for the lowest two doublets are plotted, because simulations show that the intra-doublet transitions in higher doublets contribute negligible intensity to the powder EPR spectrum. The trace labeled 'distribution' is the distribution in rhombicity resulting from a joint normal distribution in D and E about the central value of $\eta = 0.090$ with $\Delta = 15.6\%$. These values apply to Fe(III)DTPA. The amplitude of the distribution is arbitrarily set to a value of 10 to fit in the rhombogram.

Figure 2. Experimental and simulated spectrum of Fe(III)DTPA. Trace A, experimental spectrum at 9.30 GHz [21]; trace B, distributed simulation; trace C, undistributed simulation (parameters in table 1).

Figure 3. Experimental and simulated spectrum of *D. vulgaris* desulfoferrodoxin. Trace A, experimental spectrum at 9.32 GHz [21]; trace B, distributed simulation; trace C, undistributed simulation (parameters in table 1).

Figure 4. Experimental and simulated spectrum of the P-cluster in thionine-oxidized *A. vinelandii* FeMo-nitrogenase. Trace A, experimental spectrum at 9.30 GHz [22]; trace B, distributed simulation; trace C, undistributed simulation (parameters in table 1).

Figure 5. Rhombogram for S = 7/2. The solid traces are for the $m_S = \pm 1/2$ doublet, the broken traces are for the $m_S = \pm 3/2$ doublet. The rhombicity distribution is for $\eta = 0.44$ with $\Delta = 8.0\%$ and applies to the P-cluster.

Figure 6. Experimental and simulated spectrum of *M. soehngenii* CO dehydrogenase. Trace A, experimental spectrum at 9.30 GHz [27]; trace B, distributed simulation; trace C, undistributed simulation (parameters in table 1).

Figure 7. Rhombogram for S = 9/2. The solid traces are for the $m_S = \pm 1/2$ doublet, the broken traces are for the $m_S = \pm 3/2$ doublet. The rhombicity distribution is for $\eta = 0.058$ with $\Delta = 15.0\%$ and applies to CO dehydrogenase. Note the turning point in one of the g-value traces for the second doublet, which is sampled by the distribution.

Wide zero field interaction distributions in the high-spin EPR of metalloproteins

WILFRED R HAGEN

Department of Biotechnology, Delft University of Technology, Julianalaan 67,

2628 BC Delft, The Netherlands

Correspndence:

Prof dr Wilfred R Hagen

Department of Biotechnology

Delft University of Technology

Julianalaan 67

2628 BC Delft

Tel: +31 15 2785051

Fax: +31 15 2782355

e-mail: w.r.hagen@tudelft.nl

Abstract

The EPR spin Hamiltonian parameters of transition ion active centres in frozen aqueous solutions of proteins are distributed as a reflection of distributions in spacial conformations. This phenomenon is generally referred to with the generic term ‘g-strain’, however, its manifestation is not limited to a distribution in the g-values. The equivalent name ‘D-strain’ applies to the situation common for biological half-integer high-spin systems whose powder EPR shape is predominantly modified through the distributive effects in the second order terms of the spin-spin interaction Hamiltonian. A simple, one-parameter model is developed to account for D-strain, and this forms the basis for an efficient and practical numerical analysis procedure for D-strained spectra in the weak field limit. Analysis of S = 5/2, 7/2, and 9/2 protein examples is used to illustrate the drastic modification of relative intensities and widths and the occurrence of extra turning points in these spectra as a consequence of D-strain.

1. Introduction

Biomacromolecules have a very high dimensional conformational space, which exhibits rather shallow absolute and relative minima reflecting the high structural flexibility that is presumably required for proper action in key biological events, e.g., catalysis, signal transduction, and regulation of gene expression [1,2]. A frozen-in distribution of conformations is apparently retained in crystallized proteins [3] but also in frozen dilute aqueous solutions [4] which is the common sample form in biomolecular EPR spectroscopy of metalloproteins. In its turn this conformational distribution leads to a distribution in spin-Hamiltonian parameters also known as 'g-strain' [4,5]. Although the ultimate cause of g-strain appears to be describable in terms of a simple, one-dimensional hydrostatic pressure, possibly related to the average size of ice microcrystals in the frozen dilute solution [6], its translated effect on paramagnetic sites through a stress-strain relation via the complex (namely: symmetry lacking) 3D structure of the protein, results in g-strain to be a tensorial quantity not colinear with the g-tensor itself [7]. This implies the existence of two independent interactions (g-strain and the electronic Zeeman interaction) that are linear in the magnetic field, and high quality multi-frequency data are required to separate these from field-independent terms, e.g., hyperfine interactions, for accurate analysis of g-strain [8,9].

It has been realized early on in the development of g-strain analysis that no a priori reason would prevent any other parameter than the g-value in the spin Hamiltonian to also be subject to distribution. However, although some resulting spectral effects have

been identified (e.g., the variation of line width over a set of hyperfine lines [10]), attempts at quantitative simulation of the powder EPR pattern have been rare [11]. This is particularly true for 'D-strain', the name given to a distribution in the axial second order term that frequently dominates the zero-field Hamiltonian [5]. Since the magnitude of |D| in metalloproteins is typically of the order of a wavenumber, a relevant data set for quantitative D-strain analysis should, in addition to X-band (9-10 Hz) spectra, also include spectra taken at significantly higher frequencies and fields. At present, and presumably also in the forseeable future, this poses a technical problem of sensitivity: as a consequence of their high molecular mass (typically 10^5 Da) the maximum concentration of (paramagnetic sites in) metalloproteins is usually limited to values near or below circa 1 mM, and in combination with the very wide field range covered by high-spin spectra when the zero-field interaction is comparable in magnitude to the electronic Zeeman interaction (i.e. at microwave frequencies of the order of 100 GHz) existent high-field/high-frequency EPR spectrometers simply do not have the concentration sensitivity required for meaningful data collection on these systems [12,13]. Indeed, the vast majority of high-field EPR data published on high-spin systems concerns powders of pure (i.e. undiluted) non-biological coordination compounds of limited molecular mass (i.e. typically 1 M or higher in concentration) [14].

Deleted: extant

In principle, the classical concept of the spin Hamiltonian could provide a framework for the development of a numerical tool for the description of distributed high-spin EPR powder patterns, equivalent to the statistical theory of g-stain for distributed S=1/2 systems. We worked out this premise in 1984 [15] in the form of a computationally intensive powder spectral simulator that would (i) employ a spin

Hamiltonian with second and higher order zero field terms, (ii) diagonalize the energy matrix for every discrete value of the magnetic field scan, (iii) search for pairs of energy levels in the spin manifold that would fit the resonance condition, and (iv) calculate the exact transition probability from the new state vectors. The resulting stick spectrum would then (v) be convoluted with an inhomogeneously broadened line in frequency space, and, finally, (vi) a distribution in the zero field Hamiltonian parameters, or 'D strain', would be implemented under a 'reasonable' model. The question of what in fact is a 'reasonable' model for inhomogeneously broadened high-spin EPR was more recently explored in some detail on multi-frequency (specifically, high-frequency) data sets from non-biological model compounds [16-18], however, this work has yet to lead to a general proposal for D strain in metalloproteins. In the intervening decades we have collected extensive data sets of high-spin powder EPR spectra from a broad variety of metalloproteins, and we have been consistently confronted with the observation that, although spectral peak positions would appear to approximately fit those predicted by appropriate spin Hamiltonians, the actual shape of the powder spectrum, in particularly the relative peak intensities and widths, and sometimes the occurrence of extra features, would remain enigmatic (cf, e.g., the data below). Other workers have also developed computationally intensive simulators (cf [19,20] and numerous references quoted therein), however, none seem to have been particularly aimed at dealing with the problem that we address here, and which is summarized as follows: Metalloproteins can typically not be prepared in concentrations above 1 mM and, consequently, high-spin EPR data of sufficient quality for spectral analysis, especially for systems of pronounced rhombicity, can frequently be obtained only in X-band. An accurate and unique determination of their

Deleted: d

zero field splitting parameters is perhaps not a priority goal of biomolecular spectroscopy where multi-frequency data sets are usually not available, and where a meaningful interpretation of the zero field splitting parameters is yet to be developed [21]. It would, however, be quite useful for the biologist to know what spectral features to group in a single spectral component in order to count the total number of detectable spectral components from multi-centre metalloproteins, and to be able to determine, to reasonable approximation, their relative and absolute stoichiometry, because the signals can then be used as quantitative flags of active centres as a function of biologically relevant parameters, e.g., solution redox potential, pH, substrate concentration, etcetera. It is against this background that we have developed the minimal hypothesis of D-strain in the weak-field limit and its derived simple and rapid simulator VisualRHOMBO for powder EPR of distributed high-spin systems described below.

Deleted: 0

2. Materials and methods

The proteins and their EPR spectra analyzed here as examples of wide D-strain distributions, have been described: the S = 5/2 systems of *Desulfovibrio vulgaris* (strain Hildenborough) desulfoferrodoxin and its spectral model Fe(III) DTPA (diethylenetriaminepentaacetate) [22]; the S = 7/2 system of *Azotobacter vinelandii* FeMo-nitrogenase [23]; the S = 9/2 system of *Methanothrix soehngenii* (proposed to be renamed as: *Methanosaeta concilii* [24]) carbonmonoxide dehydrogenase [25,26].

Deleted: 1
Deleted: 2
Deleted: 3
Deleted: 4
Deleted: 5

The code for the wide D-strain simulator has been written in FORTRAN 90/95 with a graphical user interface for the Windows operating system (Windows XP and later). The program was compiled with the Intel Visual FORTRAN 9.1 compiler integrated in the Microsoft Visual Studio 2005 developer environment. The PC program VisualRHOMBO can be downloaded free of charge as a stand-alone application (single-file executable) from the Departmental website: www.bt.tudelft.nl > Research > Download centre.

3. Results and discussion

3.1. *Theory of D-strain and its numerical analysis*

The minimal spin Hamiltonian for biological half-integer high-spin systems is

$$H = \mathrm{S}\cdot\mathbf{D}\cdot\mathrm{S} + \beta\mathrm{B}\cdot\mathbf{g}\cdot\mathrm{S} \qquad (1)$$

in which **D** and **g** are tensorial quantities. At X-band frequencies ($h\upsilon \approx 0.3$ cm^{-1}) the first term, the second-order zero field interaction, is generally dominant (typically by an order of magnitude or more) over the second term, the electronic Zeeman interaction (six-coordinate Mn^{2+} is the exception [27]). Hyperfine interaction is ignored here (but should be included as a perturbation for, e.g., high-spin Co^{2+} proteins). Superhyperfine interactions are usually not resolved in the type of spectra discussed here. Higher-order terms in the zero-field interactions can, and will, occur for $S \geq 2$, however, it is usually not possible in half-integer high-spin powder spectra in the weak-field limit at a single frequency to separate their contribution from the second-order terms (for integer spins they may not be ignored, cf [15,28]). Iron is the most ubiquitous metal in biological high-spin systems, and all the examples worked out below, are from iron proteins. This means that deviations of the principal elements of the g-matrix from the free electron value, $g_e = 2.00$, will be insignificant for Fe^{3+}, and will be relatively small for clusters containing Fe^{3+} (i.e. all examples, below), and so will be the g-anisotropy, and therefore, the unstrained spin Hamiltonian reduces to

Deleted: 6

Deleted: 7

$$H = D[S_z^2 - S(S+1)/3] + E(S_x^2 - S_y^2) + g_{iso}\beta BS \qquad (2)$$

In which the zero field interaction term has been written in its two-parameter form implied by the traceless nature of the D-tensor, and thus all coefficients, D, E, g, are now scalars. The spin manifold in zero field factors into Kramers doublets (i.e. degenerate level pairs).

In the weak-field limit ($D >> 0.3$ cm^{-1}) only intra-doublet transitions are possible in such systems, and it has been realized more than four decades ago, that the resulting (S+1/2) spectra are highly *in*sensitive to the absolute value of D, but very sensitive to the

ratio E/D, also known as the rhombicity η [29]. For high-spin Fe^{3+} ($g \approx g_e$) this means that the complete S = 5/2 sum spectrum is determined by a single parameter, which, however, does not necessarily imply that the spectra are simple in appearance: each intra-doublet transition is describable by an *effective* spin 1/2 with *three* effective principal g-values. In principle, this affords for S = 5/2 nine, or in general 3(S+1/2) different detectable features (turning points or effective g-values) in the overall powder spectrum. On the other hand, the analysis of these spectra is straightforward with the help of a vertical ruler moved over plots of effective g-values versus rhombicity (also known as rhombograms, cf [30] and in Figure 1).

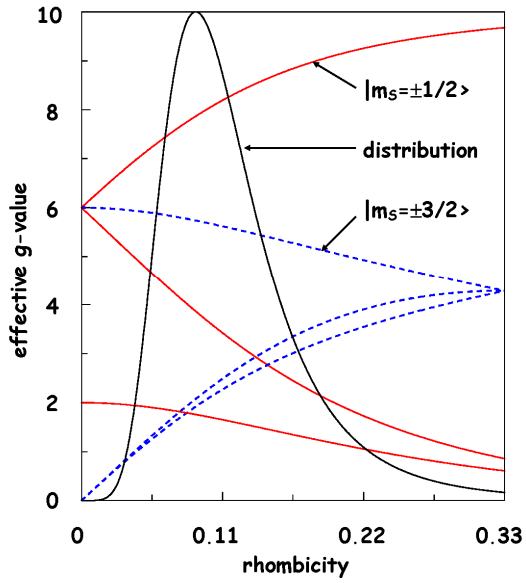
Deleted: 8

Deleted: one

Deleted: only

Deleted: 29

1



The fact that each individual intra-doublet transition is describable as an effective spin one-half system, i.e. that for each transition the effective Hamiltonian

$$H = \beta B \cdot g_{eff} \cdot S_{eff} \quad (3)$$

applies, means a very considerable simplification of the numerical spectral analysis because the Hamiltonian in eq (2) has to be diagonalized only three times, namely along each of the prinicipal (molecular) axes, and the resonance condition $h\upsilon = g_{eff}\beta B$ for all intermediate orientations, and therefore the complete powder pattern, can be solved analytically as

$$g_{eff} = (\sum l_i^2 g_{i\,eff}^2)^{0.5} \quad (4)$$

with transition probability

$$I = g^{-1}_{eff} \sum [g_{i\,eff}^2 - (l_i^2 g_{i\,eff}^4 / g^2_{eff})] \quad (5)$$

in which the summation is over the molecular axes x, y, z, and l_i are the direction cosines between the magnetic field vector B and the molecular axes (cf [28,30]).

With the implementation of a few earlier developed numerical 'tricks', notably, tabulation of the line-shape function in a large array to avoid massively repetitive calculations of exponentials and of floating point interpolations [31], smart coding (do-loop unrolling), and the use of an efficient interpolation scheme to scan the unit sphere (the 'Igloo' method due to R.L. Belford's group, cf paragraph 5.3.2 in [32]), a simulator for *un*distributed half-integer spin spectra can be set up that typically generates a complete powder pattern for up to S = 9/2 in less than one second on an off-the-shelf PC. Real-time fitting of D-strain spectra now becomes practical: a distribution in rhombicity is set up as follows. Both zero field interaction parameters, D and E, are subject to a normal distribution with standard deviation Δ_D and Δ_E expressed as percentage of the

Deleted: 7
Deleted: 29
Deleted: 0
Deleted: 1

average values of D and E. A full positive correlation between these Δ's would lead to no broadening at all in the weak-field EPR spectra as they depend only on the rhombicity, the ratio E/D. On the other hand, the most pronounced broadening would occur in case of full negative correlation, $\Delta_D = -\Delta_E$, and this definition reduces the fitting parameters in the strain model to one. Below, it is shown that this model affords semi-quantitative fits to the powder EPR spectra of several high-spin metalloproteins. Note also that in previous extensive multi-frequency analyses of g-strain in metalloprotein EPR full negative correlation of strain parameters was found to be the predominant, if not the only case of relevance [6,7,9,31].

In practice, the average of D is set at a high dummy value, $D >> 0.3$ cm^{-1}, for example $|D| \equiv 10$ cm^{-1}, the average of E is defined through the rhombicity, $E = \eta D$, and the sign of E is taken to be opposite to that of D. The fully negatively correlated normal distributions in D and E now lead to an asymmetric distribution on the rhombicity $|\eta|$, defined by a single parameter Δ (= $|\Delta_D| = |\Delta_E|$). An example is given in figure 1 for η = 0.09 and Δ = 15.6%, and with the individual actual values of D and E undetermined.

To obtain the strained EPR powder spectrum, for each set of (D,E) values eq (2) is diagonalized, and eqs (4) and (5) are solved. All resulting spectra are then summed with proper weighing according to the intensity of the rhombicity distribution to obtain the D-strained spectrum, which on a standard PC takes circa one minute or less for up to S = 9/2. Since in many practical cases the rhombicity is readily read out, or at least approximately so, from experimental peak positions, the complete D-strain simulation is de facto dependent only on a single parameter, Δ. In this approach the line width of the individual line shape turns into a 'semi-parameter': its value becomes increasingly

Deleted: 0

Deleted: m

irrelevant (hence the label: 'dummy' width in table 1, below) with increasing D-strain, except perhaps for turning point features at low field (high effective g-value) which prove to be relatively insensitive to D-strain (see below).

In practice then, the procedure is as follows: (i) an approximate real g-value is estimated from EPR theory (e.g., $g \approx 2.00$ for hs Fe^{3+}, $S = 5/2$; $g \approx 2.0 \pm 0.1$ for hs Fe/S clusters, $S \leq 9/2$; $g \approx 2.2 \pm 0.1$ for hs Co^{2+}, $S = 3/2$, etc.); (ii) a fairly accurate estimate of the average rhombicity, $\eta = E/D$, is made from the peak positions in the experimental spectrum (in particular the low-field ones) using rhombograms; (iii) the individual inhomogeneous line width, γ, is set at a dummy value significantly less than the apparent line width, Γ, of the lowest field feature of the experimental D-strained spectrum (see below); (iv) the distribution width Δ is determined interactively by repeated powder spectrum simulation with the VisualRHOMBO program; (v) the fit may be fine tuned with minor adjustments to g_{real}, η, γ, and subsequently Δ, and perhaps with minor adjuistments (increases) of the number of sampled molecular orientations (typically 10^4-10^5) and the number of sampled values in the rhombicity distribution (typically 10^2) to eliminate mosaic artifacts. The whole interactive fitting procedure should converge in well less than an hour; otherwise a mistake is indicated in the assignment of the system spin.

Note that there is an ambiguity in the 'units' in which Δ can be expressed. Since in the weak-field limit the actual value of $|D|$ is undetermined, and $\eta = E/D$ is dimensionless, Δ can be conveniently expressed as a percentage of (D,E). However, when the system is of (near) axial symmetry, i.e. $E \approx 0$, and thus $\eta \approx 0$, then Δ becomes the (very large) percentage deviation of a very small number. For these systems Δ may

perhaps preferably be expressed in energy units (e.g., reciprocal centimeters), i.e. in absolute standard deviations of D and E. The range of possible physically meaningful values for the rhombicity is theoretically limited to $0 \le \eta \le 1/3$ [29]. In all the examples worked out below the average rhombicity is $\eta > 0.05$, and Δ is reported as percentage of (D,E).

The key relevance of the above D-strain analysis is to provide insight in the shape of powder EPR spectra from metalloproteins and their models, in particular as to which spectral features combined make up the spectrum of a single high-spin system, and what are their relative amplitudes. In its turn this information affords assignment of spectra to active centers, to quantitate their concentrations (i.e. spin counting) and to monitor amplitudes of assigned peaks as function of biochemically relevant external parameters, e.g., redox potentials.

3.2. *An S = 5/2 example: multiplicity of D-strained spectra*

Sulfate reducing micororganism contain a small (2x14 kDa homodimer) protein called desulfoferrodoxin [33], which functions as a superoxide reductase [34]. The enzyme carries two iron ions per subunit: center I with coordination by four cysteine ligands, $Fe(S^{\gamma}_{Cys})_4$, and center II in which the Fe is coordinated by one cysteine and four histidines, $FeS^{\gamma}_{Cys}N^{\delta}_{His}(N^{\varepsilon}_{His})_3$ [35]. The iron of center II, the site of superoxide activation, is in the ferrous Fe(II) state before reaction with the substrate, but becomes

Deleted: 8
Deleted: 2
Deleted: 3
Deleted: 4

Fe(III)-hydroxide at high pH [36]. The iron is also slowly oxidized by air during aerobic protein purification.

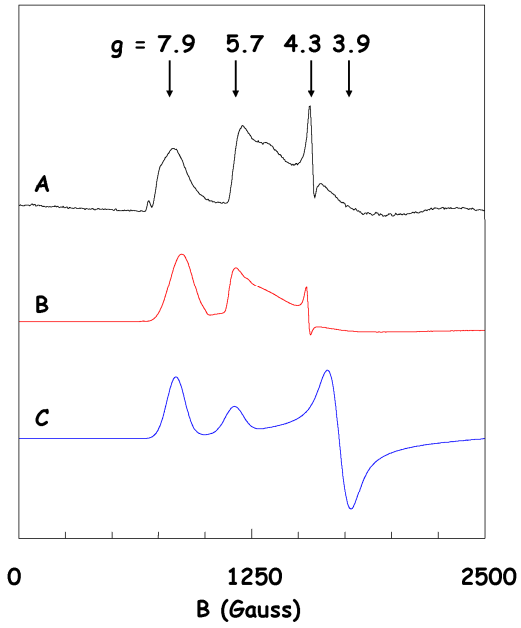
For the ferric Center II the low-field peak positions have been assigned to an S = 5/2 system with rhombicity η = E/D ≈ 0.08 [33]. The spectrum was later found to be remarkably similar to that of Fe(III) DTPA (diethylenetriaminepentaacetic acid) [22], which implies η ≈ 0.08 also for this spectroscopic model, however, confirmation of these assignments by powder spectral simulation was not attempted. In point of fact, as can be seen in figures 2 and 3, traces C,

Deleted: 5

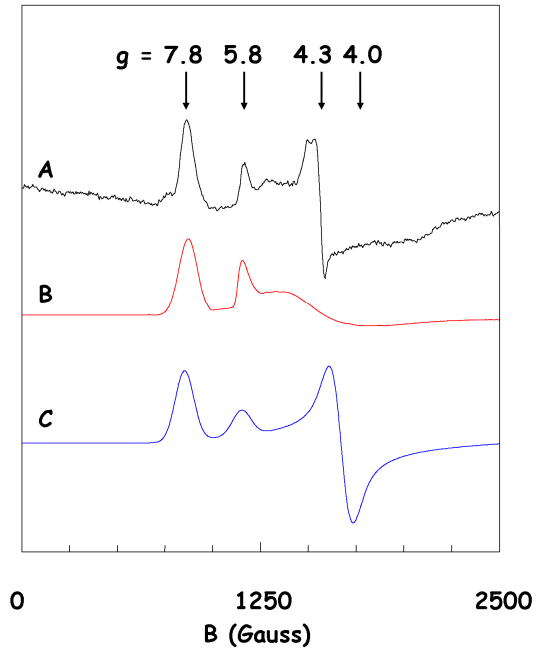
Deleted: 2

Deleted: 1

2



3



a simulation based on this rhombicity value may approximately reproduce the field positions of the first two low-field lines, but the relative width and, for FeDTPA, the intensity of the second line is poorly reproduced. Furthermore, a strong, derivative shaped feature at higher field (g_{eff} = 4.1), generated in the simulations, is absent in the experimental spectrum of the protein and the model.

The introduction of a Gaussian distribution in (D,E) around their central moments, corresponding to the rhombicity value originally obtained from rhombogram

analysis of low-field peaks, drastically changes the overall simulated spectral shape and affords a semi-quantitative reproduction of the experimental spectra (figures 2 & 3, traces B; table 1). In particular, a sharp feature at g_{eff} = 4.3 in the FeDTPA spectrum is reproduced attesting to the fact that this line is not from a contaminant of high rhombicity, but rather reflects the wide nature of the distribution that essentially samples *all* possible rhombicities (cf. figure 1). The occurrence of a g = 4.3 line as an intrinsic part of the distributed spectra of high-spin S = 5/2 model compounds has been recognized before [37], however, notice that the line is *absent* in the simulated spectrum of the desulfoferrodoxin protein, although the experimental spectrum exhibits a relatively strong feature at this effective g-value. This is consistent with our previous observation that the second iron, center I, is partially oxidized in the as isolated protein and has a 'rhombic' EPR spectrum centered around g = 4.3 [22]. Apparently, D-strain simulation is a necessity to be able to discriminate between a contaminating and an intrinsic g = 4.3 line.

Deleted: 6

Deleted: 1

3.3. *An S = 7/2 example: quantification of D-strained spectra*

Two decades ago the P-cluster prosthetic group in nitrogenase was audaciously proposed to be a 'supercluster' containing eight iron ions. The proposal was based on temperature-dependent, quantitative EPR spectroscopy of thionine-oxidized enzyme, in which the cluster exhibits S = 7/2 paramagnetism with D = -3.7 cm^{-1} [23]. The assignment of S = 7/2 was based on a rhombogram fit to the peak positions at low field (g_{eff} = 10.4, 5.8, 5.5), and spin counting was done by singly integrating the lowest-field absorption shaped

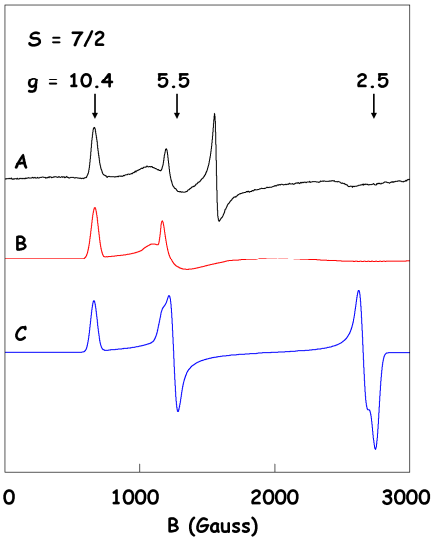
Deleted: 2

peak at g = 10.4 using the single-peak integration method propsed by Aasa and Vänngård [38] with reference to the EPR spectrum of an external standard of known concentration. This afforded approximately two P-clusters per nitrogenase, and with circa 16 Fe availabe based on analytical chemistry (in addition to the iron in the FeMo-cofactor), the first identification of a biological 8Fe cluster was straightforward.

Deleted: 7

However, the conclusion would perhaps have been less straightforward if one would have tried to actually reproduce the powder spectrum by simulation as shown in figure 4, trace C:

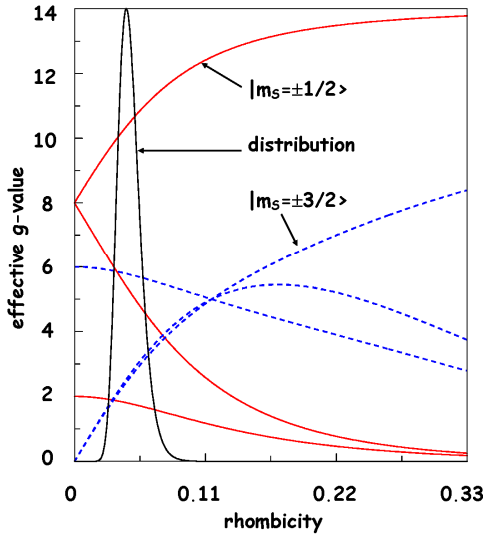
4



low field peak positions do approximately fit, but the shape around g = 5.5 is poorly reproduced. Furthermore, the simulation shows a very pronounced feature at g = 2.5 where the experimental spectrum in trace A only exhibits perhaps a minor declination. Introduction of a widely distributed rhombicity in the simulation (trace B) comes to the

rescue of the spin assignment, where the spectrum is well reproduced on basis of an S = 7/2 with rhombicity $\eta = 0.044$ and a standard deviation of $\Delta(D,E) = 8\%$ (cf figure 5).

5



Furthermore, comparison of the distributed (trace B) and undistributed (trace C) simulation shows that the spectral effects of D-strain are more pronounced towards higher fields (lower effective g-values), and that the lowest-field peak at g = 10.4 is the only feature whose shape remains essentially unaffected, which implies that the original quantification incidentally has been carried out properly, and that the 'supercluster' proposal is corroborated. This argument has been put on a quantitative footing in the last column of Table 1, which gives the intensity ratio of the simulated lowest field peak in the presence and the absence of strain: a ratio of unity implies the Aasa-Vänngård single-

peak integration method to be applicable also to the strained spectrum. Generalizing; spin counting using D-strained spectra should be done only on the first integral of the lowest-field peak (or, alternatively, via full simulation of the D-strained powder spectrum).

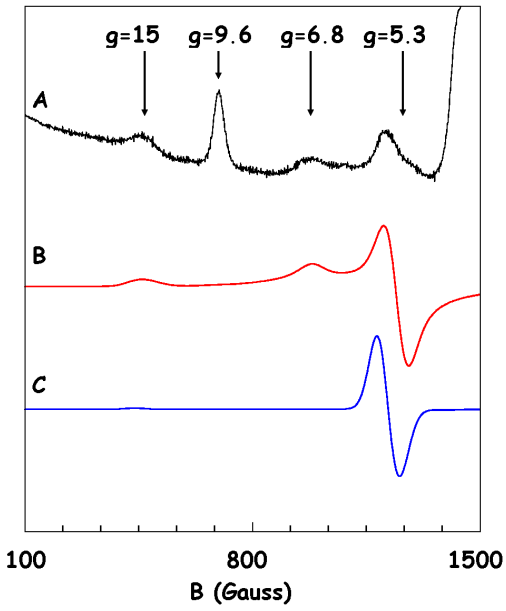
3.4. *An S = 9/2 example: intensities and extra turning points in D-strained spectra*

The unusually high spin S = 9/2 has been identified repeatedly in proteins [25,39,40], however, the basis of the assignment has not always been fully convincing in these experimentally demanding systems of high structural complexity. Not only do the S = 9/2 spectra typically exhibit several features of relatively low intensity over wide field ranges, they are also partially blocked out by interference of overlapping spectra from other prosthetic groups or contaminants in the same enzyme with S < 9/2 covering a less wide field range and, therefore, usually exhibiting a higher amplitude than the S = 9/2 spectra.

A characteristic example is the complex $\alpha_2\beta_2$ heterotetrameric enzyme CO dehydrogenase (acetyl-coenzymeA cleaving) of *M. soehngenii* [28,29]. The EPR spectrum has several low-field peaks of relatively low intensity compared to a rhombic S = 5/2 contaminant (g = 9.6, 4.3; $\eta \approx 1/3$, cf figure 1) that partially blocks the S = 9/2 spectrum and deforms a derivative feature at g ≈ 5.3 (figure 6).

Deleted: 4
Deleted: 8
Deleted: 39
Deleted: 7
Deleted: 8

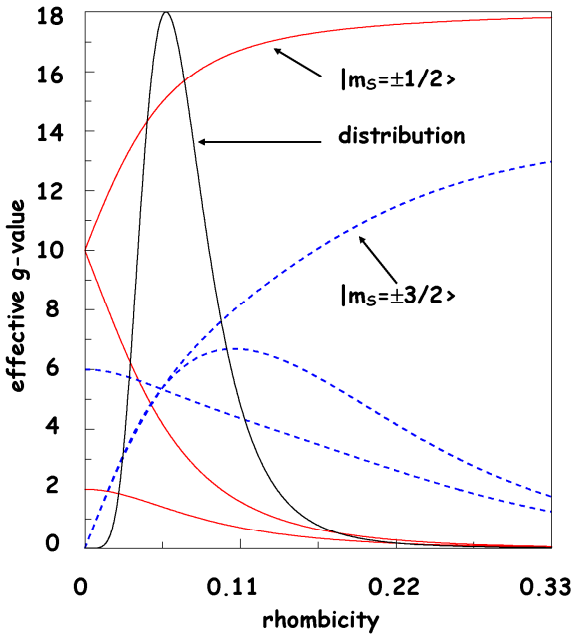
6



The S = 9/2 assignment is indicated by rhombogram analysis predicting g_{eff} = 15 and 5.3 for the lowest-field peak of the |±1/2> and the |±3/2> intradoublet transition, respectively, for η = 0.047 ([28] and figure 7).

Deleted: 7

7



Remains a peak at g_{eff} = 6.8 which was taken at the time to indicate some form of inhomogeneity and was left unassigned [25]. The undistributed and the distributed S = 9/2 simulations based on η = 0.058 and Δ = 15% are presented in traces C and B of figure 6. Introduction of D-strain does not only correct an improper signal intensity ratio of the g = 15 and 5.3 peaks in the undistributed simulation, but it also generates a peak at g = 6.8, eliminating the need to consider spectral inhomogeneity. The physical nature of the g = 6.8 feature in the powder spectrum can be traced back in figure 7 to originate in a turning point at intermediate rhombicity (η ≈ 0.11) of one of the rhombogram traces for S

Deleted: 4

= 9/2, $m_S = \pm 3/2$. Such a turning point implies the effective g-value to be invariant in small changes in the rhombicity, and this will lead to a relatively high EPR amplitude in a broad sampling of rhombicities. In fact, the g = 6.8 line is an S = 9/2 equivalent of the g = 4.3 line that frequently occurs in distributed S = 5/2 systems of intermediate rhombicity (e.g., in figure 2).

3.5. *Biological relevance of wide strain distributions*

A biological interpretation of wide D-strain, if any, is yet to be established. The concept of strain, notably in strained protein conformation, is not unfamiliar to the biochemist, and can perhaps be traced back to the original proposal of Vallee and Williams of an 'entatic' (i.e. strained) state of active centres, based on the observation of unusual optical and EPR properties, to explain, e.g., the apparent minimal reorganization energy of the Cu(I/II) redox transition in blue copper proteins [41,42]. It has later been questioned, on the basis of x-ray data and theoretical calculations, whether blue copper centres are indeed entatic at all [43]. From the perspective of the present work one can wonder whether the width of a distribution in strain (the width of a distributed state whose average conformation may or may not be entatic) would not be of equal, or even greater biological importance than its average magnitude. In terms of biological functionality this would translate into the question whether the nature and width of the distribution is such that it can sample, with significant frequency, geometrical conformations close to the transition state. Such a hypothesis should in principle be testable by relating, e.g., EPR-

Deleted: 0
Deleted: 1
Deleted: 2

detected D-strain to properties, such as reduction potential values, hitherto related to the existence of a single entatic state.

4. Conclusions

D-strain analysis of half-integer high-spin powder EPR data from biological systems and their models, based on the minimal hypothesis of a distribution in rhombicity, derived from a joint Gaussian distribution in D and E, defined in the weak-field limit, affords numerically rapid, semi-quantitative reproduction of experimental spectra. The determination of spin concentration, or spin counting, should be based on the first integral of the lowest-field peak or on full simulation of the D-strained spectrum. D-strain analysis is a mandatory requirement for an understanding of relative intensities and widths of peaks in the powder spectra as well as the appearance of extra features due to turning points in the rhombograms. Such an understanding is required for a biologically meaningful interpretation of data in terms of stoichiometry and multiplicity of protein prosthetic groups. Finally, strain in metalloprotein EPR reflects a conformational distribution that is possibly of relevance for biological functionality namely as a route to sample transition state geometry.

References

[1] R.H. Austin, K.W. Beeston, L. Eisenstein, H. Frauenfelder, I.C. Gunsalus. *Biochemistry*, **14**, 5355 (1975).

[2] H. Frauenfelder. *Proc. Natl. Acad. Sci. USA*, **99**, 2479 (2002).

[3] H. Frauenfelder, G.A. Petsko, D. Tsernoglou. *Nature*, **280**, 558 (1979).

[4] J. Fritz, R. Anderson, J. Fee, G. Palmer, R.H. Sands, J.C.M. Tsibris, I.C. Gunsalus, W.H. Orme-Johnson, H. Beinert. *Biochim. Biophys. Acta*, **253**, 110 (1971).

[5] W.R. Hagen. *J. Magn. Reson.*, **44**, 447 (1981).

[6] W.R.Hagen. In *Advanced EPR; Applications in Biology and Biochemistry*, A.J. Hoff (Ed), p. 785, Elsevier, Amsterdam (1989).

[7] W.R. Hagen, D.O. Hearshen, R.H. Sands, W.R. Dunham. *J. Magn. Reson.*, **61**, 220 (1985).

[8] W.R. Hagen, S.P.J. Albracht. *Biochim. Biophys. Acta*, **702**, 61 (1982).

[9] D.O. Hearshen, W.R. Hagen, R.H. Sands, H.J. Grande, H.L. Crespi, I.C. Gunsalus, W.R. Dunham. *J. Magn. Reson.*, **69**, 440 (1986).

[10] W. Froncisz, J.S. Hyde. *J. Chem. Phys.*, **73**, 3123 (1980).

[11] W.R. Hagen. *EPR spectroscopy of metalloproteins with emphasis on respiratory chain complexes*. PhD thesis, University of Amsterdam. (1982).

[12] W.R. Hagen. *Coord. Chem. Rev.*, **190-192**, 209 (1999).

[13] W.R. Hagen. In *High Magnetic Fields; Science and Technology Vol. 3*, F. Herlach and N. Miura (Eds), p. 207, World Scientific, Signapore (2006).

[14] J. Krzystek, A. Ozarowski, J. Telser. *Coord. Chem. Rev.*, **250**, 2308 (2006).

[15] W.R. Hagen, W.R. Dunham, R.H. Sands, R.W. Shaw, H. Beinert. *Biochim. Biophys. Acta*, **765**, 399 (1984).

[16] P.J. van Dam, A.A.K. Klaassen, E.J. Reijerse, W.R. Hagen. *J. Magn. Reson.*, **130**, 140 (1998).

[17] A. Priem, P.J.M. van Bentum, W.R. Hagen, E.J. Reijerse. *Appl. Magn. Reson.*, **21**, 535 (2001).

[18] A. Priem. *Explorations in high-frequency EPR*, PhD thesis, University Nijmegen, http://webdoc.ubn.kun.nl/mono/p/priem_a/explinhif.pdf (2002).

[19] G.R. Hanson, K.E. Gates, C.J. Noble, M. Griffin, A. Mitchell, S. Benson. *J. Inorg. Biochem.*, **98**, 903 (2004).

[20] S. Stoll, A. Schweiger, *J. Magn. Reson.*, **178**, 42 (2005).

[21] F. Neese. *J. Am. Chem. Soc.*, **128**, 10213 (2006).

[22] M.F.J.M. Verhagen, W.G.B. Voorhorst, J.A. Kolkman, R.B.G. Wolbert, W.R. Hagen. *FEBS Lett.*, **336**, 13 (1993).

[23] W.R. Hagen, H. Wassink, R.R. Eady, B.E. Smith, H. Haaker. *Eur. J. Biochem.*, **169**, 457 (1987).

[24] P. de Vos, H.G. Trüper, B.J. Tindall. *Int. J. Syst. Evol. Microbiol.*, **55**, 525.

[25] M.S.M. Jetten, A.J. Pierik, W.R. Hagen. *Eur. J. Biochem.*, **202**, 1291 (1991).

[26] R.I.L. Eggen, R. van Kranenburg, , A.J.M. Vriesema, A.C.M. Geerling, M.F.J.M. Verhagen, W.R. Hagen, W.M. de Vos. *J. Biol. Chem.*, **271**, 14256 (1996).

[27] J.S. Shaffer, H.A. Farach, C.P. Poole, Jr. *Phys. Rev. B*, **13**, 1869 (1976).

[28] W.R. Hagen. *Dalton Trans.*, **2006**, 4415 (2006).

Deleted: r

Deleted: 0

Deleted: 1

Deleted: 2

Deleted: 3

Deleted: 4

Deleted: 5

Deleted: 6

Deleted: 7

[29] G.J. Troup, D.R. Hutton. *Brit. J. Appl. Phys.*, **15**, 1493 (1964).

[30] W.R. Hagen. *Adv. Inorg. Chem.*, **38**, 165 (1992).

[31] W.R. Hagen, D.O. Hearshen, L.J. Harding, W.R. Dunham. *J. Magn. Reson.*, **61**, 233 (1985).

[32] J.R. Pilbrow. *Transition Ion Electron Paramagnetic Resonance*. Clarendon Press, Oxford (1990).

[33] I. Moura, P. Tavares, J.J.G. Moura. *J. Biol. Chem.*, **265**, 21596 (1990)

[34] C.V. Romão, M.Y. Liu, J. Le Gall, C.M. Gomes, V. Braga, I. Pacheco, A.V. Xavier, M. Teixeira. *Eur. J. Biochem.*, **261**, 438 (1999).

[35] A.V. Coelho, P.M. Matias, V. Fülöp, A. Thompson, A. Gonzalez, M.A. Carrondo, *J. Biol. Inorg. Chem.*, **2**, 680 (1997).

[36] C. Mathe, V. Niviere, T.A. Mattioli. *J. Am. Chem. Soc.*, **127**, 16436 (2005).

[37] J.T. Weisser, M.J. Nilges, M.J. Sever, J.J. Wilker. *Inorg. Chem.*, **45**, 7736 (2006).

[38] R. Aasa, T. Vänngård. *J. Magn. Reson.*, **19**, 308 (1975).

[39] A.J. Pierik, W.R. Hagen. *Eur. J. Biochem.*, **195**, 505 (1991).

[40] A.J. Pierik, W.R. Hagen, W.R. Dunham, R.H. Sands. *Eur. J. Biochem.*, **206**, 705 (1992).

[41] B.L. Vallee, R.J.P. Williams. *Proc. Natl. Acad. Sci. USA*, **59**, 498 (1968).

[42] R.J.P. Williams. *Eur. J. Biochem.*, **234**, 363 (1995).

[43] U. Ryde, M.H.M. Olsson, K. Pierloot, B.O. Roos. *J. Mol. Biol.*, **261**, 586 (1996).

Deleted: 8
Deleted: 29
Deleted: 0
Deleted: 1
Deleted: 2
Deleted: 3
Deleted: 4
Deleted: 5
Deleted: 6
Deleted: 7
Deleted: 8
Formatted: Dutch (Netherlands)
Deleted: 39
Deleted: 0
Deleted: 1
Deleted: 2

Table 1: Simulation parameters of wide D-strain distributions in the powder EPR of high-spin metalloproteins

Species	rhombicity η	strain width Δ (%)	dummy width[a] γ (Gauss)	line width[b] Γ (Gauss)	intensity ratio[c]
S = 5/2					
Desulfoferrodoxin	0.085	12.0	5	40	1.04
Fe-DTPA	0.090	15.6	5	40	0.84
S = 7/2					
nitrogenase P-cluster	0.044	8.0	5	20	0.96
S = 9/2					
CO dehydrogenase	0.058	15.0	25	25	0.79

Note: a) line width of distributed spectrum (the label: 'dummy' is explained in the main text); b) line width of undistributed spectrum; c) ratio of intensities of the lowest field peak in distributed over undistributed spectra.

Deleted: 1

Legends to the figures

Figure 1. Rhombogram for S = 5/2: effective g-values as a function of rhombicity, η = E/D. The solid traces are the three effective g-values for the $m_S = \pm 1/2$ doublet; the broken traces are the g-values for the $m_S = \pm 3/2$ doublet. In this and following rhombograms only g-values for the lowest two doublets are plotted, because simulations show that the intra-doublet transitions in higher doublets contribute negligible intensity to the powder EPR spectrum. The trace labeled 'distribution' is the distribution in rhombicity resulting from a joint normal distribution in D and E about the central value of η = 0.090 with Δ = 15.6%. These values apply to Fe(III)DTPA. The amplitude of the distribution is arbitrarily set to a value of 10 to fit in the rhombogram.

Figure 2. Experimental and simulated spectrum of Fe(III)DTPA. Trace A, experimental spectrum at 9.30 GHz [22]; trace B, distributed simulation; trace C, undistributed simulation (parameters in table 1).

Deleted: 1

Figure 3. Experimental and simulated spectrum of *D. vulgaris* desulfoferrodoxin. Trace A, experimental spectrum at 9.32 GHz [22]; trace B, distributed simulation; trace C, undistributed simulation (parameters in table 1).

Deleted: 1

Figure 4. Experimental and simulated spectrum of the P-cluster in thionine-oxidized *A. vinelandii* FeMo-nitrogenase. Trace A, experimental spectrum at 9.30 GHz [23]; trace B, distributed simulation; trace C, undistributed simulation (parameters in table 1).

Deleted: 2

Figure 5. Rhombogram for S = 7/2. The solid traces are for the $m_S = \pm 1/2$ doublet, the broken traces are for the $m_S = \pm 3/2$ doublet. The rhombicity distribution is for $\eta = 0.44$ with $\Delta = 8.0\%$ and applies to the P-cluster.

Figure 6. Experimental and simulated spectrum of *M. soehngenii* CO dehydrogenase. Trace A, experimental spectrum at 9.30 GHz [28]; trace B, distributed simulation; trace C, undistributed simulation (parameters in table 1).

Deleted: 7

Figure 7. Rhombogram for S = 9/2. The solid traces are for the $m_S = \pm 1/2$ doublet, the broken traces are for the $m_S = \pm 3/2$ doublet. The rhombicity distribution is for $\eta = 0.058$ with $\Delta = 15.0\%$ and applies to CO dehydrogenase. Note the turning point in one of the g-value traces for the second doublet, which is sampled by the distribution.