Vol. 38 — N° 1 03 14 P. 73 – 94

GEOSTANDARDS and GEOANALYTICAL RESEARCH

Towards a Higher Comparability of Geothermometric Data obtained by Raman Spectroscopy of Carbonaceous Material. Part I: Evaluation of Biasing Factors

Nils K. **Lünsdorf** (1)*, Istvan **Dunkl** (1), Burkhard C. **Schmidt** (2), Gerd **Rantitsch** (3) and Hilmar von **Eynatten** (1)

(1) Department of Sedimentology and Environmental Geology, Geoscience Center Georg-August-Universität Göttingen, 37077 Göttingen, Germany
(2) Department of Experimental and Applied Mineralogy, Geoscience Center Georg-August-Universität Göttingen, 37077 Göttingen, Germany
(3) Department of Applied Geosciences and Geophysics, University of Leoben, 8700 Leoben, Austria
* Corresponding author. e-mail: kluensd@gwdg.de

Raman spectroscopy of carbonaceous material (RSCM) is frequently used to determine metamorphic peak temperatures from the structural order of carbonaceous material enclosed in metasediments. This method provides a quick, robust and relatively cheap geothermometer. However, the comparability of the RSCM parameter is low as there are at least three major sources of biasing factors. These sources are the spectral curve-fitting procedure, the sample characteristics itself and the experimental design including the used Raman system. To assess the impacts of the biasing factors on RSCM, a series of experiments was performed. The experiments showed that curve-fitting is strongly influenced by individual operator-bias and the degrees of freedom in the model, implying the need for a standardised curve-fitting procedure. Due to the diversity of components (optics, light detection device, gratings, etc.) and their combinations within the Raman systems, different Raman instruments generally give differing results. Consequently, to estimate comparable metamorphic temperatures from RSCM data, every Raman instrument needs its own calibration. This demands a reference material series that covers the entire temperature calibration range. Although sample heterogeneity will still induce some variation, a reference material series combined with standardised curve-fitting procedures will significantly increase the overall comparability of RSCM data from different laboratories.

Keywords: Raman spectroscopy, carbonaceous material, comparability, geothermometer.

La spectroscopie Raman du matériel carboné (RSCM) est fréquemment utilisée pour déterminer les températures des pics métamorphiques à partir de l'organisation structurale du matériau carboné des métasédiments. Cette méthode fournit un géothermomètre rapide, robuste et relativement peu onéreux. Toutefois, la comparabilité du paramètre RSCM est faible car il y a au moins trois grandes sources de biais. Ces sources sont la procédure d'ajustement de la courbe spectrale, les caractéristiques de l'échantillon lui-même et la conception expérimentale, y compris le système Raman utilisé. Pour évaluer les impacts des facteurs de biais sur la RSCM, une série d'expériences a été effectuée. Les expériences ont montré que la procédure d'ajustement de la courbe spectrale est fortement influencée par les biais de l'opérateur et les degrés de liberté du modèle, ce qui implique la nécessité d'une procédure d'ajustement de type courbe standardisée. En raison de la diversité des composants (optique, dispositif de détection de la lumière, grilles, etc.) et leurs combinaisons au sein des systèmes Raman, différents instruments Raman donnent généralement des résultats différents. Par conséquent, pour estimer des températures métamorphiques comparables à partir des données RSCM, chaque instrument Raman a besoin de son propre étalonnage. Cela exige l'utilisation d'une série de matériaux de référence qui couvre toute la gamme d'étalonnage de la température. Bien que l'hétérogénéité des échantillons induise encore une certaine variation, une série de matériaux de référence combinée avec des procédures d'ajustement de courbes standardisées permettra d'accroître considérablement la comparabilité globale des données de RSCM provenant de différents laboratoires.

Mots-clés : spectroscopie Raman, matériel carbonée, comparabilité, géothermomètre.

Received 18 Feb 13 – Accepted 25 Jul 13

doi: 10.1111/j.1751-908X.2013.12011.x

Geothermometry by Raman spectroscopy of carbonaceous material (RSCM; Beyssac *et al.* 2002a) is becoming more and more popular due to its ease of applicability and non-destructiveness (e.g., Rantitsch *et al.* 2004, Guedes *et al.* 2005, Forer *et al.* 2009, Huang *et al.* 2010, Wiederkehr *et al.* 2011, Endo *et al.* 2012). The Raman spectrum of carbonaceous material (CM) enclosed in metasediments changes systematically with increasing degree of metamorphism (e.g., Pasteris and Wopenka 1991, Jehlička and Bény 1992, Wopenka and Pasteris 1993). Beyssac *et al.* (2002a) showed that those changes are mainly controlled by temperature and calibrated a geothermometer based on CM Raman spectra. This geothermometer is based on parameters calculated by spectral curve-fitting of the Raman bands (Figure 1). However, there are a multitude of different curve-fitting strategies using variable numbers (2–5) of model components (Lorentzian-, Voigt-, Gaussian- or Breit–Wigner–Fano–functions), fitting the acquired spectrum (e.g., Beyssac *et al.* 2002a, Quirico *et al.* 2003, Sadezky *et al.* 2005, Lahfid *et al.* 2010). Important function parameters (position, height, full width at half maximum FWHM and area), reflecting the thermal transformation of CM, are derived from the extracted components. This allows the calibration of different parameter ratios against metamorphic temperature (Beyssac *et al.* 2002a, Rantitsch *et al.* 2004, Rahl *et al.* 2005, Baziotis *et al.* 2006, Aoya *et al.* 2010, Lahfid *et al.* 2010).

As the parameter ratios are the condensates of many steps (e.g., sample preparation, Raman measurement and spectrum evaluation), they accumulate several biasing factors. The sources of bias can be grouped into three categories: (a) bias intrinsic to spectral curve-fitting (b) bias intrinsic to the CM and (c) bias intrinsic to the experimental design and the specific Raman system used.

Factors of the first category are different baseline corrections, different mathematical functions (Gaussian, Lorentzian, Voigt, etc.) used for peak-fitting, and the different number of components used to model the Raman spectra.

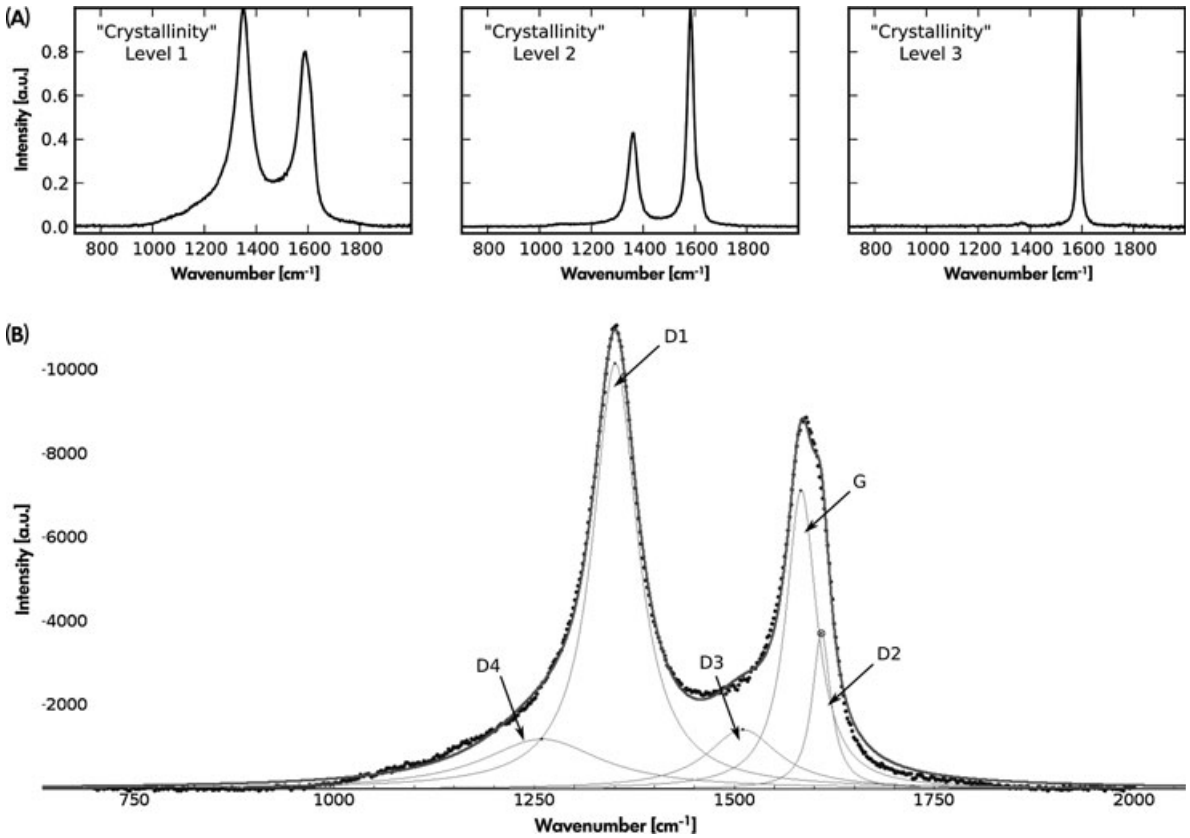

Figure 1. (A) Representative Raman spectra of the three 'crystallinity levels'. (B) An example for the decomposition of a 'crystallinity level 1' Raman spectrum by five components according to Lahfid *et al.* (2010). The two bands of the first-order region are described by D1, D2, D3, D4 and G components. D4 widens the low wavenumber side of the D1 band, while D2 appears only as a weak shoulder on the high wavenumber side of the G band.

Examples for the second category are structural anisotropy, sample preparation and sample heterogeneity. Factors of the third category include used excitation wavelength, spectral grating and light detection device, among others.

In this paper, many of the above-mentioned potential sources of bias are evaluated by a suite of simple experiments in which, at best, only one parameter influences the experimental results.

Methods and samples

Samples

The sample set (Table 1) covers a wide structural range from low to high 'crystallinity' degree of CM. The studied samples derive from the Triassic flysch of the Tethyan Himalayan sequence of SE Tibet (Dunkl *et al.* 2011), from the Eastern Alps (Rantitsch *et al.* 2004) and from the Thuringian Forest (Germany) (Kunert 1999).

In order to enrich CM, every sample has been treated chemically. The samples were crushed to particles smaller than 5–10 mm. The rock-chips were initially placed in a 1:1 solution of 37% hydrochloric acid to dissolve carbonates and after decantation mixed with 1:1 diluted 48% hydrofluoric acid to dissolve silicates. After hydrofluoric acid treatment, the sample suspensions were decanted and diluted with de-ionised water until a pH-value of 5–6 was reached. Remaining fluids were evaporated in a drying oven at 50 °C. About 10–20 mg of the dried CM was mixed with 1–2 ml de-ionised water in a small glass vial and placed into an ultrasonic bath for about 60 s in order to disperse the carbonaceous material. This suspension was deposited on a glass slide.

Raman spectroscopy

All Raman measurements were performed with a Horiba Jobin Yvon HR800-UV spectrometer, with attached Olympus BX41 microscope, if not stated otherwise. The general measurement configuration used a 488 nm Ar^+ laser for excitation, a spectral grating with 600 l mm^{-1}, a long working distance 100 × objective with a numerical aperture of 0.8, and the diameter of the confocal hole was set to 100 μm. If not stated otherwise the laser light was circular polarised, a spectral range of 700–2000 cm^{-1} was recorded in one spectral window in 3–5 accumulations of 10–30 s. The laser power on the sample surface was controlled by density filters to 0.3–0.5 mW to exclude thermal alteration of the sample. Fifteen measurements per sample were conducted on different sample spots. The Raman system was calibrated against the 520.4 cm^{-1} line of a Si-waver.

Table 1.
List of samples

Sample	Long	Lat	Region	R1	*s*	R2	*s*	RA1	*s*	RA2	*s*	n
L29	87.1829	28.6685	SE Tibet	ND	ND	ND	ND	0.61	0.01	1.56	0.05	15
DB45	91.1048	29.0483	SE Tibet	ND	ND	ND	ND	0.58	0.02	1.40	0.10	15
DB21	91.6362	28.9271	SE Tibet	ND	ND	ND	ND	0.59	0.02	1.47	0.12	15
DB26	92.1574	29.1036	SE Tibet	0.73	0.06	0.49	0.01	ND	ND	ND	ND	15
DB28	92.0433	29.1454	SE Tibet	0.65	0.06	0.45	0.02	ND	ND	ND	ND	15
DB36	91.6764	28.9881	SE Tibet	0.63	0.03	0.46	0.02	ND	ND	ND	ND	15
L1	~ 87.332	~ 29.040	SE Tibet	0.40	0.05	0.37	0.02	ND	ND	ND	ND	15
DB16	91.1906	28.7015	SE Tibet	0.38	0.03	0.36	0.01	ND	ND	ND	ND	15
TU2	92.2606	28.8136	SE Tibet	0.32	0.04	0.33	0.02	ND	ND	ND	ND	15
L45	88.0787	28.8651	SE Tibet	0.27	0.03	0.30	0.02	ND	ND	ND	ND	15
L57	88.1518	28.8372	SE Tibet	0.15	0.02	0.22	0.02	ND	ND	ND	ND	15
Kohl1	47.6617	15.6575	Austria	ND	ND	ND	ND	0.63	0.01	1.74	0.07	15
MAU	47.0451	13.2452	Austria	0.44	0.03	0.40	0.02	ND	ND	ND	ND	15
KL2-2	50.3944	11.4014	Thuringian Forest	ND	ND	ND	ND	0.61	0.00	1.54	0.02	30
KL2-3	50.4009	11.3529	Thuringian Forest	ND	ND	ND	ND	0.60	0.00	1.47	0.02	30
KL2-4	50.362	11.4057	Thuringian Forest	ND	ND	ND	ND	0.60	0.01	1.50	0.04	30
KL2-11	50.2632	11.5218	Thuringian Forest	ND	ND	ND	ND	0.63	0.00	1.67	0.02	30
KL2-17	50.3549	11.5109	Thuringian Forest	ND	ND	ND	ND	0.63	0.00	1.73	0.01	30
KL2-18	50.3268	11.3777	Thuringian Forest	ND	ND	ND	ND	0.57	0.01	1.32	0.03	30

Long., longitude; Lat., latitude; *s*, standard deviation; n, number of measurements; ND, not determined.

Evolution of the first-order Raman spectrum of CM

After deposition and early diagenesis, the organic content of sedimentary rocks constitutes a heterogeneous mixture of organic compounds. During organic maturation, mainly O, H, N and to a lesser degree C are expelled from the organic material, changing the chemical composition and structure of the residual organic material. This process leads to an enrichment of aromatic species (for a review, see Vandenbroucke and Largeau 2007).

The aromatic species form so-called 'basic structural units' (BSU) of polyaromatic (4–10 cycles) layers, isolated or piled up by 2–3 units (Oberlin 1989). The nanometre-sized BSU is described by the mean stacking height (L_c) and the mean basal plane diameter (L_a). During the early stages of diagenesis and catagenesis, the BSUs are randomly oriented, but start to synchronise their orientation to form molecular orientation domains (Bustin *et al.* 1995, Vandenbroucke and Largeau 2007). During graphitisation, L_c and L_a progressively increase, while at the same time, the number of defects and the interplanar spacing between the graphene layers is reduced (Buseck and Huang 1985, Wopenka and Pasteris 1993). Thus, graphitic material of high 'crystallinity' has few structural defects, large L_a and L_c values and a low interplanar spacing.

In the Raman spectra of CM, the above-outlined transformation process is reflected by the change in shape (Figure 1A) of the most prominent Raman bands in the first-order spectrum (~ 700–2000 cm^{-1}). Overall, there are at least five Raman bands in the first-order spectrum of CM (Figure 1B). Following Sadezky *et al.* (2005) and Marshall *et al.* (2010), these bands are denominated as D1 (~ 1350 cm^{-1}), D2 (~ 1620 cm^{-1}), D3 (~ 1500 cm^{-1}), D4 (~ 1250 cm^{-1}) and G (~ 1580 cm^{-1}). The G band is assigned to the Raman active E2 g optical phonon in graphite (Tuinstra and Koenig 1970, Reich and Thomsen 2004). The D1- and D2 bands are defect-induced (Pimenta *et al.* 2007) and depend on the excitation energy due to double-resonant Raman scattering (Reich and Thomsen 2004). For more information, see Pócsik *et al.* (1998), Matthews *et al.* (1999), Thomsen and Reich (2000), Saito *et al.* (2002), Reich and Thomsen (2004) and Pimenta *et al.* (2007). The D3 band supposedly originates from amorphous carbons, and D4 band is attributed to sp^2-sp^3 bonds or C-C and C=C stretching vibrations of polyene-like structures (Sadezky *et al.* 2005 and references therein).

Generally, the number of Raman bands decreases from low to high metamorphic conditions (Wopenka and Pasteris 1993, Yui *et al.* 1996, Beyssac *et al.* 2002b, see Figure 1A). As the performed mode of spectral curve-fitting changes with 'crystallinity level', the recorded spectrum is first evaluated 'by eye' by a rough qualitative classification (Figure 1):

'Crystallinity level 1': This level describes poorly crystalline CM that exhibits a rather complex spectrum in which two broad, overlapping Raman bands at ~ 1350 cm^{-1} (D1) and ~ 1580 to 1600 cm^{-1} (G + D2) and a third band at ~ 1250 cm^{-1} (D4) as shoulder on the 1350 cm^{-1} band are present.

'Crystallinity level 2': This level describes moderately to well crystalline CM. Here, the spectra are less complex, as the band at ~ 1250 cm^{-1} (D4) is absent. The intensities of the 1350 cm^{-1} band (D1) and the overlapping region between 1350 cm^{-1} and 1580 cm^{-1} are decreasing, while the 1580–1600 cm^{-1} band (G) gets more intense and narrow. Moreover, a new band at ~ 1620 cm^{-1} (D2) appears as a clear shoulder.

'Crystallinity level 3': This level describes well crystalline CM and graphite. The spectra are simple with only the 1350 cm^{-1} and 1580 cm^{-1} bands present. The 1350 cm^{-1} band is broad and of low intensity, while the 1580 cm^{-1} band is intense and sharp (low FWHM). In case of pure graphite, only the G band appears (Tuinstra and Koenig 1970).

Spectral evaluation

Before fitting the first-order Raman spectrum of CM, a background correction is essential. The background is usually modelled as a linear, polynomial or spline function. The mode of such baseline is crucial for spectral curve-fitting, as all peak parameters are influenced by the baseline function. As manual baseline correction is very susceptible to subjectivity, a linear baseline with two control points is proposed to yield the most reproducible results. To increase the reproducibility of manual baseline correction, the control points, which define the slope of the linear baseline function, are placed in the spectral region of 800–900 cm^{-1} and 1800–1900 cm^{-1} for all 'crystallinity levels'.

In this study, the peak- and curve-fitting software Fityk (Wojdyr 2010; http://fityk.nieto.pl) is used for deconvolution of the Raman spectrum of CM into the different components (D and G bands). However, any other peak-fitting software can be used for this purpose. In our approach, the position and shape of the components are detected automatically by the software. If this is not successful, the components are located manually. Because Voigt- and Lorentzian functions are most commonly used in RSCM thermometry, all components are modelled here as Voigt- or Lorentzian functions with unfixed peak parameters (FWHM, height, position, area

and shape). For 'crystallinity level 1', five components (D1, D2, D3, D4, G) result in a good fit (e.g., Sadezky *et al.* 2005, Lahfid *et al.* 2010, see Figure 1B). For 'crystallinity level 2', a good solution is obtained with 3–4 components (D1, D2, G, [D3]; see Beyssac *et al.* 2002a), and for 'crystallinity level 3', only two components (D1, G) are needed. The components are assigned sequentially to the model, which is also sequentially fitted to the data with by the Levenberg–Marquardt method (Moré 1978). If components are displaced during fitting, take unlikely shapes or are in any other way inconsistent, the solution is rejected. Subsequently, the component parameters have to be changed and the model has to be solved again. This procedure is repeated until a satisfying fit is obtained. The complete fitting protocol is available in the Appendix S1.

Once the parameters of all components are obtained, different ratios can be calculated which correlate with the maximum metamorphic temperature (Beyssac *et al.* 2002a, Rantitsch *et al.* 2004, Rahl *et al.* 2005, Baziotis *et al.* 2006, Aoya *et al.* 2010, Lahfid *et al.* 2010). The most common are the R1- and R2 ratio (Beyssac *et al.* 2002a, Equations 1 and 2) and the RA1- and RA2 ratio (Lahfid *et al.* 2010, Equations 3 and 4). The R1 ratio corresponds to the height of the D1 component divided by the height of the G component. The integrated area of the D1 component divided by the sum of the integrated areas of the D1-, D2- and G component gives the R2 ratio. The sum of the integrated areas of the D1- and D4 component divided by the sum of the integrated areas of the D1-, D2-, D3-, D4- and G components forms the RA1 ratio, and the sum of the integrated areas of the D1- and D4 component divided by the sum of the integrated areas of the D2-, D3- and G component is the RA2 ratio.

$$R1 = \left(\frac{D1}{G}\right)_{\text{Intensity}} \quad (1)$$

$$R2 = \left(\frac{D1}{(D1 + D2 + D3)}\right)_{\text{Area}} \quad (2)$$

$$\text{RA1} = \left(\frac{(\text{D1} + \text{D4})}{(\text{D1} + \text{D2} + \text{D3} + \text{D4} + \text{G})}\right)_{\text{Area}} \quad (3)$$

$$\text{RA2} = \left(\frac{(D1 + D4)}{(D2 + D3 + G)}\right)_{\text{Area}} \quad (4)$$

Systematic tests of the biasing factors

In order to estimate the impact of different biasing factors, a series of experiments was designed. All experiments focus on selected factors, while other factors are kept constant (see Table 2).

To compare the accuracy of the measurements of the different samples, the percental fraction of the standard deviation of the mean, that is, the relative standard deviation, is used as a comparative index. According to the calibration range of the calibration curves of Beyssac *et al.* (2002a) and Lahfid *et al.* (2010), it is calculated that an increment of 0.01 in the commonly used parameter ratios R2, RA1 and RA2 (Beyssac *et al.* 2002a, Lahfid *et al.* 2010) is equivalent to 4, 12 and 2 °C, respectively. These values estimate the significance of each experiment with respect to the initial calibration uncertainty, which is ± 50 °C for Beyssac *et al.* (2002a).

Spectral processing bias

Test 1 – Influence of curve-fitting strategy on parameter ratios: The factors that influence the fitting are the signal-to-noise ratio, the position and slope of the baseline, type of function used (Voigt, Lorentzian, Gaussian, etc.) and the start position and shape of the inserted peak. Lahfid *et al.* (2010) suggested that CM spectra of low-grade metamorphic rocks ('crystallinity level 1' of this study) should be fitted by Lorentzian functions and not by non-converging Voigt functions. When 'crystallinity level 2 and 3' samples are fitted, Voigt functions should be used according to Beyssac *et al.* (2002a).

In this experiment, the amount of scatter in the parameter ratios due to the fitting procedure is quantified. From the different 'crystallinity levels', single spectra with a high signal-to-noise were repeatedly (twenty times) evaluated, using the above-outlined fitting strategy. To exclude a variation due to the baseline correction, baseline-corrected spectra were used for each repetition. As the system often does not converge with automatically positioned Voigt functions, a strict fitting protocol has been used in which the functions are added manually (see Appendix S1). The same procedure has been performed with Lorentzian functions. Since the signal-to-noise ratio and slope of baseline are constant, the only variable is given by the initial conditions (i.e., position and shape of the components).

This experiment shows that in the 'crystallinity level 1', the fitting strategy has a significant influence on the parameter ratios. The repeated evaluation of the same spectrum results in different RA1- and RA2 ratios when Lorentzian- and Voigt functions are used (Table 3). Additionally, the variation in RA1- and RA2 ratios of the twenty evaluations of the same spectrum is greater when only Voigt functions are used (Table 3).

In the 'crystallinity levels 2 and 3', the choice of the function does not influence the results (Table 3). This is in line

Table 2.
Summary of the different tests. The influencing factors are listed in columns, and the single experiments are shown in rows. The tested factors are emphasised by grey colour, and empty fields indicate irrelevant factors in the given test

Test	Influence of…	Person	Sample (per cryst. level)	Spot	Measurement (per spot)	Type (function) of components	Baseline correction	Software	Sample prep. (chemically enriched CM vs. rock chips)	Orientation	Spectrometer
1	Spectral fitting strategy	1	1	1	1	M	1				
2	Baseline correction	1	1	1	1	1	M				
3	Empirical scatter	1	1	1	M	1	1				
4	Evalutation software	1	M	M	1	1	1	M			
5	Personal fitting strategy	M	M	M	1	1	1	1			
6	Sample preparation	1	M	M	1	1	1	1	M	1	1
7	Sample heterogeneity	1	M	M	M	1	1	1	1	1	1
8	Anisotropy	1	1	M	1	1	1	1	1	M	1
9	Raman system	M	1	M	1	1	1	1	1	1	M

1, single; M, multiple.

Table 3.
Results of test 1: comparison of R1, R2, RA1 and RA2 of the three 'crystallinity levels' fitted with Lorentzian and with Voigt functions

Sample	Crystallinity level 1 – Lorentzian				Crystallinity level 1 – Voigt			
	R1	R2	RA1	RA2	R1	R2	RA1	RA2
L29	1.701	0.671	0.615	1.599	1.374	0.748	0.643	1.803
RSD (%)	0.1	0.0	0.0	0.0	10.2	3.4	1.7	4.6
DB45	1.281	0.641	0.586	1.414	1.200	0.679	0.618	1.619
RSD (%)	0.0	0.0	0.0	0.0	2.3	0.8	0.2	0.6
DB21	1.423	0.612	0.603	1.521	1.391	0.638	0.633	1.721
RSD (%)	0.0	0.0	0.0	0.0	1.1	0.5	0.2	0.6
	Crystallinity level 2 – Lorentzian				Crystallinity level 2 – Voigt			
DB36	0.673	0.477	–	–	0.668	0.477	–	–
RSD (%)	0.0	0.0	–	–	0.0	0.0	–	–
DB28	0.675	0.455	–	–	0.668	0.453	–	–
RSD (%)	0.0	0.0	–	–	0.0	0.0	–	–
L1	0.391	0.371	–	–	0.39	0.371	–	–
RSD (%)	0.0	0.0	–	–	0.0	0.0	–	–
	Crystallinity level 3 – Lorentzian				Crystallinity level 3 – Voigt			
DB16	0.016	0.048	–	–	0.016	0.049	–	–
RSD (%)	0.0	0.0	–	–	0.0	0.1	–	–
DB28	0.019	0.064	–	–	0.019	0.065	–	–
RSD (%)	0.0	0.0	–	–	0.0	0.0	–	–
TU2	0.025	0.073	–	–	0.025	0.073	–	–
RSD (%)	0.0	0.0	–	–	0.0	0.0	–	–

RSD, relative standard deviation.

with previous results by Lahfid *et al.* (2010). However, the Lorentzian and Voigt fits of 'crystallinity levels 2 and 3' samples result in basically the same R1 and R2 ratios. Thus, there seems to be no justification for changing from Lorentzian- to Voigt functions for better crystalline samples. Instead, for the sake of linearity, we suggest to use the same function for processing samples of all 'crystallinity levels'. As the fits with Lorentzian functions perform well for low and high crystalline samples, this is the function type to be used for all the fits.

Test 2 – Influence of baseline correction: The baseline correction significantly influences the results. As stated above, a linear baseline is more reproducible because only two control points are needed to define the line slope. It is tested here to what extend the varying position of the baseline points will influence the results. Three scenarios were tested by a Python (Pilgrim 2004) script using the extensions Numpy and Scipy (Ascher *et al.* 2001, Oliphant 2007): (1) the low wavenumber control point is fixed, while the high wavenumber control point is moved. (2) Is the inversion of (1), (3) both control points are moved towards each other with the same increment. Every 10th data point was chosen as a new control point for baseline correction. Obviously, this procedure is only valid if the background can be approximated by a linear baseline. If the background is curved by fluorescence, the background cannot be described by a linear function and this procedure fails.

In every experiment, ten baselines were calculated. As the spectral resolution is ~ 2 cm^{-1}, the moving control point was varied in a range of about 200 wavenumbers (Figure 2A–C).

Another frequently used type of baseline correction is represented by polynomial and spline functions, both result in curved baselines. To account for this, another script was used in which the first and last n data points, choosing n = {10, 20,..., 100}, were selected for ten different spline interpolations (Figure 2D). To compensate for noise effects, the spectra were smoothed by using a penalised least squares approach (Eilers (2003). For reasons given above (see Test 1), all fits were performed using Lorentzian functions. For 'crystallinity levels 1 and 2', the experiments demonstrate that the model uncertainty of a linear baseline results in a ratio scatter below 1% RSD (relative standard deviation, Table 4). Only R2 of 'crystallinity level 3' shows a greater scatter. The curved baseline leads to an increased variation in RA1 and RA2 for 'crystallinity level 1' samples. For 'crystallinity level 2', the variation is

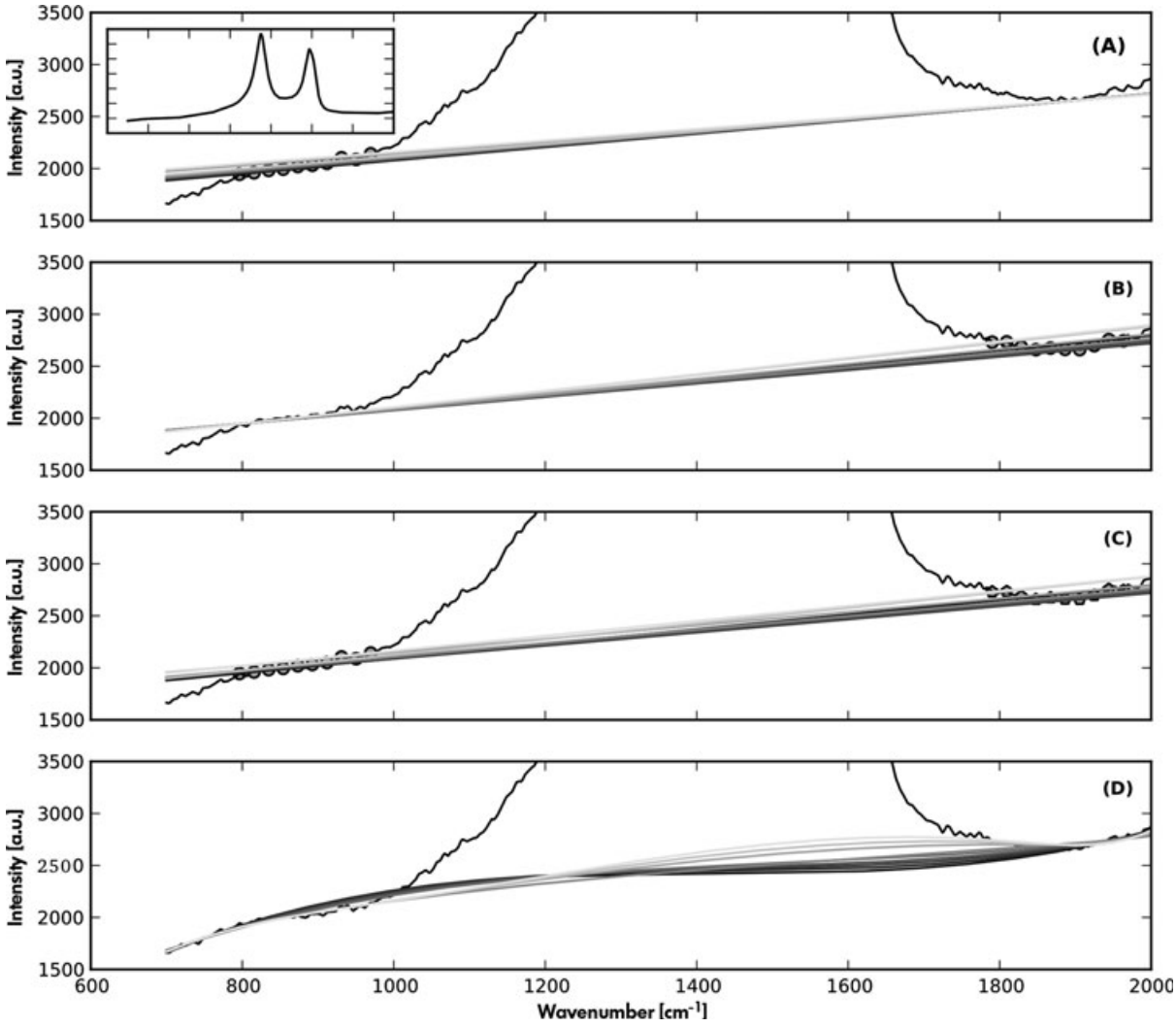

Figure 2. Results of test 2; the baseline correction is tested by varying the baseline slope and baseline type for a 'crystallinity level 1' spectrum. (A) The low wavenumber side control point (circle) is variable. (B) The high wavenumber side control point is variable. (C) Both control points are variable. (D) The same spectrum as in A–C but with curved baselines derived from spline interpolation. a.u., arbitrary units. The inset in (A) shows the complete spectrum, which has been used in A–D.

negligible, and for 'crystallinity level 3', the RSD of R2 is again high (Table 4).

Consequently, the reproducibility of the RA1 and RA2 ratios in 'crystallinity level 1' is higher when a linear baseline is used. Also, the positioning of control points within the same wavenumber intervals is of high importance.

Translating the R2 and RA1 ratios of the curved baseline-corrected spectra into a metamorphic temperature using the equations given in Beyssac *et al.* (2002a) and Lahfid *et al.* (2010), the resulting temperature variation is about 5–10 °C for level 3 samples and about 20–30 °C for level 1 samples. For linear baseline-corrected samples, the temperature variation of R2 of level 3 samples is < 5 °C and for RA1 of level 1 samples *ca.* 10 °C.

Test 3 – Reproducibility (repeated analysis of the same spots): In this test series, the total empirical scatter generated by the Raman system (i.e., minor variations in beam intensity, dispersion effects in the optical path, etc.) is examined. Ten spectra, each representing the mean of two acquisitions of 30 s, were recorded successively on the same spots, and the evaluations were made strictly in the same way (using linear baseline and applying the evaluation protocol, see sections Test 1 – Influence of curve-fitting

Table 4.
Results of test 2: influence of baseline correction. R1, R2, RA1 and RA2 ratios for the three 'crystallinity levels' with varying linear baselines and spline interpolated, curved baselines

	Cryst. Lvl 1				Cryst. Lvl 2		Cryst. Lvl 3	
	R1	R2	RA1	RA2	R1	R2	R1	R2
Low wavenumber side control point variable								
Average	1.432	0.614	0.602	1.512	0.597	0.437	0.016	0.034
RSD (%)	0.2	0.1	0.4	1.0	0.1	0.1	0.6	6.3
High wavenumber side control point variable								
Average	1.440	0.616	0.607	1.542	0.596	0.437	0.016	0.038
RSD (%)	0.6	0.5	0.3	0.8	0.1	0.2	0.4	8.8
Both control points variable								
Average	1.441	0.617	0.605	1.530	0.597	0.437	0.016	0.036
RSD (%)	0.6	0.5	0.3	0.7	0.1	0.2	0.6	12.1
Curved baseline – Spline interpolation								
Average	1.435	0.613	0.595	1.469	0.595	0.435	0.016	0.045
RSD (%)	0.9	1.0	2.1	5.2	0.0	0.2	0.7	14.4

Cryst. Lvl, 'Crystallinity level', RSD, relative standard deviation.

strategy on parameter ratios and Test 2 – Influence of baseline correction).

From Figure 3, it is obvious that the behaviour of 'crystallinity level 1' differs from the more ordered material. The Raman parameters of the less ordered organic material are highly sensitive to the absorbed laser energy. The RA1 and RA2 ratios of 'crystallinity level 1' decrease during the first three measurements; after those, the data scatter randomly. This phenomenon was already detected by Quirico *et al.* (2005) who suggested that most likely photo-oxidation of the sample controls the fluorescence and Raman signal of the sample, even if very low energy is applied. Additionally, they demonstrated that the signal (fluorescence + Raman) stabilises faster when measured in an inert (Argon) atmosphere. However, the changing fluorescence cannot be ruled out completely (Quirico *et al.* 2005).

The time-dependent change in RA1- and RA2 ratios of 'crystallinity level 1' samples demonstrates that the Raman analysis of such organic material needs special care. Even in case of low laser energy, the possible transformation of the organic material or the declining fluorescence change the signal by time, resulting in highly affected results by repeated measurements on the same spot. Therefore, the empirical scatter of 'crystallinity level 1' material cannot be determined by repeated analyses on the same spot.

In level 2 and 3 samples, the parameter ratios (R1, R2) do not change significantly with time (Figure 3). Thus, the variability of the fitting results is a reliable measure of the reproducibility by repeated analyses. However, the scatter generated by the spectral decomposition (incl. baseline correction) is included and forms an indistinguishable part of the total scatter. Thus, the left panel of Table 5 shows the scatter of spectral decomposition (Tests 1 and 2), while the right panel shows the total empirical scatter. The lesson of this experiment is that the evaluation of level 3 is extremely sensitive to minor modifications in the fitting procedure (for instance baseline correction).

This artificial uncertainty, which is introduced by spectral processing, is in the range of *ca.* 20% RSD of R2. Translating the calculated R2 ratios into temperature (Beyssac *et al.* 2002a) results in a spread of *ca.* 10 °C for the same measuring spot and adds to the calibration uncertainty of the RSCM geothermometer.

Test 4 – Influence of spectral evaluation software on parameter ratios: In this experiment, thirteen samples from low to high 'crystallinity' were measured at five random spots holding the measuring conditions and the background correction constant. The spectra were evaluated by three different computer programs: Fityk (Wojdyr 2010), Labspec 5 (Horiba Scientific, Edison, New Jersey 08820, USA) and Peakfit (SeaSolve Software, Inc., Framingham, Massachusetts 01702, USA). Thus, differences in R1, R2, RA1 and RA2 ratios are generated exclusively by the used evaluation software and applied fitting strategy. The contribution to the scatter from sample heterogeneity can be neglected, because identical spectra were evaluated.

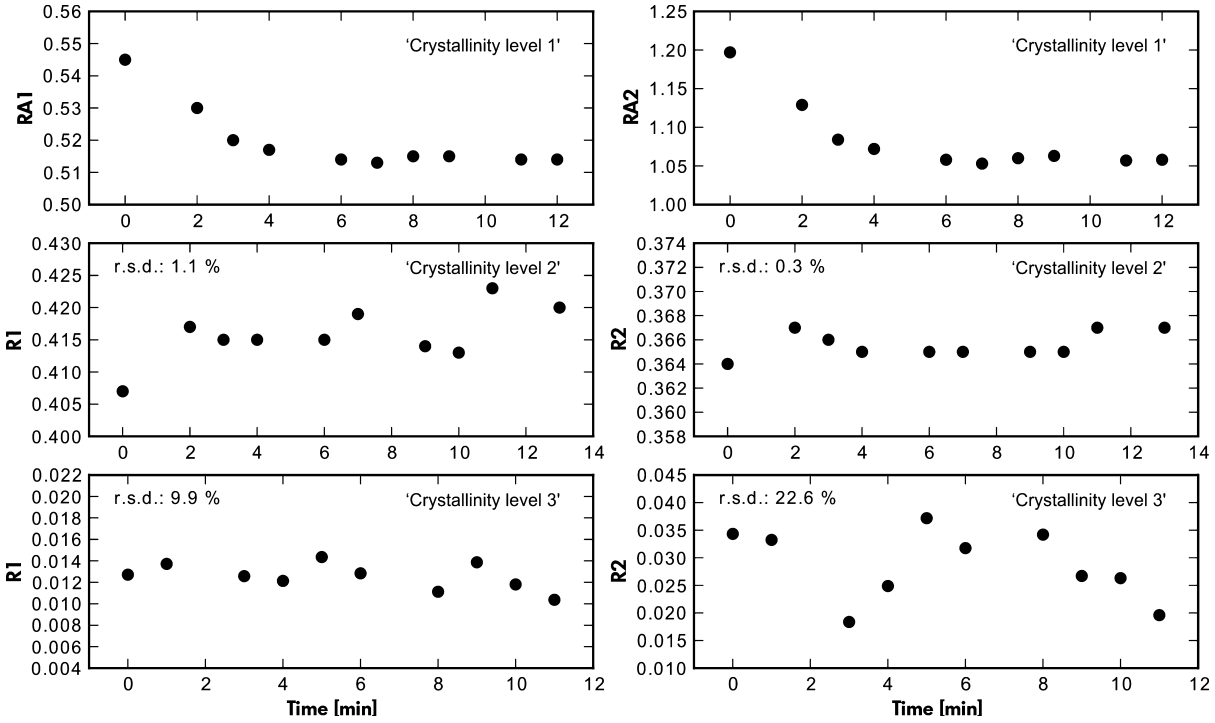

Figure 3. Results of test 3; the reproducibility by repeated analysis. The subplots show time series of the RA1, RA2, R1 and R2 ratios (according to Beyssac *et al.* 2002a, Lahfid *et al.* 2010) for samples of the three 'crystallinity levels'. All measurements were performed consecutively at the same spot. The first three measurements of 'crystallinity level 1' are influenced by luminescence that changes strongly with time, but afterwards, the detected ratios are basically stable.

Table 5.
Results of test 3: Empirical scatter of repeated measurements

	Scatter generated by spectral decomposition with linear baseline (see Test 2)		Scatter generated by repeated analysis of the same spot (with linear baseline correction)	
	Cryst. Lvl. 2	Cryst. Lvl. 3	Cryst. Lvl. 2	Cryst. Lvl. 3
R1 RSD (%)	0.1	0.6	1.1	9.9
R2 RSD (%)	0.2	12.1	0.3	22.6

From Figure 4, it is evident that there are essentially no differences in R1 and R2 for the whole sample set. In case of RA1 and RA2, Fityk generates systematically higher ratios than Labspec and Peakfit; however, the differences are well within one standard deviation (except for sample L29).

Test 5 – Influence of operator's personal fitting strategy on parameter ratios: In this test, the personal bias in curve-fitting is analysed. Therefore, the same five spectra per sample obtained from DB21, DB36 and L45 were handed to four different operators. Spectral processing was first conducted by the operators without any recommendation, and afterwards, the same spectra were processed by the operators according to the recommended fitting protocol (see Appendix S1). The spectral processing was performed once with Lorentzian functions only and once with Voigt functions only. That is, for each sample, five Lorentzian and five Voigt R1, R2, RA1 and RA2 values were derived by each operator. The five values of a given ratio were summed, and then, the absolute difference of the sums of two operators for a given sample was used for comparison. This approach is valid since identical spectra were evaluated, and if the operators produced the same ratios, the absolute difference between these operators must be zero. It can be deduced from Table 6 that there are variable and sometimes significant differences between the operators. The summed absolute differences are reduced by the spectral processing protocol when only Voigt functions are used (Table 6). If only Lorentzian functions are used a slight increase in the differences can be observed. This is most likely due to differences in the baseline correction performed by the operators, because if the type of function is

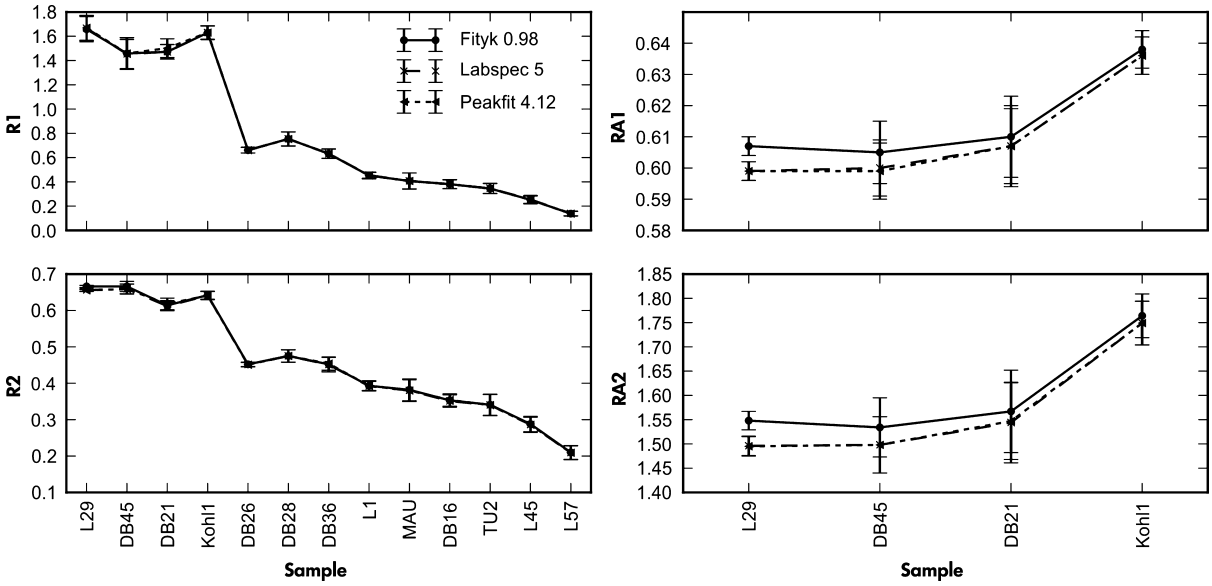

Figure 4. Results of test 4; the influence of spectral evaluation software. Comparison of the R1, R2, RA1 and RA2 ratios (Beyssac *et al.* 2002a, Lahfid *et al.* 2010) determined with different programmes. The sample 'crystallinity' increases from left to right. The error bars indicate one standard deviation of five measurements.

Table 6.
The subtables A and B show the sums of the absolute differences for the R1, R2, RA1 and RA2 ratios between the four operators

Compared operators	**Sum of differences – "crystallinity" level 1 – Voigt before protocol**		**Sum of differences – "crystallinity" level 1 – Voigt after protocol**		**Sum of differences – "crystallinity" level 1 – Lorentz before protocol**		**Sum of differences – "crystallinity" level 1 – Lorentz after protocol**	
	RA1	***RA2***	***RA1***	***RA2***	***RA1***	***RA2***	***RA1***	***RA2***
(A)								
1 – 2	NA	NA	*0.124*	*0.994*	NA	NA	*0.021*	*0.142*
1 – 3	0.056	0.402	0.005	0.019	0.041	0.257	0.041	0.275
1 – 4	0.090	0.610	0.021	0.153	0.030	0.190	0.018	0.116
2 – 3	NA	NA	*0.129*	*0.975*	NA	NA	*0.020*	*0.133*
2 – 4	NA	NA	*0.145*	*1.148*	NA	NA	*0.039*	*0.258*
3 – 4	0.033	0.208	0.016	0.172	0.012	0.080	0.059	0.391
Sum (excluding NA and italics)	0.180	1.220	0.042	0.344	0.083	0.527	0.119	0.782
Compared operators	**Sum of differences – "crystallinity" level 2 – Voigt before protocol**		**Sum of differences – "crystallinity" level 2 – Voigt after protocol**		**Sum of differences – "crystallinity" level 2 – Lorentz before protocol**		**Sum of differences – "crystallinity" level 2 – Lorentz after protocol**	
	R1	***R2***	***R1***	***R2***	***R1***	***R2***	***R1***	***R2***
(B)								
1 – 2	0.047	0.185	0.002	0.001	NA	NA	*0.025*	*0.064*
1 – 3	0.075	0.007	0.033	0.101	0.031	0.061	0.072	0.080
1 – 4	0.064	0.024	0.036	0.006	0.024	0.064	0.041	0.033
2 – 3	0.028	0.192	0.036	0.101	NA	NA	*0.047*	*0.016*
2 – 4	0.017	0.161	0.034	0.007	NA	NA	*0.066*	*0.031*
3 – 4	0.011	0.031	0.070	0.094	0.008	0.003	0.114	0.047
Sum (excluding NA and italics)	0.243	0.600	0.211	0.310	0.062	0.129	0.227	0.159

In each row, the ratios of two operators are compared. If the sum of an operator is not available, the values (printed in italics) of the compared operator are not used in the total sum.
NA, not available.

the only variable, no change in the parameter ratios is observed in the case of Lorentzian functions, which is shown in section Test 1 – Influence of curve-fitting strategy on parameter ratios.

Especially Figure 5 shows that the differences between the operators can be significant (see insert A in Figure 5). That is for a given spectrum, the difference in RA1 can be as high as 0.05 (Figure 6), which would translate to *ca.* 60 °C. The differences in R2 of 'crystallinity level 2' samples between the operators are in the range of about 0.01 to 0.03 (Figure 5), which corresponds to a temperature range of *ca.* 4 °C to 12 °C.

These observations imply that the way of spectral processing is a key feature in the achievement of parameter ratios and should be defined very precisely. At the same time, it is obvious that the used protocol needs refinement which is documented, for instance by the increase in the differences for Lorentzian functions.

Sources of biasing factors intrinsic to the carbonaceous material

Test 6 – Influence of sample preparation on parameter ratios: The goal of this test is to check if there are deviations in the parameter ratios due to sample preparation. Untreated rock fragments and HCl-HF-digestion-enriched CM were compared. The sample set of the Thuringian Forest (Table 1) was selected for this test. These black shales experienced subgreenschist facies metamorphism (Kunert 1999). Therefore, the curve-fitting method after

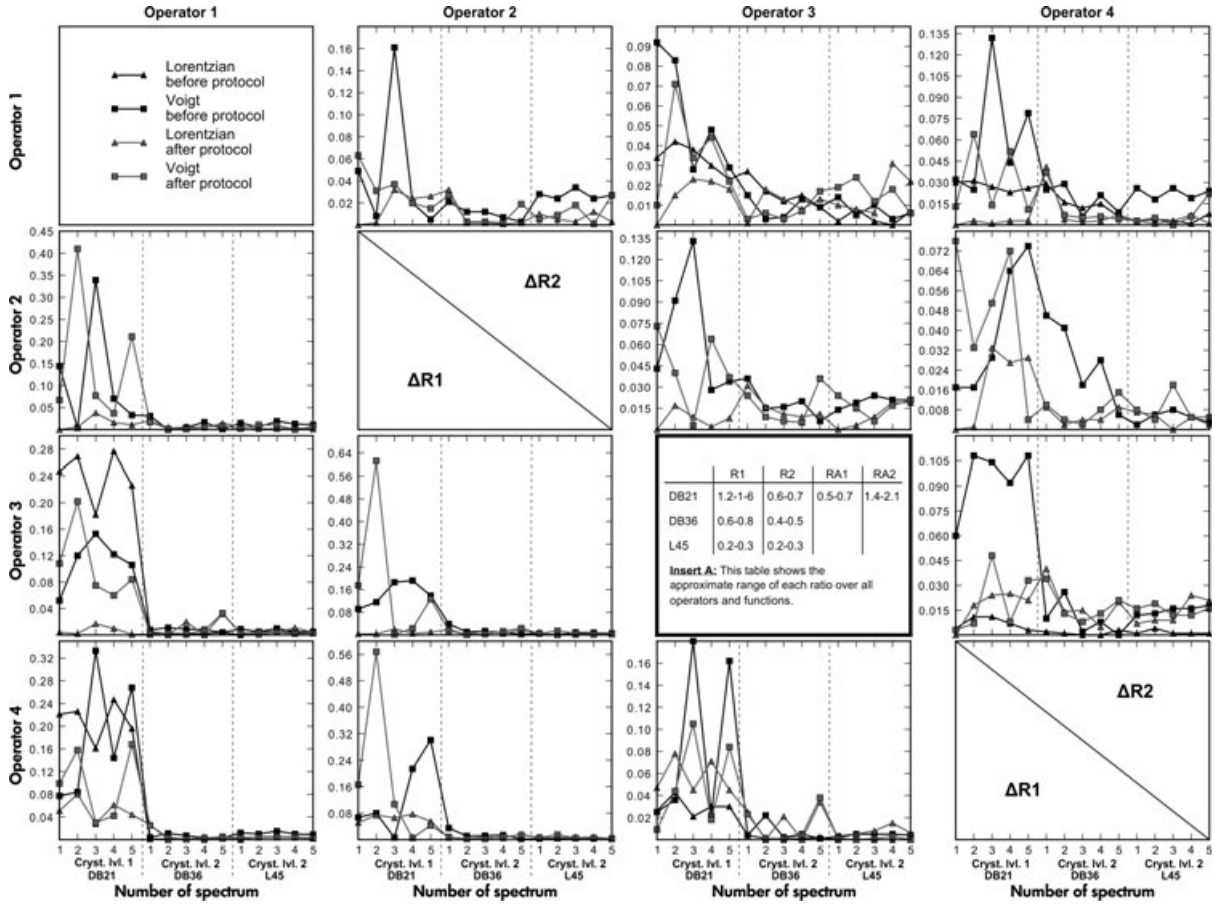

Figure 5. Results of test 5; this matrix plot shows the influence of operator's personal fitting strategy. The lower left triangle of the matrix shows the absolute differences in R1 (termed ΔR1), and the upper right triangle shows the absolute differences in R2 (termed ΔR2) between the operators (e.g., column one gives ΔR1 between operator 1 and the others, while row one gives ΔR2 between operator 1 and the others). Each plot in the matrix represents the absolute differences between two operators for the given 15 spectra, which are listed along each x-axis. In most cases, especially for 'crystallinity level 1', the differences are reduced by the protocol. Insert A shows the range of the ratios in order to the estimation of the impact of the personal bias.

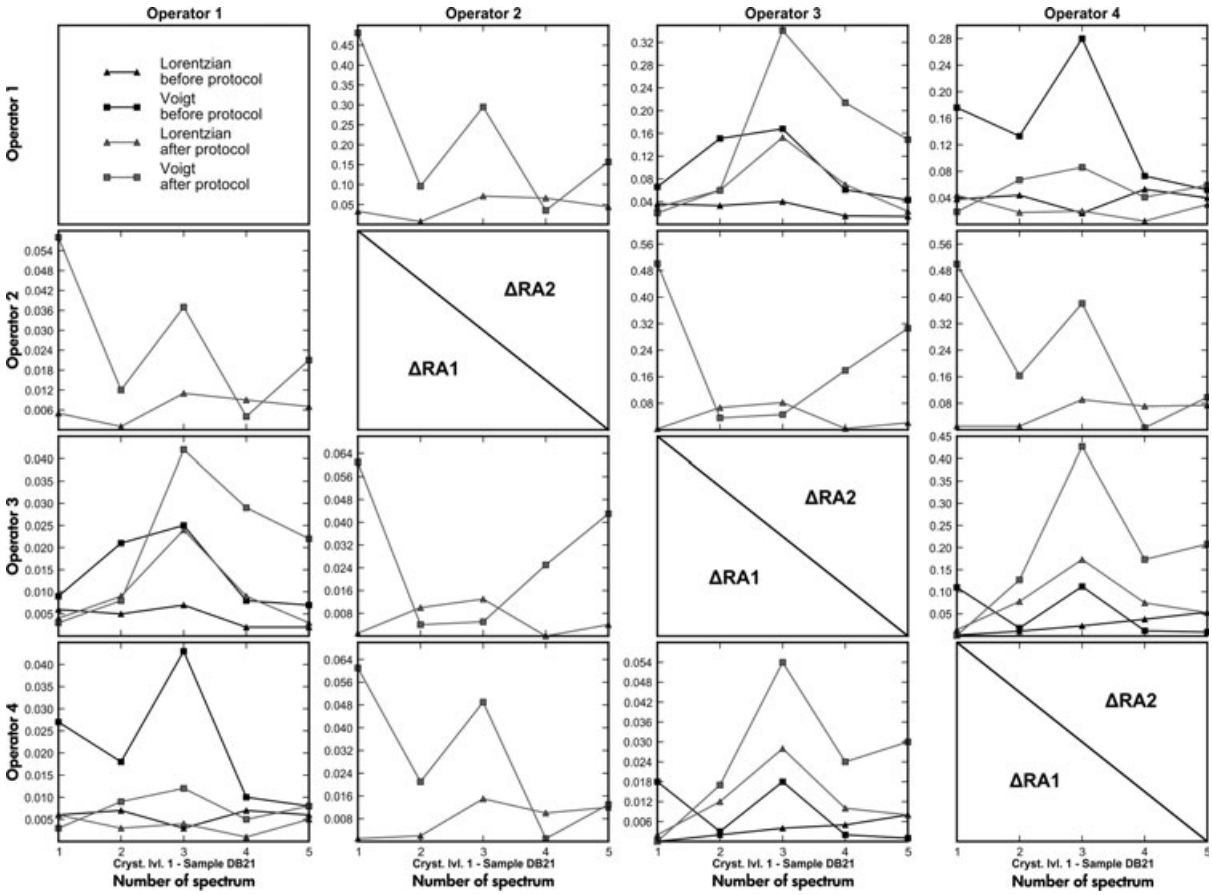

Figure 6. Results of test 5; this matrix plot shows the influence of operator's personal fitting strategy. The lower left triangle of the matrix shows the absolute differences in RA1 (termed ΔRA1), and the upper right triangle shows the absolute differences in RA2 (termed ΔRA2) between the operators (e.g., column one gives ΔRA1 between operator 1 and the others, while row one gives ΔRA2 between operator 1 and the others). Each plot in the matrix represents the absolute differences between two operators for the given five spectra of a 'crystallinity' level 1 sample. The results are ambiguous as in some cases the differences are reduced and in other cases increased by the use of the protocol. The range of the ratios is given in insert A in Figure 5.

Lahfid *et al.* (2010) with five Lorentzian components was used.

Each sample was measured directly as rock fragment with a laser orientation perpendicular to the schistosity. This was done because it is assumed that the typically planar CM particles settle parallel to the glass slide surface, resulting in a laser orientation more or less perpendicular to the CM.

From Figure 7, it is evident that HF samples produced higher RA ratios, that is, RA1 is about 0.01 and RA2 is about 0.1 higher, which translates to temperature differences of *ca.* 13 °C and 22 °C, respectively. Furthermore, the standard deviations of the HF samples are slightly smaller. Even if the surface of the rock fragments follows the schistosity, it still has a given roughness which might affect the orientation of the CM particles to the laser beam and thus explain a larger spread of parameter ratios. Furthermore, it is speculated that the systematically higher, less scattered RA ratios of the HF samples might be due to reduction in functional groups or lower clay mineral content due to hydrofluoric acid treatment. It has been shown by Villanueva *et al.* (2008) that 'clay' introduces a strong Raman background and Ahn *et al.* (1999) showed that carbonaceous material can be interstratified within illite. Thus, HF-treatment is thought to minimise strong Raman background from CM interconnected with clay minerals.

Test 7 – Influence of sample heterogeneity on parameter ratios: Sample heterogeneity in terms of RSCM is reflected in the variability of a given parameter ratio.

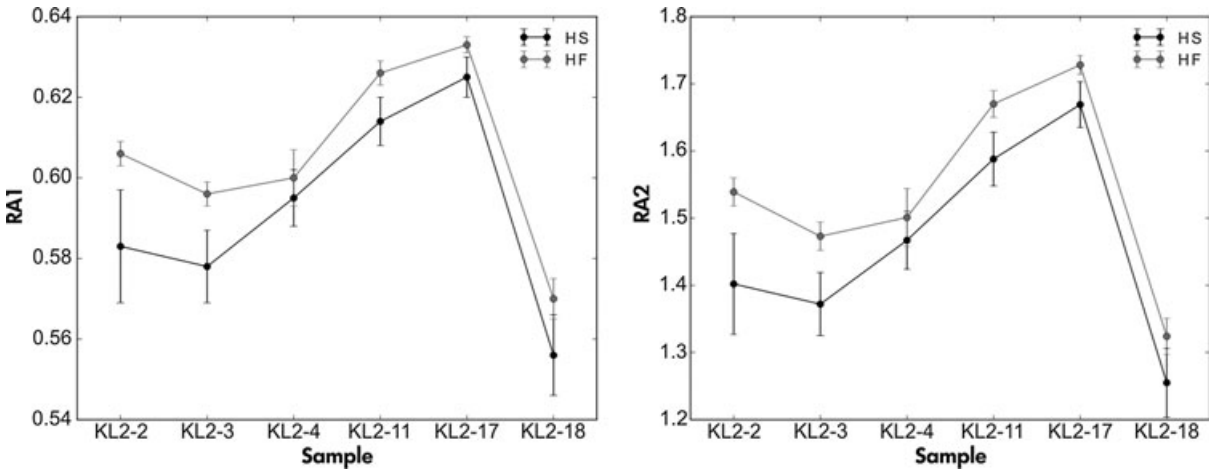

Figure 7. Results of test 6; the influence of sample preparation. The RA1 and RA2 ratios of the hydrofluoric acid-treated samples (HF) are systematically higher and have a lower standard deviation than the ratios of the hand specimen (HS).

Supposedly, the different chemical composition of the precursor organic materials yields different Raman spectra (Kribek *et al.* 1994, Valentim *et al.* 2004, Guedes *et al.* 2010). Considering that the organic content of a given sedimentary rock is a mixture of two or more maceral groups (Hutton *et al.* 1994), a spread of parameter ratios will result. This has a greater impact on parameter ratios of CM dispersed in low-grade metamorphic rocks. Furthermore, recycled detrital organic grains can carry signals of a higher degree of organic maturation from the sediment source areas. Therefore, sample heterogeneity is combining various factors and has to be considered.

As shown in Test 1, the scatter due to spectral evaluation is basically excluded if the fitting is performed with Lorentzian functions only and scatter due to variation of the linear baseline is low (Tables 3 and 4) for 'crystallinity levels 1 and 2'. Thus, the scatter resulting from multigrain analysis is mainly inherited by the sample heterogeneity itself. To estimate the sample heterogeneity, 30 randomly chosen sample spots were analysed in four HF-HCl-treated CM samples of 'crystallinity level 2'. For each sample spot, five successive spectra (each the average of five accumulations with a duration of ten seconds) were recorded to estimate the total empirical scatter.

The mean square of weighted deviates (MSWD) was used to quantify sample heterogeneity. The MSWD is widely used in geochronology to determine whether the observed scatter in a data set is consistent with the calculated analytical uncertainty of the individual measurements (Powell *et al.* 2002). It is a measure of the ratio of the observed scatter of the data points (t_i) around the mean value (t_m) to the expected scatter from the assigned errors (σ_i) (Vermeesch 2010). If the assigned errors are the only cause of scatter, the MSWD will tend towards unity and *'MSWD values much greater than unity generally indicate either underestimated analytical errors, or the presence of non-analytical scatter'* (Ludwig 2003).

$$\mathrm{MSWD} = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(t_i - t_m)^2}{\sigma_i^2} \qquad (5)$$

By solving Equation (5) where t_i is the arithmetic mean of the repeated measurements in one grain, t_m is the mean of all grains, and σ_i is the standard deviation of the data obtained in one grain, it is estimated to which extend the overall scatter is influenced by the empirical scatter. Figure 8 shows the R1 and R2 ratios plotted in measuring order and their associated MSWD. MSWD values are remarkably high (MSWD = 47–318, Figure 8), indicating the presence of non-analytical scatter in all samples, that is, thought to be mainly caused by sample heterogeneity.

In several acid-treated CM samples, the heterogeneity is already well detectable by reflected light microscopy. Frequently, two particle types of CM are present. "Large particulate CM" is formed by particles with a grain size range of *ca.* 20–100 μm and a bright grey to white reflective light appearance (Figure 9). In contrast, 'disperse CM' has a grain size of less than 10 μm and is of dark grey colour in reflected light (Figure 9).

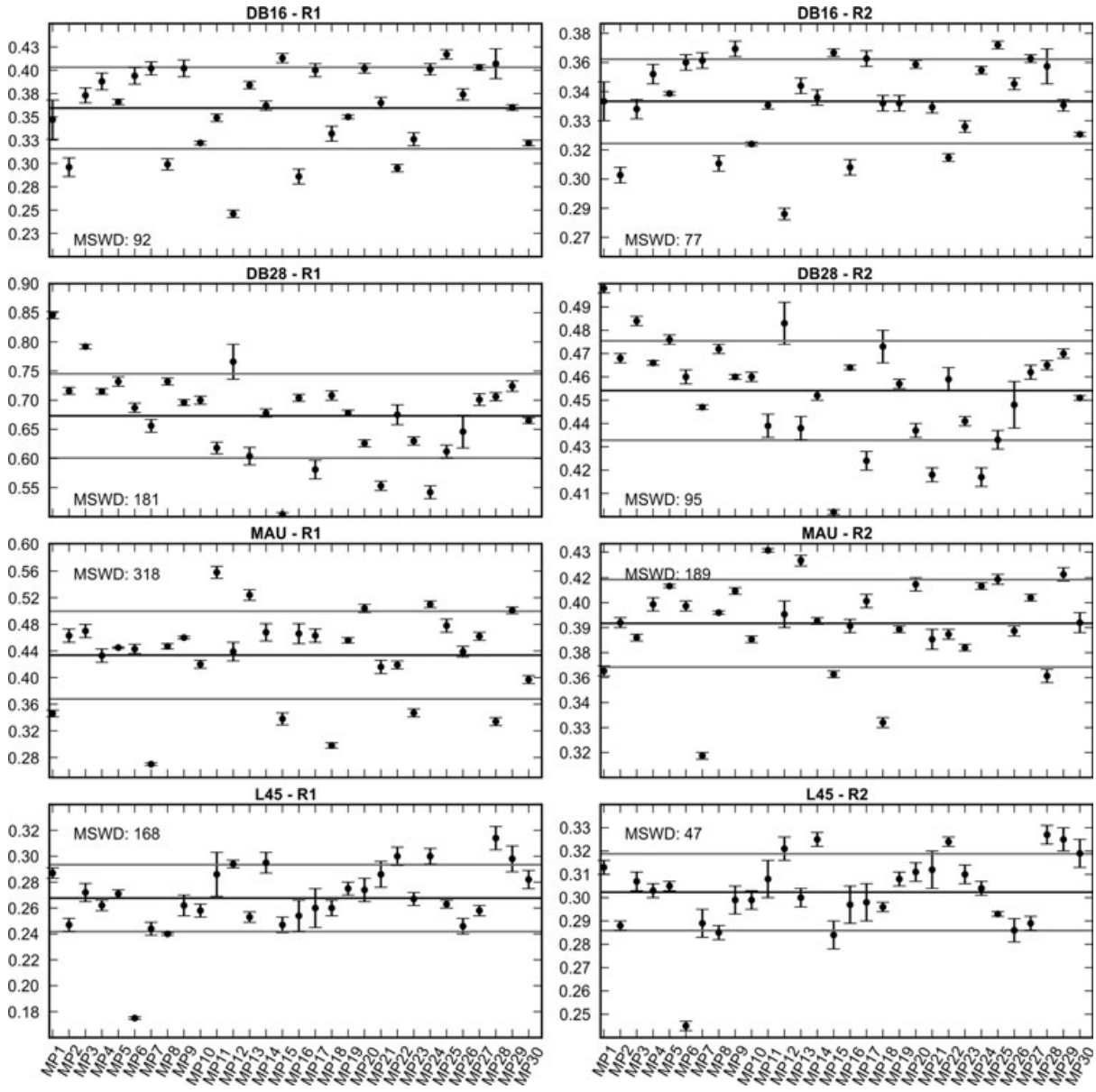

Figure 8. Results of test 7; the influence of sample heterogeneity. The mean square weighted deviation (MSWD, according to Vermeesch 2010) has been calculated for four samples in the 'crystallinity level 2'. Five spectra were recorded at each of the thirty randomly chosen measuring points (MP). If the value of the MSWD is high, then the overall scatter of the points is dominated by the sample heterogeneity. If the MSWD value is low, then uncertainty of the individual measuring points represents well the overall scatter of the population. The solid black line represents the average of all measurements of one sample, and the solid grey lines are the associated positive and negative standard deviation. Error bars represent the standard deviation of individual measuring points.

Table 7 shows the R1 and R2 ratios with associated relative standard deviation of large particulate and disperse CM for the entire sample set. It is obvious that the Raman parameter ratios of the large particulate CM differ strongly from the disperse CM of the same sample, and the scatter in R1 and R2 of large particulate CM is much higher. A similar observation is made in the rock fragments of the Thuringian black shales comparing coarse organoclasts and dispersed CM; however, the differences of the parameter ratios are minor (*ca.* 0.01 in RA1).

These observations imply that one has to be aware that already matured, detrital CM populations may be present in a sample, which is also an important issue in vitrinite

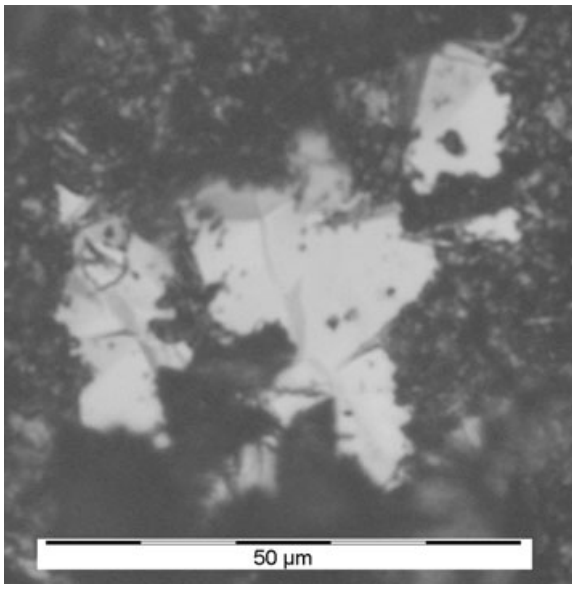

Figure 9. Photomicrograph showing large particulate carbonaceous material (CM; light grey) surrounded by fine-grained disperse CM (dark grey) in HCl-HF concentrated organic material deposited on a glass slide.

reflectance measurements (e.g., Nzoussi-Mbassani *et al.* 2005). Because of the irreversibility of the thermal maturation of CM, the population of lowest maturity is usually chosen (Nzoussi-Mbassani *et al.* 2005). In this case, a proper number of observations (Raman spectra) are necessary to identify the different populations. Aoya *et al.* (2010) proposed a minimum of *ca.* 25 measurements per samples to reflect sample heterogeneity.

Test 8 – Influence of structural anisotropy on parameter ratios: Due to its strong structural anisotropy, Raman spectra of graphite and very well crystalline CM depend on the orientation of the laser beam to the crystallographic *c*-axis (Katagiri *et al.* 1988, Wang *et al.* 1989, Compagnini *et al.* 1997). Katagiri *et al.* (1988) showed that the intensity ratio of the D1- and G band is minimal when the incident laser beam is parallel to the *c*-axis of the graphitic material, and the ratio is at its maximum when the laser beam is perpendicular to the *c*-axis.

Natural graphite occurs usually as microcrystallites of a few 100 nm length (Beyssac and Lazzeri 2012). The optical *c*-axis of microcrystallites enclosed in rocks is generally orientated perpendicular to the main fabric (Beyssac and Lazzeri 2012). To quantify the effect of coherent microcrystallite orientation on the R1 and R2 ratios in natural samples, a sample was rotated directly under the laser beam.

To check if there are systematic deviations due to polarisation effects, a spectrum with linear and a spectrum with circular polarised laser light were recorded at each measuring point. From a sample (DB26) with well-developed schistosity, a rock cylinder of 6 mm diameter was drilled parallel to the schistosity. The cylinder was then mounted into a sample holder and rotated by an increment of 20° underneath the laser beam. Thus, the orientation of the CM with respect to the laser beam is varied, and the change in R1 and R2 ratios is observed as a function of schistosity orientation.

In this experiment, the generation of structural defects in the CM by the drilling of the rock core has to be considered. However, Crespo *et al.* (2006) showed that the graphitic structure is insignificantly changed by short grinding periods (< 10 min). We therefore assume that the water-cooled,

Table 7.
Comparison of R1 and R2 ratios of disperse and large particulate CM (five measurements per sample)

Sample	Large particulate CM				Disperse CM			
	R1	RSD (%)	R2	RSD (%)	R1	RSD (%)	R2	RSD (%)
L29	–	–	–	–	0.96	7	0.60	16
DB45	0.01	137	0.06	143	1.22	7	0.66	9
DB21	0.34	104	0.42	48	1.30	5	0.64	3
Kohl1	1.42	3	0.63	0	1.67	8	0.68	3
DB26	0.18	63	0.25	49	0.78	4	0.50	3
DB28	0.19	62	0.26	58	0.61	26	0.44	14
DB36	0.01	100	0.02	106	0.62	8	0.45	4
MAU	0.22	33	0.26	17	0.36	19	0.34	11
L1	0.35	18	0.34	8	0.46	53	0.39	24
DB16	0.15	51	0.20	46	0.35	14	0.33	8
TU2	0.15	115	0.18	79	0.27	8	0.29	6
L45	0.08	94	0.14	59	0.21	13	0.25	8
L57	0.03	79	0.08	50	0.12	13	0.18	11

CM, carbonaceous material; RSD, relative standard deviation.

gentle drilling process applied here has generated insignificant modification of CM on the cylinder surface of the drilled rock core.

Measurements were performed across five sections along the sample cylinder, and the mean of R1 and R2 of the five measurements for each 20°-step was calculated (Figure 10). It is obvious that the R1 and R2 ratios increase when the CM is rotated from 0° to 90° (i.e., from a perpendicular orientation to a parallel orientation). This indicates a virtual decrease in 'crystallinity'. The expected increase from 0 to 90° is followed by a decrease in R1 and R2 from about 100 to 180° which implies a better virtual 'crystallinity' which is again the result of a perpendicular orientation of the CM to the laser beam. Also, the polarisation of the laser light does not lead to systematic differences in the R1 and R2 ratios (Figure 10). However, due to a high standard deviation (Figure 10), the significance of this preliminary test is low.

Consequently, it is deduced that sample anisotropy varies the R1 ratio by ~ 0.1 and the R2 ratio by ~ 0.04 (*ca.* 15 °C). Of course, this is only valid for the sample tested here. Nevertheless, this test points out that sample orientation should be kept in mind when performing Raman measurements on higher crystalline carbonaceous materials.

Sources of variation intrinsic to the experimental design and Raman system

Test 9 – Influence of Raman systems on parameter ratios (interlaboratory comparison): As many different configurations of Raman systems exist, a small-scale interlaboratory comparison was performed to elucidate the variation in results (i.e., R1, R2) with respect to the Raman instruments and their major components, for instance the spectral grating.

The effect of different gratings of the Horiba Jobin Yvon LabRAM HR800-UV spectrometer in Göttingen on the Raman spectrum of sample DB26 is shown in Figure 11. For comparison, the intensities are normalised to the height of G band. It is demonstrated that a finer grating generates a more intense D1 band, which leads to differences in the R1 and R2 ratios. Spectrum (a) in Figure 11 gives 0.52 and 0.39 while spectrum (b) gives 0.66 and 0.44 R1 and R2 ratios, respectively. This indicates, if a linear relation between parameter ratio and grating is assumed, that finer gratings produce higher ratios and that a difference in grating of 1800 l mm^{-1} generates a shift in R1 and R2 of 0.14 and 0.05, corresponding to a temperature range of up to 20 °C.

For the interlaboratory comparison, the samples DB16, DB26 and DB36 were measured with three external Raman

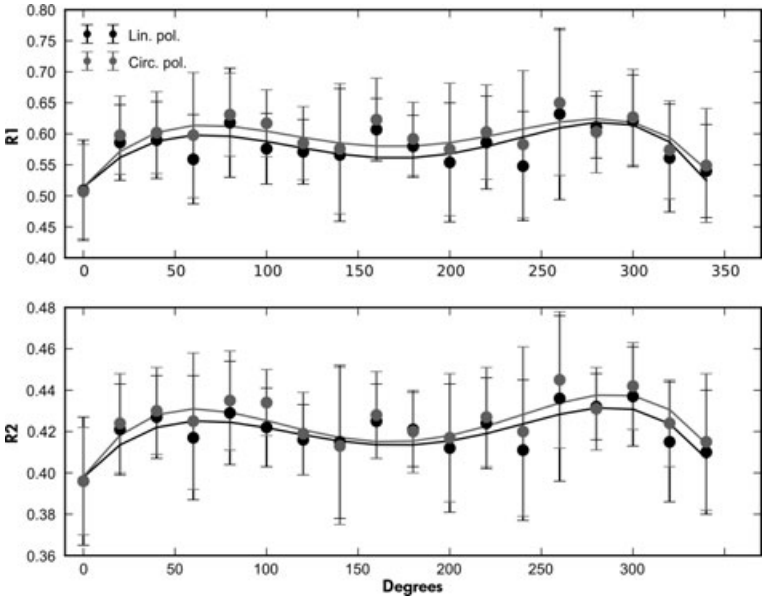

Figure 10. Results of test 8; the influence of sample preparation. At 0° and 180°, the laser beam is perpendicular to the schistosity, and at 90° and 270°, the orientation is parallel. The solid lines are weighted polynomial regressions which represent the trend in the data better. Lin. pol., linear polarised laser light; circ. pol., circular polarised laser light.

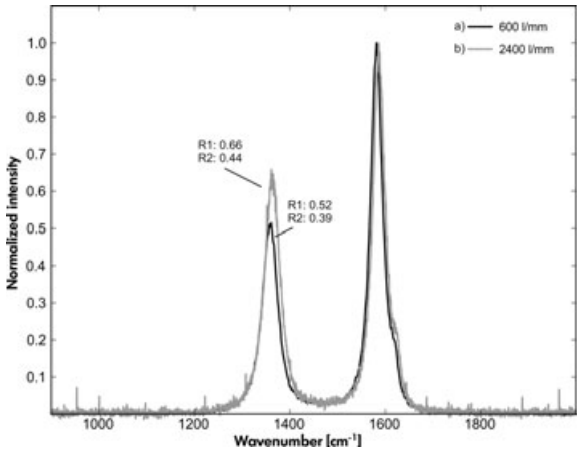

Figure 11. Results of test 9; Raman spectra of the same measuring spot acquired by different gratings. The R1 ratios of spectrum a) and b) are 0.52 and 0.66, and the R2 ratios are 0.39 and 0.44.

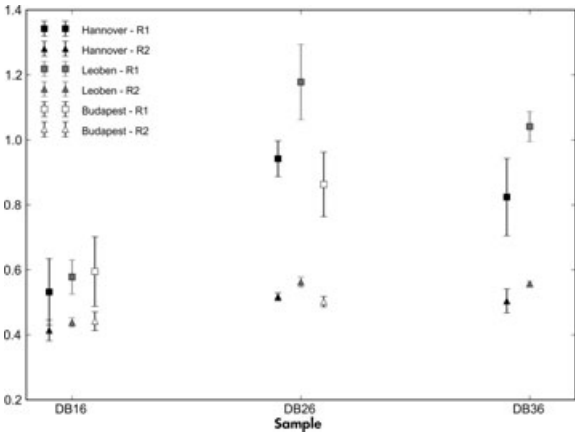

Figure 12. Results of test 9; the spread of the R values obtained by the interlaboratory comparison. The same excitation wavelength (532 nm), but different Raman systems were used in the laboratories. Depending on the sample, significant differences are observed between the laboratories, especially in the R1 ratio.

systems. As the excitation wavelength has a prime influence on the Raman spectra of CM (Vidano *et al.* 1981, Wang *et al.* 1990, Matthews *et al.* 1999, Sato *et al.* 2006), only spectrometers using the same excitation wavelength were tested. Therefore, three samples were measured with a Dilor LabRAM spectrometer (Department of Applied Geosciences and Geophysics, University Leoben, Austria) equipped with a 532 nm Nd-YAG laser, a spectral grating of 1800 l mm^{-1} and a 40 × objective with a NA of 0.55, a Bruker Senterra spectrometer (Department of Mineralogy, University Hannover, Germany) also equipped with a 532 nm laser, a spectral grating of 1200 l mm^{-1} and a 50 × objective with a NA of 0.75 and a Horiba Jobin Yvon LabRAM HR800 UV (Department of Mineralogy, Eötvös Loránd University, Budapest) also with a 532 nm laser, a spectral grating of 600 l mm^{-1} and a 100 × objective with a NA of 0.9.

To obtain a rough estimation of the sample heterogeneity, five spots were measured per sample with each system. As the difference in grating is 600 and 1200 l mm^{-1} for the three spectrometers tested, a difference of about 0.046 (R1) and 0.017 (R2) is expected for 600 l mm^{-1} difference and 0.093 (R1) and 0.033 (R2) for 1200 l mm^{-1} difference. Comparing the R1 and R2 ratios of the three samples, a far greater difference between the three spectrometers is evident, especially for the R1 ratio (Figure 12). Although some sample heterogeneity must also be considered here, the dominant reason for the observed variation must be the different spectrometer compounds like gratings or CCD cameras.

Consequently, this test demonstrates that some deviation in the Raman spectroscopic results of CM between different Raman systems cannot be avoided because of the large diversity of different components in the spectrometers that influence the spectral resolution, the signal intensity and peak positions.

Conclusions and outlook

The most important information derived from our study of Raman spectroscopy of carbonaceous material can be summarised as follows:

1 Spectral curve-fitting induces significant variation. The amount of variation depends on the degree of 'crystallinity' (i.e., structural order) of the CM sample and is highest in low crystalline material. The software used for curve-fitting does not influence the results.

2 Personal bias has a significant impact on the results. Thus, a standardised fitting procedure is proposed (see Appendix S1) for studies involving RSCM.

3 In very low-grade, organic-rich metasediments, the separation of CM by hydrochloric and hydrofluoric acid treatments generates a systematic shift in peak ratios when compared to measurements performed on untreated rock samples.

4 Sample heterogeneity appears in a high proportion of samples; the scatter of the Raman parameters in CM grain populations is much higher than the variation of

measurements within a grain. This represents a significant source of error in RSCM geothermometry and can be controlled only by enhanced number of measurements and by avoiding the detrital grains that carry inherited, high maturation signal.

5 Spectra of the same sample recorded with different Raman systems show distinct contrasts, due to differences in Raman system components.

6 When measuring higher grade crystalline samples (level 2 and level 3), the structural anisotropy of CM needs to be considered.

The presented results of nine experiments prove a rather high empirical deviation between RSCM geothermometrical data generated by different Raman instruments and data processing procedures. Different laser excitation wavelength generates even more deviation (Vidano *et al.* 1981, Wang *et al.* 1990, Matthews *et al.* 1999, Sato *et al.* 2006).

The total error due to the listed biasing factors is difficult to assess, but from the conducted experiments, it can be estimated that the error, especially for higher 'crystalline' samples, is in the range of 10–30 °C if the temperature is derived from the R2 ratio. For the RA1- and RA2 ratio, the error range is even larger. For the R2 ratio, this error range is below the given error of ± 50 °C of the RSCM thermometer (Beyssac *et al.* 2002a). However, if the same calibration is used by different laboratories, the above-stated biasing factors will influence the R2 ratios, and thus, the temperature estimates may differ approximately by the above-stated error range.

More consistent results can be expected if samples, measurements and thermometer calibrations are performed with the same instrumental set-up and data processing procedure.

Therefore, we suggest to improve the RSCM method by developing and distributing a natural sample series that can be used to calibrate a laboratory. When the calibration is based on measurements of reference materials, any instrumentation-related bias or drift can be corrected and even different excitation wavelengths could be used. In this case, every laboratory can generate its own calibration equation before and after a measuring session. Such a calibration system is used and widely accepted for stable isotope studies (V-SMOW, Gonfiantini 1978), for fission track and *in situ* U/Pb geochronology (Hurford and Green 1983, Compston *et al.* 1984, Frei and Gerdes 2009) or for the determination of 'illite crystallinity' (Kisch *et al.* 2004).

A reference material series is under preparation. The criteria used for selection of the reference materials are (a) independent and well-constrained maximum temperature, (b) fresh character, (c) homogeneity of the CM particles and (d) availability of the sample locality. This will be reported in detail in a coming communication (part 2).

Acknowledgements

We thank L. Laake and H.-J. Schubert (Metal Workshop of the Geoscience Center, University of Göttingen) for the construction of degree separator, M. Zuilen (Institut de Physique du Globe de Paris), T. Váczi (Eötvös Loránd University, Budapest), H. Behrens and A. Welsch (both Mineralogical Institute, Leibniz University Hannover) for Raman measurements in their laboratories and/or for the participation in the 'blind-test' evaluation of the spectra. We also thank two anonymous reviewers for their constructive critiques and thoughts that helped to improve the manuscript.

References

Ahn J.H., Cho M. and Buseck P.R. (1999)
Interstratification of carbonaceous material within illite. **American Mineralogist, 84**, 1967–1970.

Aoya M., Kouketsu Y., Endo S., Shimizu H., Mizukami T., Nakamura D. and Wallis S. (2010)
Extending the applicability of the Raman carbonaceous-material geothermometer using data from contact metamorphic rocks. **Journal of Metamorphic Geology, 28**, 895–914.

Ascher D., Dubois P.F., Hinsen K., Hugunin J. and Oliphant T.E. (2001)
Numerical Python, tech. report UCRL-MA-128569, **Lawrence Livermore National Laboratory**, http://numpy.scipy.org.

Baziotis I., Mposkos E. and Skarpelis N. (2006)
Raman micro-spectroscopy of carbonaceous material using the 633 nm line of a He-Ne laser: Application to the metamorphic rocks of Attica. **Geophysical Research Abstracts, 8**, 2.

Beyssac O. and Lazzeri M. (2012)
Application of Raman spectroscopy to the study of graphitic carbons in earth sciences. **In: Dubessy J., Caumon M.-C. and Rull F. (eds), Raman spectroscopy applied to Earth sciences and cultural heritage. European Mineralogical Union and Mineralogical Society of Great Britain and Ireland (London)**, 415–454.

references

Beyssac O., Goffé B., Chopin C. and Rouzaud J.N. (2002a)
Raman spectra of carbonaceous material in metasediments: A new geothermometer. Journal of Metamorphic Geology, 20, 859–871.

Beyssac O., Rouzaud J.N., Goffé B., Brunet F. and Chopin C. (2002b)
Graphitization in a high-pressure, low-temperature metamorphic gradient: A Raman microspectroscopy and HRTEM study. Contributions to Mineralogy and Petrology, 143, 19–31.

Buseck P.R. and Huang B.J. (1985)
Conversion of carbonaceous material to graphite during metamorphism. Geochimica et Cosmochimica Acta, 49, 2003–2016.

Bustin R.M., Ross J.V. and Rouzaud J.-N. (1995)
Mechanisms of graphite formation from kerogen: Experimental evidence. International Journal of Coal Geology, 28, 1–36.

Compagnini G., Puglisi O. and Foti G. (1997)
Raman spectra of virgin and damaged edge planes graphite. Carbon, 35, 1793–1797.

Compston W., Williams I.S. and Meyer C. (1984)
U-Pb geochronology of zircons from lunar breccia 73217 using a sensitive high mass-resolution ion microprobe. Journal of Geophysical Research, 89, B525–B534.

Crespo E., Luque F.J., Barrenechea J.F. and Rodas M. (2006)
Influence of grinding on graphite crystallinity from experimental and natural data: Implications for graphite thermometry and sample preparation. Mineralogical Magazine, 70, 697–707.

Dunkl I., Antolín B., Wemmer K., Rantitsch G., Kienast M., Montomoli C., Ding L., Carosi R., Appel E., El Bay R., Xu Q. and von Eynatten H. (2011)
Metamorphic evolution of the Tethyan Himalayan flysch in SE Tibet. Geological Society, London, Special Publications, 353, 45–69.

Eilers P.H.C. (2003)
A perfect smoother. Analytical Chemistry, 75, 3631–3636.

Endo S., Wallis S.R., Tsuboi M., Torres De León R. and Solari L.A. (2012)
Metamorphic evolution of lawsonite eclogites from the southern Motagua fault zone, Guatemala: Insights from phase equilibria and Raman spectroscopy. Journal of Metamorphic Geology, 30, 143–164.

Forer U., Kaindl R., Tropper P. and Mair V. (2009)
Mikro-Raman-spektroskopische Thermometrie an graphithaltigen Bündnerschiefern am Tauern Südrand im Ahrntal (S-Tirol, Italien). Mitteilungen der Österreichischen Mineralogischen Gesellschaft, 155, 183–188.

Frei D. and Gerdes A. (2009)
Precise and accurate *in situ* U-Pb dating of zircon with high sample throughput by automated LA-SF-ICP-MS. Chemical Geology, 261, 261–270.

Gonfiantini R. (1978)
Standards for stable isotope measurements in natural compounds. Nature, 273, 534–536.

Guedes A., Noronha F. and Prieto A.C. (2005)
Characterisation of dispersed organic matter from lower Palaeozoic metasedimentary rocks by organic petrography, X-ray diffraction and micro-Raman spectroscopy analyses. International Journal of Coal Geology, 62, 237–249.

Guedes A., Valentim B., Prieto A.C., Rodrigues S. and Noronha F. (2010)
Micro-Raman spectroscopy of collotelinite, fusinite and macrinite. International Journal of Coal Geology, 83, 415–422.

Huang E.P., Huang E., Yu S.C., Chen Y.H., Lee J.S. and Fang J.N. (2010)
In situ Raman spectroscopy on kerogen at high temperatures and high pressures. Physics and Chemistry of Minerals, 37, 593–600.

Hurford A.J. and Green P.F. (1983)
The zeta age calibration of fission-track dating. Isotope Geoscience, 1, 285–317.

Hutton A., Bharati S. and Robl T. (1994)
Chemical and petrographic classification of kerogen/macerals. Energy and Fuels, 8, 1478–1488.

Jehlička J. and Bény C. (1992)
Application of Raman microspectrometry in the study of structural changes in Precambrian kerogens during regional metamorphism. Organic Geochemistry, 18, 211–213.

Katagiri G., Ishida H. and Ishitani A. (1988)
Raman spectra of graphite edge planes. Carbon, 26, 565–571.

Kisch H., Árkai P. and Brime C. (2004)
On the calibration of the illite Kübler index (illite "crystallinity"). Schweizer Mineralogische und Petrographische Mitteilungen, 84, 323–331.

Kribek B., Hrabal J., Landais P. and Hladíková J. (1994)
The association of poorly ordered graphite, coke and bitumens in greenschist facies rocks of the Ponikla group, Lugicum, Czech-republic - the result of graphitization of various types of carbonaceous matter. Journal of Metamorphic Geology, 12, 493–503.

Kunert V. (1999)
Die Frankenwälder Querzone: Entwicklung einer thermischen Anomalie im Saxothuringikum. PhD thesis, Justus-Liebig-Universität Gießen, 171pp.

Lahfid A., Beyssac O., Deville E., Negro F., Chopin C. and Goffé B. (2010)
Evolution of the Raman spectrum of carbonaceous material in low-grade metasediments of the Glarus Alps (Switzerland). Terra Nova, 22, 354–360.

Ludwig K.R. (2003)
User's manual for Isoplot 3.00 – A geochronological toolkit for Microsoft Excel. Berkeley Geochronology Center Special Publication, 4, 71pp.

references

Marshall C.P., Howell G.M. and Jehlicka J. (2010)
Understanding the application of raman spectroscopy to the detection of traces of life. Astrobiology, 10, 229–243.

Matthews M.J., Pimenta M.A., Dresselhaus G., Dresselhaus M.S. and Endo M. (1999)
Origin of dispersive effects of the Raman D band in carbon materials. Physical Review B, 59, R6585–R6588.

Moré J.J. (1978)
The Levenberg-Marquardt algorithm: Implementation and theory. In: Watson G.A. (ed.), Numerical Analysis – Proceedings of the Biennial Conference Held at Dundee, June 28–July 1, 1977. Lecture Notes in Mathematics, 630, Springer (Berlin Heidelberg), 105–116.

Nzoussi-Mbassani P., Copard Y. and Disnar J.R. (2005)
Vitrinite recycling: Diagnostic criteria and reflectance changes during weathering and reburial. International Journal of Coal Geology, 61, 223–239.

Oberlin A. (1989)
High-resolution TEM studies of carbonization and graphitization. In: Thrower P.A. (ed.), Chemistry and Physics of carbon – A Series of Advances, 22, Dekker (New York and Basel), 1–143.

Oliphant T.E. (2007)
Python for scientific computing. Computing in Science and Engineering, 9, 10–20.

Pasteris J.D. and Wopenka B. (1991)
Raman spectra of graphite as indicators of degree of metamorphism. Canadian Mineralogist, 29, 1–9.

Pilgrim M. (2004)
Dive into Python. http://www.diveintopython.net/

Pimenta M.A., Dresselhaus G., Dresselhaus M.S., Cancado L.G., Jorio A. and Saito R. (2007)
Studying disorder in graphite-based systems by Raman spectroscopy. Physical Chemistry Chemical Physics, 9, 1276–1291.

Pócsik I., Hundhausen M., Koós M. and Ley L. (1998)
Origin of the D peak in the Raman spectrum of microcrystalline graphite. Journal of Non-crystalline Solids, 227–230, 1083–1086.

Powell R., Hergt J. and Woodhead J. (2002)
Improving isochron calculations with robust statistics and the bootstrap. Chemical Geology, 185, 191–204.

Quirico E., Raynal P. and Bourot-Denise M. (2003)
Metamorphic grade of organic matter in six unequilibrated ordinary chondrites. Meteorics and Planetary Science, 38, 795–811.

Quirico E., Rouzaud J.N., Bonal L. and Montagnac G. (2005)
Maturation grade of coals as revealed by Raman spectroscopy: Progress and problems. Spectrochimica Acta Part A, 61, 2368–2377.

Rahl J.M., Anderson K.M., Brandon M.T. and Fassoulas C. (2005)
Raman spectroscopic carbonaceous material thermometry of low-grade metamorphic rocks: Calibration and application to tectonic exhumation in Crete, Greece. Earth and Planetary Science Letters, 240, 339–354.

Rantitsch G., Grogger W., Teichert C., Ebner F., Hofer C., Maurer E.M., Schaffer B. and Toth M. (2004)
Conversion of carbonaceous material to graphite within the Greywacke Zone of the Eastern Alps. International Journal of Earth Sciences, 93, 959–973.

Reich S. and Thomsen C. (2004)
Raman spectroscopy of graphite. Philosophical Transactions of the Royal Society A, 362, 2271–2288.

Sadezky A., Muckenhuber H., Grothe H., Niessner R. and Pöschl U. (2005)
Raman microspectroscopy of soot and related carbonaceous materials: Spectral analysis and structural information. Carbon, 43, 1731–1742.

Saito R., Jorio A., Souza Filho A.G., Dresselhaus G., Dresselhaus M.S. and Pimenta M.A. (2002)
Probing phonon dispersion relations of graphite by double resonance Raman scattering. Physical Review Letters, 88, 027401-1–027401-4.

Sato K., Saito R., Oyama Y., Jiang J., Cancado L.G., Pimenta M.A., Jorio A., Samsonidze G.G., Dresselhaus G. and Dresselhaus M.S. (2006)
D-band Raman intensity of graphitic materials as a function of laser energy and crystallite size. Chemical Physics Letters, 427, 117–121.

Thomsen C. and Reich S. (2000)
Double Resonant Raman Scattering in Graphite. Physical Review Letters, 85, 5214–5217.

Tuinstra F. and Koenig J.L. (1970)
Raman spectrum of graphite. The Journal of Chemical Physics, 53, 1126–1130.

Valentim B., Guedes A., Prieto A.C. and Lemos de Sousa M.J. (2004)
Characterization of coal maceral groups by micro-Raman spectroscopy. Revista da Faculdade de Ciencias, 1, 318–325.

Vandenbroucke M. and Largeau C. (2007)
Kerogen origin, evolution and structure. Organic Geochemistry, 38, 719–833.

Vermeesch P. (2010)
HelioPlot, and the treatment of overdispersed (U–Th–Sm)/He data. Chemical Geology, 271, 108–111.

references

Vidano R.P., Fischbach D.B., Willis L.J. and Loehr T.M. (1981)
Observation of Raman band shifting with excitation wavelength for carbons and graphites. Solid State Communications, 39, 341–344.

Villanueva U., Raposo J.C., Castro K., de Diego A., Arana G. and Madariaga J.M. (2008)
Raman spectroscopy speciation of natural and anthropogenic solid phases in river and estuarine sediments with appreciable amount of clay and organic matter. Journal of Raman Spectroscopy, 39, 1195–1203.

Wang A., Dhamelincourt P., Dubessy J., Guerard D., Landais P. and Lelaurain M. (1989)
Characterization of graphite alteration in an uranium deposit by micro-Raman spectroscopy, X-Ray diffraction, transmission electron microscopy and scanning electron microscopy. Carbon, 27, 209–218.

Wang Y., Alsmeyer D.C. and Mccreery R.L. (1990)
Raman-spectroscopy of carbon materials - Structural basis of observed spectra. Chemistry of Materials, 2, 557–563.

Wiederkehr M., Bousquet R., Ziemann M.A., Berger A. and Schmid S.M. (2011)
3-D assessment of peak-metamorphic conditions by Raman spectroscopy of carbonaceous material: An example from the margin of the Lepontine dome (Swiss Central Alps). International Journal of Earth Science, 100, 1029–1063.

Wojdyr M. (2010)
Fityk: A general-purpose peak fitting program. Journal of Applied Crystallography, 43, 1126–1128.

Wopenka B. and Pasteris J.D. (1993)
Structural characterization of kerogens to granulite-facies graphite - Applicability of Raman microprobe spectroscopy. American Mineralogist, 78, 533–557.

Yui T.F., Huang E. and Xu J. (1996)
Raman spectrum of carbonaceous material: A possible metamorphic grade indicator for low-grade metamorphic rocks. Journal of Metamorphic Geology, 14, 115–124.

Supporting information

Additional Supporting information may be found in the online version of this article:

Appendix S1. Evaluation Protocol.

This material is available as part of the online article from: http://onlinelibrary.wiley.com/doi/10.1111/j.1751-908X.2013.12011.x/abstract (This link will take you to the article abstract).