In Quest of the Alanine R3 Radical: Multivariate EPR Spectral Analyses of X-Irradiated Alanine in the Solid State

Eirik O. Jåstad[†], Turid Torheim[†#], Kathleen M. Villeneuve[‡,‖], Knut Kvaal[†], Eli O. Hole[‡], Einar Sagstuen[‡,*], Eirik Malinen[‡,§] and Cecilia M. Futsaether[†,*]

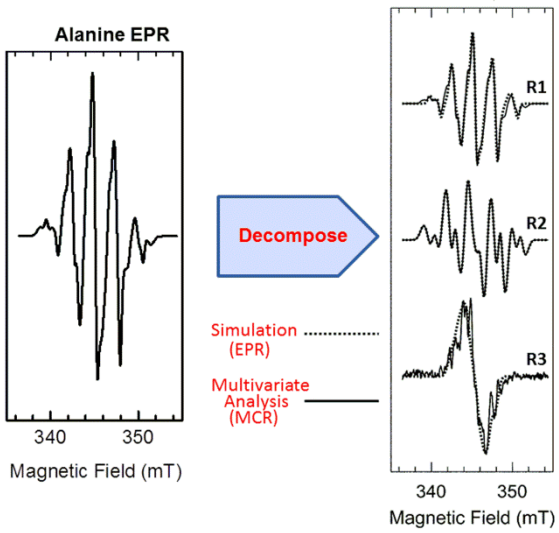
[†]Faculty of Science and Technology, Norwegian University of Life Sciences, Ås, Norway; [‡]Department of Physics, University of Oslo, Oslo, Norway; [§]Department of Medical Physics, Oslo University Hospital, Oslo, Norway

* Corresponding authors, e-mail: cecilia.futsaether@nmbu.no, tel. +47 672 31563
e-mail: einar.sagstuen@fys.uio.no, tel. +47 228 55653

\# Present address: Cancer Research UK Cambridge Research Institute: Cambridge, Cambridgeshire, UK.

‖ Present address: Department of Physics and Astronomy, The University of Calgary, Alberta, Canada

ABSTRACT



The amino acid *l*-α-alanine is the most commonly used material for solid-state Electron Paramagnetic Resonance (EPR) dosimetry, due to the formation of highly stable radicals upon irradiation, with yields proportional to the radiation dose. Two major alanine radical components designated **R1** and **R2** have previously been uniquely characterized from EPR and electron-nuclear double resonance (ENDOR) studies as well as from quantum chemical calculations. There is also convincing experimental evidence of a third minor radical component **R3** and a tentative radical structure has been suggested, even though no well-defined spectral signature has been observed experimentally. In the present study, temperature dependent EPR spectra of x-ray irradiated polycrystalline alanine were analyzed using five multivariate methods in further attempts to understand the composite nature of the alanine dosimeter EPR spectrum. Principal Component Analysis (PCA), Maximum Likelihood Common Factor Analysis (MLCFA), Independent Component Analysis (ICA), Self-modelling Mixture Analysis (SMA) and Multivariate Curve Resolution (MCR) were used to extract pure radical spectra and their fractional contributions from the experimental EPR spectra. All methods yielded spectral estimates resembling the established **R1** spectrum. Furthermore, SMA and MCR consistently predicted both the established **R2** spectrum and the shape of the **R3** spectrum. The predicted shape of the **R3** spectrum corresponded well with the proposed tentative

spectrum derived from spectrum simulations. Thus, results from two independent multivariate data analysis techniques strongly support the previous evidence that three radicals are indeed present in irradiated alanine samples.

1. INTRODUCTION

Radiation dosimetry by Electron Paramagnetic Resonance (EPR) spectroscopy has during the last 3-4 decades gained acceptance as a precise dosimetry system for a range of physical, biological, medical and industrial applications. In EPR dosimetry, the number of radiation induced radicals is used as a measure of the absorbed radiation dose. It was early realized that radicals in high yields can be formed in organic substances upon radiation exposure, and that in the solid state these species can be very stable at room temperature.[1-5] Furthermore, within a given dose range, the yields of these stable radicals are linearly proportional to the radiation dose. Already in 1962, Bradshaw and coworkers[6] suggested that such compounds together with EPR spectroscopy could be used for radiation dosimetry. It was further suggested that the amino acid *l*-α-alanine could be an excellent candidate for solid state EPR dosimetry. [6]

As a non-destructive quantitative method, the concept of EPR/alanine dosimetry attracted interest – an interest that gained considerable momentum in the mid-1980ies by the careful investigations by Hansen and Olsen and Regulla, Deffner and Wieser.[7-11] The longevity of the induced radicals suggested additional benefits such as cumulative radiation measurement procedures and long-term dosimeter storage for quality control and metastudies. Further corroboration was obtained from a number of subsequent very detailed studies by others using many different radiation qualities and environmental conditions.[12-14] Today, alanine is the most used material for solid-state EPR dosimetry purposes. Extensive international dosimetry services based on EPR/alanine dosimetry are now offered by e.g. NIST (USA), PBL (Germany) and NPL (UK).[15-17]

Based on early EPR studies of x- or γ-irradiated alanine crystals, it was assumed that the EPR spectrum was due to one single radical species, often designated the SAR

("stable alanine radical") or radical **R1**. However, Miyagawa and co-workers employing EPR and Electron Nuclear Double Resonance (ENDOR) analyses of irradiated alanine crystals,[18-19] pointed out evidence for at least one additional radical component in the spectra. Callens and co-workers similarly provided evidence for at least one, possibly two other components using multivariate spectral analysis.[20-22] In a series of papers using EPR, ENDOR, ENDOR Induced EPR (EIE) and spectral simulations, Sagstuen, Hole, Malinen and co-workers uniquely characterized the second major component designated **R2** and in addition launched convincing evidence of a third (minor) radical component, **R3.**[23-27] A tentative radical structure was assigned to this third component although no well-defined spectral signature had been observed experimentally.[25] The experimental EPR/ENDOR data enabled precise simulations of the spectra for both the **R1** and **R2** components. However, as the EPR/ENDOR data for **R3** was less complete, only a tentative simulation of the **R3** spectrum was obtained. [25]

The proposed chemical structures of these three radical species have later been carefully analyzed using quantum chemistry modelling – and in each instance the proposed structures were found to agree well with the available experimental data. These structures were also found to be energetically and mechanistically the most probable products. [28-35]

There is still not sufficient experimental evidence for the actual spectral fingerprint of the **R3** radical. During the last decennium, multivariate analyses of series of inherently related data have shown considerable progress and a variety of methods and algorithms have been established for determining hidden common features or traits in such data series.[21-22, 36-40] Some of these methods can be used to extract pure component spectra and component contribution fractions from spectra of mixtures consisting of several different

components, without requiring any prior knowledge of the component spectra or the sample composition.

The relative amounts of at least two of the radical species (**R1** and **R2**) in alanine are known to vary strongly with temperature.[21, 25] This temperature dependence is illustrated in Figure 1. A corresponding graph, showing the un-normalized spectra, is given in the Supplemental Information (Figure S1.1). Thus, in further attempts to isolate the spectral signature of the **R3** radical, temperature dependent EPR spectra of x-ray irradiated alanine were studied using a suite of multivariate methods.

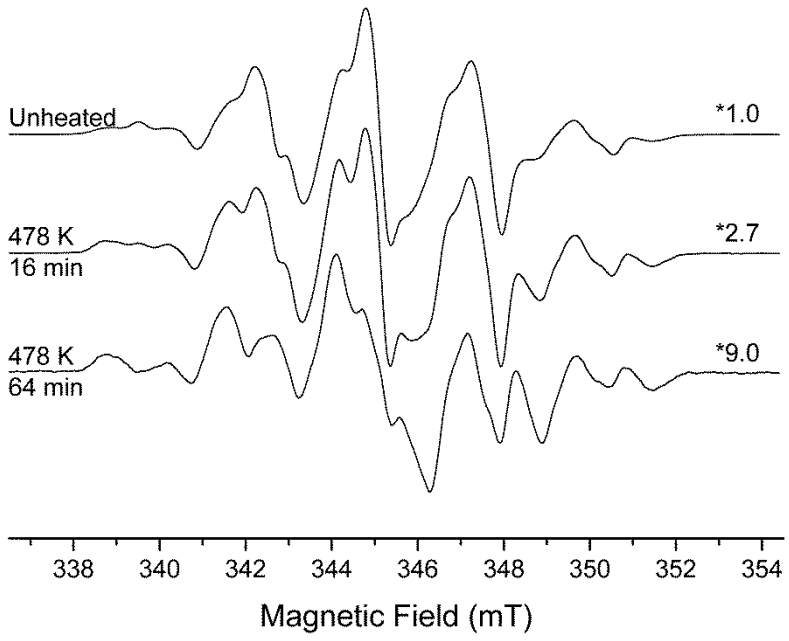


Figure 1: Alanine dosimeter EPR spectra as a function of annealing temperature and time. The top spectrum is recorded from a sample irradiated at 296 K and not exposed to any heat treatment. The middle and lower spectra were recorded at 296 K after heat treatments as indicated. The spectra have been scaled to the spectrum of the untreated sample, and the scaling factors are given with each spectrum.

2. EXPERIMENTAL METHODS

Twenty *l*-α-alanine cylindrical dosimeter pellets (BRUKER BioSpin ES200-2106/ E2044562, batch T020604, weight 64.5 ± 0.5 mg, diameter 4.8 mm, height 3.0 mm), were irradiated at 296 K using 60 kV x-rays from a Philips 100 kV chromium anode tube. A homogeneous dose in the pellets was ensured by inserting an additional 200 μm aluminum filter in front of the anode and flipping the pellets by 180° halfway through the irradiation. The dose rate was 404 ± 7 Gy/min (separate and independent alanine dosimetry). Pellets were irradiated for a total of 13 minutes, resulting in a dose of 5.25 ± 0.09 kGy.

Based on previous work,[24-25] thermal annealing was used to create a sufficient number of independent EPR spectra of alanine samples with variable relative contents of **R1**, **R2** and **R3**. The following annealing temperatures and times were used: 470 K for 1, 30, 60, 90, 120, or 150 mins; 478 K for 1, 16, 32, 48, 64 or 80 mins; 486 K for 1, 10, 20, 30, 40, or 50 mins. One pellet was used at each temperature/annealing time setting, and two pellets were kept at 296 K and used as controls. The oven was preheated at the chosen temperature. After placing the pellets in the oven, the temperature was slightly unstable for about 10 minutes before firmly stabilizing at the set temperature. All heated pellets were cooled to 296 K before EPR measurements were made. *l*-α-alanine radicals are expected to be stable after annealing and cooling to 296 K.[24]

EPR spectra were recorded using a BRUKER Elexsys E560 X-band spectrometer equipped with a standard TE_{102} rectangular cavity. Reproducible pellet positions at the center of the cavity were obtained using a system of teflon and quartz pedestals and a quartz guidance tube for the 5 mm ID quartz tube containing the pellet. The microwave power was set at 2.0 mW incident to the cavity (20 dB attenuation). The magnetic field sweep width was 18 mT at 1024 points resolution, at 0.215 mT/s sweep rate, a 0.33 s time

constant and a 0.2 mT modulation width. The microwave frequency was about 9676.7 MHz and varied only slightly (within 1 MHz).

3. COMPUTATIONAL METHODS

The data set

The twenty EPR sample spectra (18 thermally annealed pellets, two controls) were organized into a data matrix **X**, where the rows and columns corresponded to the samples and 1024 magnetic field variables, respectively, giving a 20 × 1024 data matrix. All spectra were adjusted by baseline correction and normalized to a common microwave frequency. For comparison purposes and to compensate for the strong overall signal decay observed upon heating,[21, 24-25] the EPR spectra were, unless explicitly specified in the text, normalized to unit area. Annealed samples contain relatively larger contributions of the **R3** radical, but have an overall much weaker EPR signal (see Figure S1.1 (Supplemental Information)). Without normalization, these samples would have less influence on the modeling, potentially resulting in poorer estimates of the **R3** spectrum.[24-25]

The **R1**, **R2** and **R3** spectra used for comparisons to the spectral estimates calculated using multivariate techniques are shown in Figure S1.2 (Supplemental Information). These radical spectra were obtained by simulations using experimental hyperfine coupling data.[25] The **R1** and **R2** radical simulated spectra are in the present study assumed to represent the true EPR spectra of these radicals. Note that in the present work, the simulated **R3** component is tentative and is also slightly different from that in previous studies, as described in Section S1.2 (Supplemental Information).

Calculations

Interactive Data Language (IDL 8.2, Exelis Visual Information Solutions, USA) was used for the common microwave frequency normalization, baseline drift and for the least squares fitting-based three-component spectral reconstruction in Section S1.2 (Supplemental Information). Other calculations were performed in MATLAB® (9.0.0, 2016a, The MathWorks, Inc., Natick, MA, USA) in combination with the Statistics and Machine Learning Toolbox (version 10.2), PLS Toolbox® (v8.2.1, 2016, Eigenvector Research Inc., Wenatchee, WA, USA) and the FastICA package.[41]

Initial spectral analysis

The initial spectral analysis was performed using least squares methods to fit a weighted sum of either two (**R1** and **R2)** or three (**R1**, **R2** and **R3)** simulated component spectra to the experimental spectra, to evaluate the goodness of the fit and to give an independent estimate of the **R1**, **R2** and **R3** composition of each spectrum.

Multivariate methods

The chosen methods consider a data matrix **X** consisting of experimental spectra of a mixture of components (radical spectra in this case), and decompose this matrix into what can be interpreted as a component matrix **S** of estimated pure component spectra and a component contribution matrix **C**, given by

$$\mathbf{X} = \mathbf{C}\mathbf{S}^{T} + \mathbf{E} \qquad (1)$$

E is a matrix of residuals not explained by the extracted component spectra **S**.

To assess the decomposition, the estimated contributions **C*** and spectra $\hat{\mathbf{S}}$ were used to reconstruct the experimental spectra, giving the reconstructed data matrix **X*** where

$$\mathbf{X}^* = \mathbf{C}^* \hat{\mathbf{S}}^T \quad (2)$$

Five decomposition methods based on different mathematical conditions or constraints were compared: (1) Principal Component Analysis (PCA);[36, 42] (2) Maximum Likelihood Common Factor Analysis (MLCFA);[21-22, 42-43] (3) Independent Component Analysis (ICA);[37, 41] (4) Self-modelling Mixture Analysis (SMA);[39, 44] (5) Multivariate Curve Resolution (MCR).[40]

Principal Component Analysis (PCA)

In PCA, high-dimensional data is transformed into a lower dimensional space, where the data is projected onto a small number of new, uncorrelated and orthogonal latent variables called principal components (PCs).[42, 45] These PCs capture, in decreasing order, the maximum variance in the data, and are linear combinations of the original variables. Thus, PCA reveals major trends in the data and provides dimensionality reduction, which can be useful in spectral analysis as neighboring variables (e.g. measurements at similar magnetic field values) are often interrelated.

The data matrix **X** is decomposed into an orthogonal set of scores **T** and orthonormal set of loadings **P**. The scores **T** are the coordinates of the samples in the new space and show how the samples are related. The loadings **P** provide the weights of the original variables in the new directions, the PCs, and show how each original variable contributes to the new directions. Equation 1 is rewritten as

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \quad (3)$$

PCA provides a unique solution, but with sign ambiguity for **T** and **P** (i.e. reversing the sign of both **T** and **P** gives the same solution).[42, 45]

In this study, PCA was used to explore the relationship between the sample spectra using score plots, which show the sample scores **T** in the new principal component space. The loadings **P** plotted as a function of the original variables, the magnetic field values, were compared and correlated to the simulated radical spectra to assess their suitability as component spectra. As PCA focuses on maximizing variance, it is standard practice to center the variables (i.e. columns) in the data matrix **X** to mean zero. The rows, i.e. the experimental spectra, in the mean-centered data matrix now show how much each sample spectrum differs from the average spectrum.

Maximum Likelihood Common Factor Analysis (MLCFA)

As in PCA, Factor Analysis (FA) reduces the original variables into fewer, but unobservable, latent variables.[43] Unlike PCA, which searches for principal components capturing the maximum variance in the data, FA factors capture the common data variance.[42, 45]. In FA, the data matrix **X** is decomposed as

$$\mathbf{X} = \mathbf{L}\mathbf{F}^{T} + \mathbf{U} \qquad (4)$$

where **F** is a matrix of common factors, **L** is a matrix of factor loadings or coefficients, and **U** is a matrix of specific factors, referred to as uniqueness.[42-43, 45] The factors **F** are linear combinations of the original variables and account for the shared variance between the variables. The specific factors in **U** capture variance specific or unique to a particular variable.

Factor analysis was performed on normalized experimental spectra using the MATLAB® FA algorithm, which calculates the maximum likelihood estimate of the factor loadings matrix **L** [42] with varimax factor rotation.

Independent Component Analysis (ICA)

In ICA the aim is to decompose, with a minimum loss of information, the data matrix **X** of signals/spectra into underlying components that are statistically independent, and not only uncorrelated as in PCA.[37-38, 41] In addition, the distribution of the components must be non-Gaussian. The ICA model is given by

$$\mathbf{X} = \mathbf{A}\mathbf{S}^{T} + \mathbf{E} \qquad (5)$$

where **A** is the mixing matrix and **S** is the source matrix of independent components. Unlike PCA, where the components are sorted in order of decreasing variance, the ordering and the variance of the independent components cannot be determined. ICA also has sign ambiguity.

Independent components were calculated using the FastICA algorithm,[41] which first mean-centers the data, then transforms the data into uncorrelated components with unit variance using eigenvalue decomposition of the centered data matrix, and finally uses a fixed-point iteration algorithm to maximize the non-Gaussian criterion for computation of the independent components.

Self-modelling Mixture Analysis (SMA)

SMA was conducted using the SIMPLISMA algorithm (SIMPLe-to-use Interactive Self-modelling Mixture Analysis), which searches for so-called pure variables that have contributions from only one component of the system.[39, 44, 46] The intensities of the pure

variables are proportional to the pure component weights in the experimental spectra and are used as component contribution estimates, giving an estimated contribution matrix $\hat{\mathbf{C}}$ for Equation 1. Using least squares, estimates $\hat{\mathbf{S}}$ of the component spectra were obtained from Equation 1 as

$$\hat{\mathbf{S}} = \mathbf{X}^{T}\,\hat{\mathbf{C}}\,(\hat{\mathbf{C}}^{T}\,\hat{\mathbf{C}})^{-1} \qquad (6)$$

New contribution profiles $\mathbf{C}^*$ were then calculated from the extracted spectra $\hat{\mathbf{S}}$ using

$$\mathbf{C}^* = \mathbf{X}\,\hat{\mathbf{S}}\,(\hat{\mathbf{S}}^{T}\,\hat{\mathbf{S}})^{-1} \qquad (7)$$

The pure variables were determined using purity spectra consisting of the purity values of all variables (i.e. magnetic field values) in the experimental spectra. The purity p_{ij} of the j^{th} variable of the i^{th} experimental spectrum is given by

$$p_{ij} = w_{ij}\frac{\sigma_i}{\mu_i+\alpha}, \quad i = 1, \ldots, no.of\ spectra \qquad (8)$$

σ_i and μ_i are the standard deviation and the mean of the i^{th} spectrum. The offset α is a noise correction factor, typically 1-10% of the maximum mean value of the experimental spectra, which prevents spectra with low signal-to-noise ratio, and hence low mean μ_i, from obtaining high purity values. w_{ij} is a determinant based weighting function which gives a measure of the dissimilarity between the i^{th} spectrum and the selected pure spectra.[44] The spectrum with the highest purity (i.e. most dissimilar to the other selected pure spectra) is selected as the next pure spectrum and the maximum of this spectrum is a pure variable. The first pure spectrum was determined by calculating the purity between each spectrum and an initial reference spectrum, selected as a line with slope one.[39] SMA was performed on normalized experimental spectra.

Multivariate Curve Resolution (MCR)

In MCR, the goal is to decompose the data matrix **X** in Equation 1 into chemically meaningful components. In similarity to PCA, MCR attempts to maximize the data variance explained by the components, but unlike PCA, the extracted components are not subjected to the orthogonality constraint.[38, 40, 47] This, however, results in a multitude of different possible estimates of the component spectra $\hat{\mathbf{S}}$ and contribution fractions **C*** that satisfactorily reproduce the original mixture spectra in **X**. Unique solutions are obtained using constraints, which can reflect known physical and chemical system properties, such as non-negative or known contribution fractions or spectra, or mathematical properties such as local rank selectivity constraints, defining, for example, zero-concentration component regions.

In this study, the iterative Multivariate Curve Resolution-Alternating Least Squares algorithm (MCR-ALS) in PLS-toolbox was used, which minimizes **E** in Equation 1 in a least square sense.[48] Contribution (concentration) profiles were assumed to be non-negative. Different equality constraints were tested, where the simulated **R1** and/or **R2** radical spectra (Figure S1.2, Supplemental Information) or the contribution fractions of specific samples were assumed to be known. Based on Heydari *et al.,*[24] the **R1**:**R2**:**R3** radical contribution ratio of the non-heated samples was assumed to be 0.60:0.35:0.05. To account for the inherent uncertainty of the simulated spectra and their relative contributions, equality constraints could be weighted as "hard" or "soft", signifying fixed or approximately set values. For example, a "soft" contribution constraint indicates that the estimated contribution fractions should be similar, but not necessarily equal to the specified values.

Evaluation of the results

Pearson's correlation r ($-1 \leq r \leq 1$), defined as

$$r = \frac{n(\Sigma_i x_i y_i) - (\Sigma_i x_i)(\Sigma_i y_i)}{\sqrt{[n\Sigma_i x_i^2 - (\Sigma_i x_i)^2][n\Sigma_i y_i^2 - (\Sigma_i y_i)^2]}}, \quad datasets\{x_i\} and \{y_i\} each with\ n\ variables$$

was used to assess the similarity between a simulated radical spectrum $\{x_i\}$ and a calculated spectral estimate $\{y_i\}$.[42] For EPR spectroscopy, the correlation coefficient is often considered to be a measure of the degree of 'overlap' between resonance lines from different radicals.

The decomposition models were assessed using the lack of fit, *LOF* [39-40] ($0 \leq LOF \leq 1$) between the original experimental spectra and the reconstructed spectra, defined as

$$LOF = \sqrt{\frac{\sum_{i,j}(x_{ij} - x_{ij}^*)^2}{\sum_{i,j} x_{ij}^2}}, \qquad i = 1, \ldots, no.of\ spectra, j = 1, \ldots, no.of\ variables$$

where x_{ij} and x_{ij}^* are elements of the original **X** and the reconstructed **X*** (Equation 2) data matrix, respectively.

4. Results

Evidence of a third radical

Figure 2 shows PCA score plots of EPR spectra (Figure 2a) and spectra normalized to unit area (Figure 2b), the latter to compensate for signal intensity decay upon heating (see Figure S1.1, Supplemental Information).[21, 25] In both cases, samples form two separate clusters reflecting heating time. Controls and samples heated for only 1 minute cluster at positive PC1, whereas samples heated for longer intervals cluster at negative PC1. Moreover, the heated sample cluster displays a clear trend from positive to negative PC2 values, which follows the sample heating time and temperature. Thus, these PCA results

indicate that heated and non-heated sample spectra display different characteristics. Furthermore, as these differences are also clear in the PCA of normalized spectra, spectral characteristics other than reduced overall intensity are important for the sample separation caused by heating. This agrees with the observed change in the EPR spectrum features with heating (see e.g. Figure 1), which most likely is due to differential thermal behaviors of the radical component spectra.

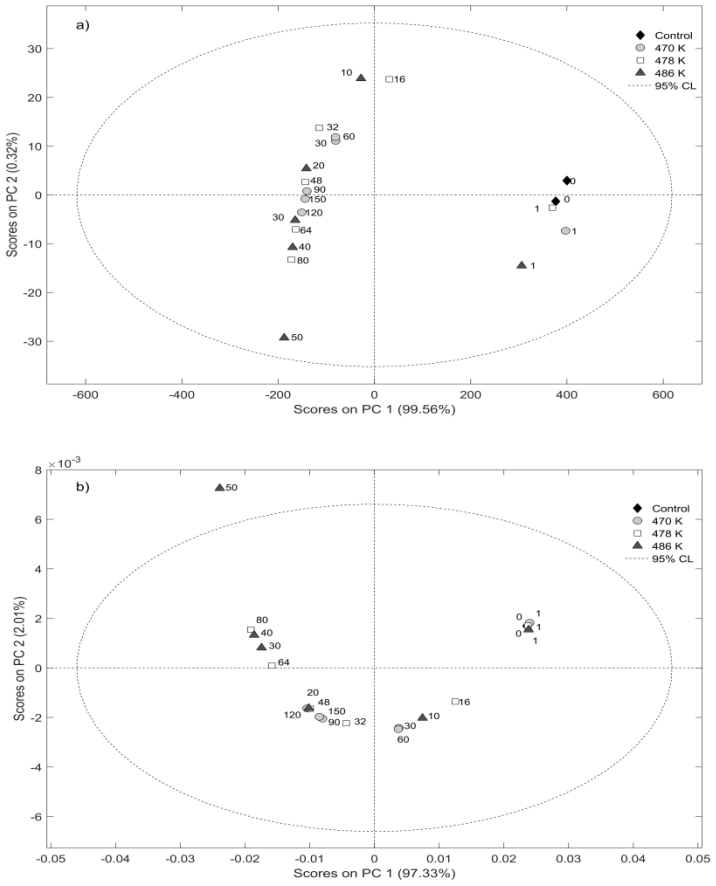


Figure 2: The PCA score plot of (a) EPR spectra and (b) EPR spectra normalized to unit area. The figures show the controls (♦) and the samples heated at 470 K (○), 478 K (□), or 486 K (▲), in the new PC space. The sample points are labelled with the heating time in minutes. The ellipse shows the 95% confidence interval of the PCA model.

The initial least square fit, presented in Table S1-3 (Supplemental Information), shows that control spectra and spectra of samples heated only for 1 minute could be reproduced fairly well using only the simulated **R1** and **R2** spectra. This similarity between control and 1 min heated samples is supported by the PCA score plots in Figure 2, where these spectra are found to cluster together. Hence, the 1 min heated spectra and the unheated spectra were grouped together and referred to as non-heated samples. On the other hand, fitting heated spectra at higher temperatures by only two components resulted in systematically increasing lack of fit (LOF) and decreasing correlation (Table S1-3, Supplemental Information), suggesting the significant presence of one or several additional radical component(s). This is also in agreement with the PCA results (Figure 2).

Good approximations of the experimental spectra were obtained using simulated **R1** and **R2** radical spectra together with the tentative simulation for a **R3** radical, as described in Section S1.3 (Supplemental Information), further supporting the three-radical hypothesis. In addition, realistic radical contribution fractions were obtained and indicated, in agreement with previous work,[21, 25] that the relative **R1** fraction decreased and the relative **R2** and **R3** fractions increased with longer annealing times.

*Estimates of an **R3** spectrum*

It is important to note, as shown in Table 1, that the simulated **R1** and **R2** component spectra are correlated, but that correlations between the tentative (simulated) **R3** spectrum and the simulated **R1** and **R2** spectra are low. Some correlation between the estimated radical spectra calculated using multivariate analysis methods is therefore to be expected. Furthermore, the simulated **R1** spectrum and the control spectra were strongly correlated, indicating the dominance of the **R1** radical in non-heated samples. Heated spectra, however, were correlated to all three simulated spectra.

Table 1. Pearson's correlations between the simulated radical spectra (**R1**, **R2**, and tentative **R3** spectrum), the mean experimental control spectrum and the mean experimental heated spectrum[a].

Correlations	R1	R2	R3	Controls	Heated
R1	1.00	0.47	0.11	0.98	0.61
R2	0.47	1.00	0.28	0.60	0.88
R3	0.11	0.28	1.00	0.18	0.61
Controls	0.98	0.60	0.18	1.00	0.72
Heated	0.61	0.88	0.61	0.72	1.00

[a] The mean experimental heated spectrum was taken as the average spectrum of samples heated at 478 K or 486 K for > 20 minutes

Neither principal component analysis (PCA), factor analysis (MLCFA) nor Independent Component Analysis (ICA) provided suitable estimates of both of the well-known simulated **R1** and **R2** radical spectra (see Table S2-1 and Figures S2.1, S2.2 and S2.3 (Supplemental Information)). Each of the three methods gave one spectral estimate (i.e., PC loadings, Common Factor, Independent Component, respectively) that correlated strongly with the dominant **R1** simulated spectrum, and two estimates that correlated either with combinations of the simulated **R2** and **R3** spectra with to neither of them. For MLCFA this was expected as it is well known[22] that the real spectral components only can be obtained by linearly combining the Common Factors. Up-weighting the heated spectra by normalization to unit area or only considering heated samples (i.e. a subset of the samples), yielded estimates that resembled combinations, rather than the pure simulated spectra.

As shown in Table 2, Self-modelling Mixture Analysis (SMA) and Multivariate Curve Resolution (MCR) provided realistic estimates of the simulated **R1** and **R2** radical spectra,

as well as the tentatively simulated **R3** spectrum. The spectral estimates found by SMA had over 90% correlation with each of the three simulated spectra. The SMA spectrum estimates had slightly higher cross correlations to other spectra as compared to the corresponding correlations between the simulated spectra (Table 1). This resulted in slight discrepancies between the estimated and simulated **R1** and **R2** spectra, as shown in Figure 3, panel a.

Table 2. The Pearson correlations between the established (simulated) **R1** and **R2** radical spectra, the tentative **R3** spectrum, and the component spectra (S) estimated using SMA and MCR[a].

Method		**R1**	**R2**	**R3**	**LOF**	**Correlation**
SMA						
Offset 8[b]	S1	-0.93	-0.51	-0.42		
	S2	-0.51	-0.93	-0.53	0.027	0.999
	S3	0.22	0.48	0.94		
MCR						
Non-negative contributions[c]	S1	0.98	0.58	0.16		
	S2	0.66	0.91	0.50	0.027	0.999
	S3	0.37	0.80	0.78		
R2 spectrum & control radical contributions[d]	S1	0.98	0.58	0.12		
	S2	0.47	1.00	0.28	0.164	0.986
	S3	0.23	0.41	0.94		

[a]The lack of fit (LOF) and the correlation between the original data matrix **X** and the reconstructed data matrix **X*** are also shown. Spectra were normalized to unit area prior to decomposition.

[b]Offset of 8% of the maximum mean value of the experimental EPR spectra.

[c]Non-negative contributions constraint only.

[d]Partially known **R2** spectrum and control radical contribution ratio (0.60:0.35:0.05 for **R1**:**R2**:**R3**, [24]), in addition to the non-negativity constraint. Samples heated for 1 min. were included in the control group. These equality constraints were weighted as "very soft" (weight factor 1, where a weight factor 10 signifies hard-constrained).

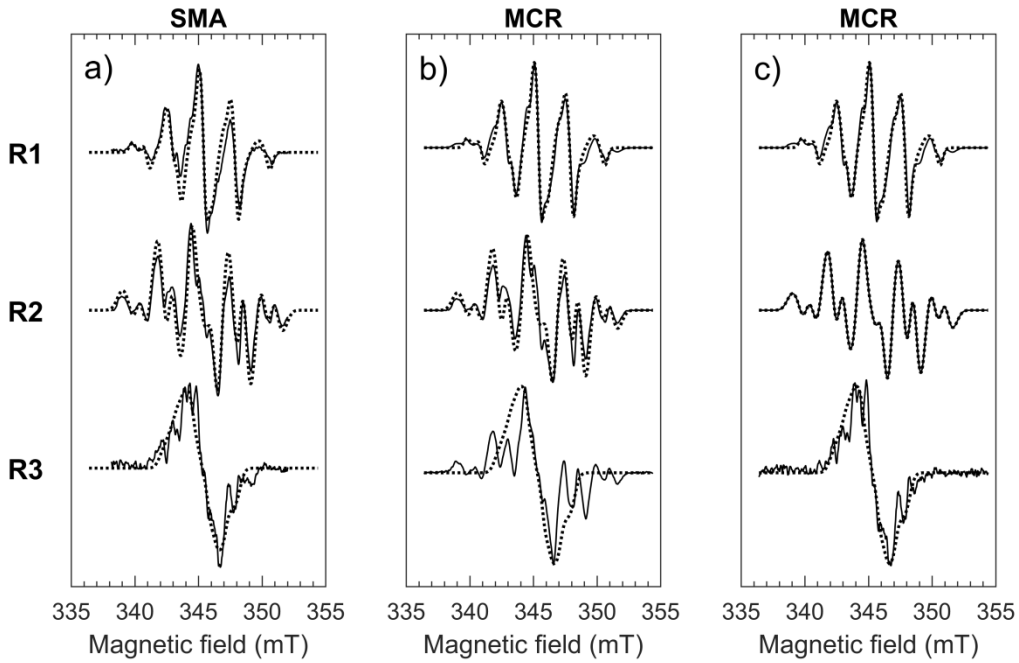


Figure 3. Spectral estimates (solid lines) for radical **R1** (top), **R2** (middle), and **R3** (bottom), obtained using (a) SMA with 8% offset, (b) MCR with non-negative contribution constraint, and (c) MCR with very soft (weight 1) equality constraints in addition to the non-negativity constraint (i.e. partially known **R2** spectrum and non-heated (including controls) sample radical contribution ratio (0.60:0.35:0.05)). The dotted lines show the simulated radical spectra (Figure S1.2 Supplemental Information). Spectra have been scaled for comparison purposes.

Using only the non-negative contribution fractions constraint, the MCR analysis resulted in **R1** and **R2** estimates with over 90% correlation with the corresponding simulated spectra, but with rather high cross correlations to the other simulated spectra (Table 2). Thus, the **R1** and **R2** estimates deviated somewhat from the simulated **R1** and **R2** spectra (Figure 3b). The estimated **R3** spectrum resembled and correlated to a combination of the simulated **R2** and **R3** spectra.

When very soft equality constraints (1 on typically a 1-10 scale) were applied for both the shape of the **R2** radical spectrum and the contribution fractions of the control samples (including, based on the PCA results in Figure 2, samples heated for 1 min), the MCR analysis yielded estimated spectra with very high correlations with the simulated spectra (Figure 3c, Table 2). The cross correlations with other spectra were low and not unlike the correlations found between the simulated spectra (Table 1).

Radical contribution fractions

Only radical contribution fractions found by SMA and MCR are presented (Figure 4), as only these methods provided realistic **R1** and **R2** spectral estimates (see Table 2 and Figure 3). Both methods gave contribution fractions that followed the single exponential function of Equation S1 (Supplemental Information) as a function of annealing time (correlation 0.99 ± 0.01). The initial contribution ratio (**R1**:**R2**:**R3**) estimated by SMA was approximately 0.40:0.55:0.05. The **R3** contribution increased from about 5% to 40-50% upon annealing, whereas the **R1** and **R2** contributions decreased to about 20-30%.

MCR estimated that the **R1** radical dominated samples heated for 1 min or less (> 95% contribution). Heating reduced the **R1** contribution to about 5-30%, depending upon temperature. The **R2** and **R3** contributions increased from about 1-2% to 20-60% after heating. The fractions calculated using MCR summed to 1.02 ± 0.04.

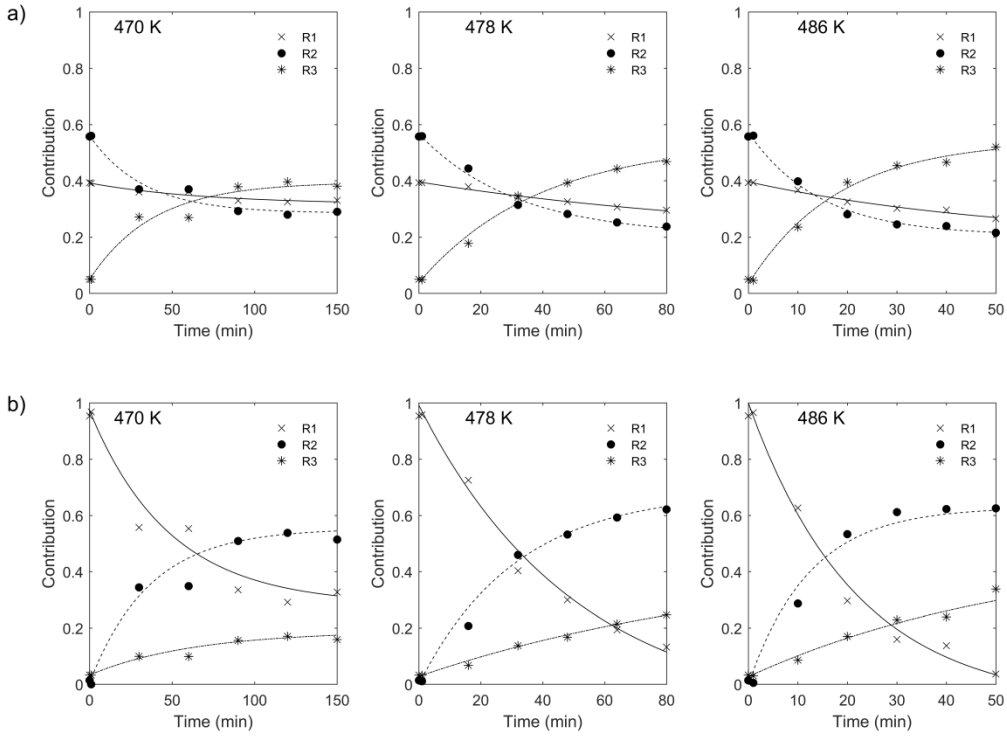


Figure 4. The contribution fractions of the radicals **R1** (x), **R2** (o) and **R3** (∗), as a function of time at 470 K, 478 K and 486 K, estimated using (a) SMA with 8% offset, and (b) MCR with the non-negative contributions constraint and very soft (weight 1) equality constraints (i.e. partially known **R2** spectrum and non-heated contribution ratio). The lines show the exponential fit (Equation S1) to the contribution data.

5. Discussion and Conclusions

Previous studies [20-27] suggest that experimental EPR spectra of irradiated alanine consist of a mixture of contributions from three different radical species. Presently, the simulated spectra of two components, the **R1** and **R2** radical spectra, are highly credible based on detailed hyperfine coupling constants obtained from experimental data.[23] However, there is currently not sufficient experimental evidence to propose a reliable EPR spectrum of the **R3** radical[24-25] although a tentative simulation has been proposed (Figure S1.2

(Supplemental Information)). In the present study, several multivariate techniques were explored to elucidate the shapes of the radical component spectra, with particular emphasis on estimating a spectrum for the assumed third radical **R3**. The challenge was to extract pure component spectra and component contribution fractions from experimental EPR spectra. The anticipated high degree of overlap of the pure radical component spectra as well as the presence of more than one radical component in all alanine samples complicated the issue.

All decomposition methods provided excellent reconstruction of the experimental spectra, as assessed by the very low lack of fit (LOF) and the high correlation between the original and reconstructed spectra. However, not all methods provide satisfactory estimates of the underlying radical component spectra. Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Maximum Likelihood Common Factor Analysis (MLCFA) provided spectral estimates resembling combinations of rather than pure simulated radical component spectra. As pointed out by Parastar *et al*.[38], pure chemical signatures will most likely not comply with the orthogonality constraint of PCA or the statistical independence condition of ICA.[36, 40] This was also the case in the current study, where the simulated radical spectra were neither uncorrelated nor orthogonal. MLFCA focuses on capturing the common variance in the data set[43, 45], which may be appropriate for estimating the dominant spectrum, but other shared spectral characteristics are combined into the remaining common factors. Thus, PCA, ICA and MLCFA were not optimal approaches for resolving all radical spectra for the present data set.

MLCFA has previously been used to study the temperature dependence of EPR spectra of irradiated alanine.[21-22] In agreement with the current study conducted using a radiation dose of 5 kGy, three common factors were found for the 1-10 kGy dose range.

Although these factors could be used to reconstruct the proposed **R1** and **R2** spectra, only one common factor bore resemblance to the simulated **R1** radical component spectra. The other two factors were combinations of the underlying radical spectra.[22] These results are in accordance with the findings of the present study where MLCFA could resolve the dominant underlying spectrum, but was unable to resolve other individual spectra. The kinetic analysis of temperature dependent real **R1** and **R2** component spectra based on MLCFA analyses by Vanhaelewyn *et al*.[22] is not unlike the present SMA results (Figure 4a) even if the **R3** spectral component was missing in that work.

In the present study, two methods, Multivariate Curve Resolutions (MCR) and Self-modelling Mixture Analysis (SMA), provided estimates of the alanine radical **R1** and **R2** spectra that were in good agreement with the simulated spectra. These estimated spectra had over 90% correlation with the corresponding simulated spectra, albeit with somewhat more cross correlation in between the spectral estimates than found between the simulated spectra (Table 1). This resulted in some deviations between the estimated and simulated spectra. As the simulated **R2** radical spectrum is slightly wider than the simulated **R1** radical spectrum (see Figure S1.2, Supplemental Information), the correlation between **R1** and **R2** radical spectra estimated using SMA and MCR can easily be spotted visually in the **R1** spectrum extreme high- and low field regions (top spectra in Figure 3). Furthermore, the radical contribution fractions as a function of heating time estimated by SMA and MCR followed the anticipated single exponential function (Figure 4). The initial contribution fractions estimated by SMA (0.40:0.55:0.05) were not unlike the 0.60:0.35:0.05 ratio proposed previously[24], albeit with a higher **R2** than **R1** contribution. A higher **R2** than **R1** contribution is contrary to the previous experimental evidence for the spectral composition of the alanine spectra [24-25, 49-50] but is comparable to

the results of Vanhaelewyn et al. obtained using MLCFA.[22] On the other hand, MCR estimated that **R1** clearly dominated the non-heated samples (>95 % contribution). This large contribution is most probably partly due to a non-negligible mixing of **R2** into **R1**. Regardless, the **R1** radical dominance agrees well with the experimental evidence as well as with the correlation analysis of Table 1 showing that **R1** dominates the control spectra.

It is very interesting to note that the estimated **R3** radical spectra predicted by SMA and weakly constrained MCR (Figures 3a and 3c, respectively, bottom spectra) are surprisingly similar even in small details. They are also very similar to the tentative simulated **R3** spectrum. If more experimental data becomes available, this similarity may be used for guiding further refinements of extended simulations of the **R3** spectrum.

A strength of the SMA technique is that it does not require any prior knowledge of the underlying spectra or sample composition to guide the method to a realistic solution.[44,51] The SMA algorithm focuses on spectral regions most characteristic of each component in the sample, the so-called pure variables, while avoiding more problematic regions with, for example, high degree of overlap.[51] SMA gives very good estimates of all three simulated model spectra whereas the contributions analysis (Figure 4) provides component contributions to the overall mixture spectra which are not in agreement with previous experimental data.[24-25, 49-50] In contrast, MCR requires the use of some constraints to reach a unique solution.[40, 45, 47] These constraints, however, enable the incorporation of prior system knowledge into the modelling and account for the flexibility of the MCR technique. Chemical information such as known concentrations/pure spectra, mass-balance conditions or various mathematical conditions can be imposed with variable strengths. Thus, the MCR model can be fine-tuned such that the results are chemically and physically meaningful. In the MCR analysis, **R2** apparently has a larger influence on the estimated **R1**

component representation than is the case for SMA, as judged by the discrepancies in the extreme high and low field regions (Figure 3 top). However, the MCR contributions analysis is overall more satisfactory with respect to previous experimental data.[24-25, 49-50] SMA and MCR are both very useful approaches, but the extra feature of allowing for constraints is probably what gives MCR the leading edge in the present work. Such features could, however, potentially bias the results.

In summary, all decomposition methods yield results that resembled the established **R1** simulated spectrum. In addition, SMA and MCR consistently predicted both the established **R2** simulated spectrum as well as the shape of the tentative **R3** simulated spectrum. Thus, the multivariate decomposition approach used in the present work provides results that support the previously proposed radical component spectra. The **R1** and **R2** component spectra were initially determined at very high credibility from experimental EPR measurements, and the radical structures arrived at from these data have been corroborated using quantum chemical calculations.[28-35]. The suggested chemical structure for the postulated third radical component **R3** is also supported by quantum chemical calculations. By using two independent multivariate data analysis techniques, the present work strongly supports the basic features of the previously proposed spectral shape of the **R3** component, thereby also strengthening the evidence that three radicals are indeed present in irradiated alanine samples. Still, more extensive experimental data for the **R3** radical remains to be obtained in order to firmly establish the detailed spectral shape.

ACKNOWLEDGEMENTS

KMV acknowledges a student exchange grant from the University of Oslo, and support from the University of Calgary. Thanks are due to Mr. Efim Brondz for excellent technical assistance.

SUPPLEMENTAL INFORMATION

Part 1: Additional data on the spectral behavior of thermally annealed irradiated alanine pellets and results from the two- and three-component spectral reconstructions.

Part 2: Additional data on the PCA, ICA and MLCFA results.

ORCID IDENTIFIERS

Cecilia Futsaether: orcid.org/0000-0001-7944-0719

Knut Kvaal: orcid.org/0000-0002-0579-7448

Eirik Malinen: orcid.org/0000-0002-1308-9871

Einar Sagstuen**:** orcid.org/0000-0003-4261-1398

Turid Torheim: orcid.org/0000-0001-6191-2036

REFERENCES

(1) Luck, C. F.; Gordy, W. Effects of X-irradiation upon some organic substances in the solid state: Simple alcohols, amines, amides and mercaptans. *J. Amer. Chem. Soc.* **1956,** *78*, 3240 - 3243.

(2) Ehrenberg, A.; Ehrenberg, L.; Zimmer, K. G. Production of magnetic centers in glycine by X-rays. *Acta Chem. Scand.* **1957,** *11*, 199-201.

(3) Shields, H.; Gordy, W. Electron spin resonance studies of radiation damage to amino acids. *J. Phys. Chem.* **1958,** *62*, 789-798.

(4) Ghosh, D. K.; Whiffen, D. H. Electron spin resonance (EPR) spectrum of a γ-irradiated single crystal of glycine. *Molec. Phys.* **1959,** *2*, 285-300.

(5) McConnell, H. M.; Heller, C.; Cole, T.; Fessenden, R. W. Radiation damage in organic crystals. I. $CH(COOH)_2$ in malonic acid. *J. Amer. Chem. Soc.* **1960,** *82*, 766-75.

(6) Bradshaw, W. W.; Cadena, D. G.; Crawford, G. W.; Spetzler, H. A. W. The use of alanine as a solid dosimeter. *Radiat. Res.* **1962,** *17*, 11 - 21.

(7) Regulla, D. F.; Deffner, U. Dosimetry by ESR spectroscopy of alanine. *Appl. Radiat. Isot.* **1982,** *33*, 1101-1114.

(8) Wieser, A.; Siegele, R.; Regulla, D. F. Influence of the irradiation temperature on the free-radical response of alanine. *Appl. Radiat. Isot.* **1989,** *40*, 957-959.

(9) Hansen, J. W.; Olsen, K. J. Theoretical and experimental radiation effectiveness of the free radical dosimeter alanine to irradiation with heavy charged particles. *Radiat. Res.* **1985,** *104*, 15 - 27.

(10) Hansen, J. W.; Olsen, K. J. Predicting decay in free-radical concentration in l-α-alanine following high-LET radiation exposures. *Appl. Radiat. Isot.* **1989,** *40*, 935-939.

(11) Hansen, J. W.; Olsen, K. J.; Wille, M. The alanine radiation detector for high and low-LET dosimetry. *Radiat Prot Dosimetry* **1987,** *19*, 43-47.

(12) Regulla, D. F. From dating to biophysics - 20 years of progress in applied ESR spectroscopy. *Appl. Radiat. Isotop.* **2000,** *52*, 1023-1030.

(13) Yordanov, N. D.; Gancheva, V. Recent development of EPR dosimetry. In *EPR of free radicals in solids. Trends in methods and applications*, Lund, A.; Shiotani, M., Eds. Kluwer Academic Publishers: Dordrecht, 2003.

(14) Malinen, E. EPR dosimetry in clinical applications. In *Application of EPR in radiation research*, Lund, A.; Shiotani, M., Eds. Springer International, Inc.: Heidelberg, 2014; pp 509 - 538.

(15) Desrosiers, M. F. Alanine-EPR high dose radiation metrology. In *Applications of EPR in radiation research*, Lund, A.; Shiotani, M., Eds. Springer International, Inc.: Heidelberg, 2014; pp 489-507.

(16) Anton, M.; Allisy-Roberts, P. J.; Kessler, C.; Burns, D. T. A blind test of the alanine dosimetry secondary standard of the PTB conducted by the BIPM. *Metrologia* **2014,** *51,Tech.Suppl. 06001*, 1-11.

(17) Sharpe, P. H. G.; Rajendran, K.; Sephton, J. P. Progress towards an alanine/ESR therapy level reference dosimetry service at NPL. *Appl. Radiat. Isotop.* **1996,** *47*, 1171-1175.

(18) Kuroda, S.-i.; Miyagawa, I. ENDOR study of an irradiated crystal of l-alanine: Environment of the stable $CH_3CHCO_2^-$ radical. *J. Chem. Phys.* **1982,** *76*, 3933 - 3944.

(19) Matsuki, K.; Miyagawa, I. ENDOR study of an irradiated crystal of l-alanine: Structure and the environment of the unstable CH_3CHCOO- radical. *J. Chem. Phys.* **1982,** *76*, 3945-3952.

(20) Callens, F.; vanLaere, K.; Mondelaers, W.; Matthys, P.; Boesman, E. A study of the composite character of the ESR spectrum of alanine. *Appl. Radiat. Isotopes* **1996,** *47*, 1241-1250.

(21) Vanhaelewyn, G. C. A. M.; Mondelaers, W. K. P. G.; Callens, F. J. Effect of temperature on the electron paramagnetic resonance spectrum of irradiated alanine. *Radiat. Res.* **1999,** *151*, 590 - 594.

(22) Vanhaelewyn, G. C. A. M.; Amira, S. A.; Mondelaers, W. K. P. G.; Callens, F. J. Decomposition study of the electron paramagnetic resonance spectrum of irradiated alanine. *Spectrochim. Acta A* **2000,** *56*, 387 - 397.

(23) Sagstuen, E.; Hole, E. O.; Haugedal, S. R.; Nelson, W. H. Alanine radicals: structure determination by EPR and ENDOR of single crystals x-irradiated at 295 K. *J. Phys. Chem.* **1997,** *A101*, 9763-9772.

(24) Heydari, M. Z.; Malinen, E.; Hole, E. O.; Sagstuen, E. Alanine radicals. 2. The composite polycrystalline alanine EPR spectrum studied by ENDOR, thermal annealing, and spectrum simulations. *J. Phys. Chem. A* **2002,** *106*, 8971-8977.

(25) Malinen, E.; Heydari, M. Z.; Sagstuen, E.; Hole, E. O. Alanine radicals, part 3: properties of the components contributing to the EPR spectrum of x-irradiated alanine dosimeters. *Radiat. Res.* **2003,** *159*, 23-32.

(26) Malinen, E.; Hult, E. A.; Hole, E. O.; Sagstuen, E. Alanine radicals, part 4: relative amounts of radical species in alanine dosimeters after exposure to 6-19 MeV electrons and 10 kV-15 MV photons. *Radiat. Res.* **2003,** *159*, 149-153.

(27) Sagstuen, E.; Sanderud, A.; Hole, E. O. The solid state radiation chemistry of simple amino acids, revisited. *Radiat. Res.* **2004,** *162*, 112 - 119.

(28) Ban, F.; Wetmore, S. D.; Boyd, R. J. A Density-functional theory investigation of the radiation products of l-α-alanine. *J. Phys. Chem. A* **1999,** *103*, 4303-4308.

(29) Lahorte, P.; De Proft, F.; Vanhaelewyn, G.; Masschaele, B.; Cauwels, P.; Callens, F.; Geerlings, P.; Mondelaers, W. Density functional calculations of hyperfine coupling constants in alanine-derived radicals. *J. Phys. Chem A* **1999,** *103*, 6650-6657.

(30) Bugay, A. A.; Onischuk, V. A.; Petrenko, T. L.; Teslenko, V. V. The mechanisms of radical formation in l-α-alanine. *Appl. Radiat. Isot.* **2000,** *52*, 1189-1193.

(31) Pauwels, E.; Van Speybroeck, V.; Lahorte, P.; Waroquier, M. Density functional calculations on alanine-derived radicals: influence of molecular environment on EPR hyperfine coupling constants. *J. Phys. Chem. A* **2001,** *105*, 8794-8804.

(32) Pauwels, E.; Van Speybroeck, V.; Waroquier, M. Application of molecular cluster models to study the amino acid l-α-alanine and its derived radicals in the crystalline state. *Int. J. Quantum Chem.* **2003,** *91*, 511-516.

(33) Petrenko, T. L. Transformation and structure of cation radicals in l-α-alanine. *J. Phys. Chem. A* **2002,** *106*, 149 - 156.

(34) Simion, C. Theoretical investigation of radical species formed from l-α-alanine under gamma-irradiation. *J. Radioanalyt. Nucl. Chem.* **2008,** *275*, 331-335.

(35) Pauwels, E.; De Cooman, H.; Waroquier, M.; Hole, E. O.; Sagstuen, E. Solved? The reductive radiation chemistry of alanine. *Phys. Chem. Chem. Phys.* **2014,** *16*, 2475-2482.

(36) Geladi, P. Chemometrics in spectroscopy. Part 1. Classical chemometrics. *Spectrochim. Acta B* **2003,** *58*, 767-782.

(37) Hyvärinen, A. Independent component analysis: recent advances. *Phil. Trans. Roy. Soc. A* **2012,** *371*.

(38) Parastar, H.; Jalali-Heravi, M.; Tauler, R. Is independent component analysis appropriate for multivariate resolution in analytical chemistry? *Trends Analyt. Chem.* **2012,** *31*, 134-143.

(39) Windig, W.; Gallagher, N. B.; Shaver, J. M.; Wise, B. M. A new approach for interactive self-modeling mixture analysis. *Chemometr. Intell. Lab. Syst.* **2005,** *77*, 85-96.

(40) de Juan, A.; Jaumot, J.; Tauler, R. Multivariate curve resolution (MCR). Solving the mixture analysis problem. *Analyt. Meth.* **2014,** *6*, 4964-4976.

(41) Hyvärinen, A.; Oja, E. Independent component analysis: algorithms and applications. *Neur. Netw.* **2000,** *13*, 411-430.

(42) Johnson, R. A.; Wichern, D. W. *Applied Multivariate Statistical Analysis*. 6th ed.; Pearson: Cloth, 2008.

(43) Persoone, P.; De Gryse, R.; De Volder, P. A new powerful transformation for maximum likelihood common factor analysis (MLCFA). *J. Electr. Spectr. Rel. Phenom.* **1995,** *71*, 225-232.

(44) Cuesta Sánchez, F.; van den Bogaert, B.; Rutan, S. C.; Massart, D. L. Multivariate peak purity approaches. *Chemometr. Intell. Lab. Syst.* **1996,** *34*, 139-171.

(45) Wehrens, R. *Chemometrics with R. Multivariate data analysis in the natural sciences and life sciences*. Springer-Verlag Berlin: Berlin Heidelberg, 2011; p 286.

(46) Windig, W.; Guilment, J. Interactive self-modeling mixture analysis. *Analyt. Chem.* **1991,** *63*, 1425-1432.

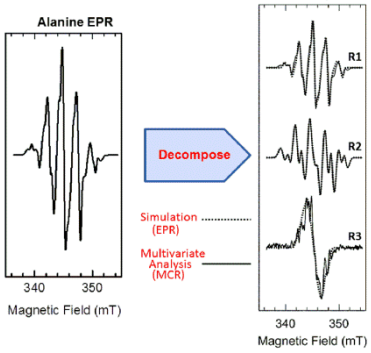
(47) Ruckebusch, C.; Blanchet, L. Multivariate curve resolution: A review of advanced and tailored applications and challenges. *Analyt. Chim. Acta* **2013,** *765*, 28-36.

(48) Jaumot, J.; de Juan, A.; Tauler, R. MCR-ALS GUI 2.0: New features and applications. *Chemometr. Intell. Lab. Syst.* **2015,** *140*, 1-12.

(49) Ciesielski, B.; Schultka, K.; Penkowski, M.; Sagstuen, E. EPR study of light illumination effects on radicals in gamma-irradiated l-alanine. *Spectrochim. Acta A* **2004,** *60*, 1327-1333.

(50) Ciesielski, B.; Tyszkowskaa, M.; Grudniewskaa, A.; Penkowski, M.; Schultka, K.; Peimel-Stuglik, Z. The effect of dose on light-sensitivity of radicals in alanine EPR dosimeters. *Spectrochim. Acta A* **2008,** *69*, 1405-1416.

(51) Bogomolov, A.; Hachey, M. Application of SIMPLISMA purity function for variable selection in multivariate regression analysis: A case study of protein secondary structure determination from infrared spectra. *Chemometr. Intell. Lab. Syst.* **2007,** *88*, 132-142.


Alanine EPR
340
350
Magnetic Field (mT)
Decompose
Simulation (EPR)
Multivariate Analysis (MCR)
R1
R2
R3
340
350
Magnetic Field (mT)

TOC Graphic