Natural Products

Deutsche Ausgabe: DOI: 10.1002/ange.201600334
Internationale Ausgabe: DOI: 10.1002/anie.201600334

Structure and Metabolic-Flow Analysis of Molecular Complexity in a ^{13}C-Labeled Tree by 2D and 3D NMR

*Takanori Komatsu, Risa Ohishi, Amiu Shino, and Jun Kikuchi**

Abstract: *Improved signal identification for biological small molecules (BSMs) in a mixture was demonstrated by using multidimensional NMR on samples from ^{13}C-enriched Rhododendron japonicum (59.5 atom%) cultivated in air containing ^{13}C-labeled carbon dioxide for 14 weeks. The resonance assignment of 386 carbon atoms and 380 hydrogen atoms in the mixture was achieved. 42 BSMs, including eight that were unlisted in the spectral databases, were identified. Comparisons between the experimental values and the ^{13}C chemical shift values calculated by density functional theory supported the identifications of unlisted BSMs. Tracing the ^{13}C/^{12}C ratio by multidimensional NMR spectra revealed faster and slower turnover ratios of BSMs involved in central metabolism and those categorized as secondary metabolites, respectively. The identification of BSMs and subsequent flow analysis provided insight into the metabolic systems of the plant.*

The comprehensive analysis of biological small molecules (BSMs) by NMR is highly important for clinical, agricultural, and industrial science.[1] However, the development of comprehensive BSM analysis has been hampered by the challenges involved in BSM identification. At present, the identification of BSMs largely relies on databases of standard spectra. Various spectral databases for use in the NMR analysis of BSMs have been developed and made publically available, including the human metabolome database (HMDB),[2] the Madison–Qingdao metabolomics consortium database (MMCD),[3] and the SpinAssign database.[4] However, currently, the number of BSMs in these databases (approximately 1000 or less) are still far smaller than the total number of BSMs found in nature.[5] Many BSMs have been identified by studies in natural products chemistry, with over 270,000 BSMs currently listed in the dictionary of natural products (as of September 4, 2015, http://dnp.chemnetbase.com/).

The aim of the present study was to enable BSM identification beyond spectral database matching. We employed the strategy of using multidimensional NMR with ^{13}C isotope labeling. While there have been several approaches using a similar strategy in bacteria, nematode, plant, silkworm, mouse, and so on,[6] in these cases, BSMs included in spectral databases were focused on as predominant targets. *Rhododendron japonicum*, an ericaceous bushy tree, was chosen as a test case for a multidimensional NMR metabolomics study because the number of BSMs is much higher in plants than in animals.

The *R. japonicum* were cultured in an acryl chamber filled with synthetic air containing ^{13}C-labeled carbon dioxide for ^{13}C incorporation via photosynthesis (Figure 1a).[7] As a result, the entire plant, including macromolecules and BSMs, exhibited up to 59.5% ^{13}C enrichment in the aerial parts of the plants cultured for 14 weeks in the ^{13}C-enriched environment; these plants were then used for NMR analyses (Figure 1b). The NMR isotopic analyses described below (shown in Figure 4) indicated that ^{13}C was evenly incorporated into each BSM, with a mean enrichment in BSMs of 73.3%.

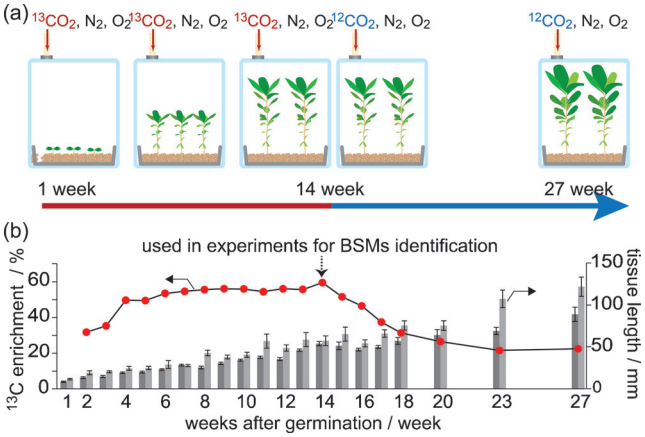

Figure 1. ^{13}C-labeling of *R. japonicum*. a) ^{13}C labeling of *R. japonicum* was performed by supplying ^{13}C-labeled carbon dioxide to an acryl chamber for 14 weeks. b) ^{13}C enrichment (points) and plant length (bars) values are shown as a function of time. ^{13}C enrichment reached 59.5% in 14th weeks, and then decreased as a function of time. Dark and light grey bars correspond to the length of the aerial part and the root of the plant, respectively. Error bars indicated the standard division of length in five plants.

^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) spectra of BSMs extracted from the aerial parts of ^{13}C-labeled *R. japonicum* were recorded using two solvent systems, a co-solvent of deuterated methanol and aqueous HEPES buffer (MeOD/HEPES; Figure S1a in the Supporting Information) and a deuterated aqueous potassium phosphate buffer (D_2O/KPi; Figure S1b), and then they were matched to the SpinAssign, HMDB, and MMCD databases. As a result, 19

[*] T. Komatsu, R. Ohishi, A. Shino, Prof. Dr. J. Kikuchi
RIKEN Center for Sustainable Resource Science
Graduate School of Medical Life Science, Yokohama City University
and Graduate School of Bioagricultural Sciences, Nagoya University
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 235-0045 (Japan)
E-mail: jun.kikuchi@riken.jp

Supporting information (including experimental details) and the ORCID identification number(s) for the author(s) of this article can be found under http://dx.doi.org/10.1002/anie.201600334.

and 27 metabolites, including amino acids, sugars, flavonoids, and terpenoids, were annotated in the spectra recorded with the MeOD/HEPES and D_2O/KPi solvents, respectively (Table S1 in the Supporting Information).

Multidimensional NMR spectra of a ^{13}C-labeled BSM mixture were used to obtain partial BSM structures, and further analysis enabled the tentative identification of seven BSMs not listed in the current databases (Figures S2, S3 and Table S2). 3D-(H)CCH-TOCSY spectra were used for elucidating partial structures of each BSM (Figure 2). Striped plots from 3D NMR spectra of ursolic acid, oleanolic acid, 6,7-dihydromonotropein, and grayanotoxin III are shown Figure 2 a–d. (H)CCH-TOCSY spectra provided most of the correlation information for both ^{1}H and ^{13}C in BSMs because of the ^{13}C-^{13}C magnetization transfer in BSMs present between quarternary carbon atoms. 3D (H)CCH-COSY spectra were also used for sequential resonance assignment using one-bond ^{13}C-^{13}C correlations. From these analyses, fundamental primary structures of BSMs were identified, including some that are not listed in the current databases.

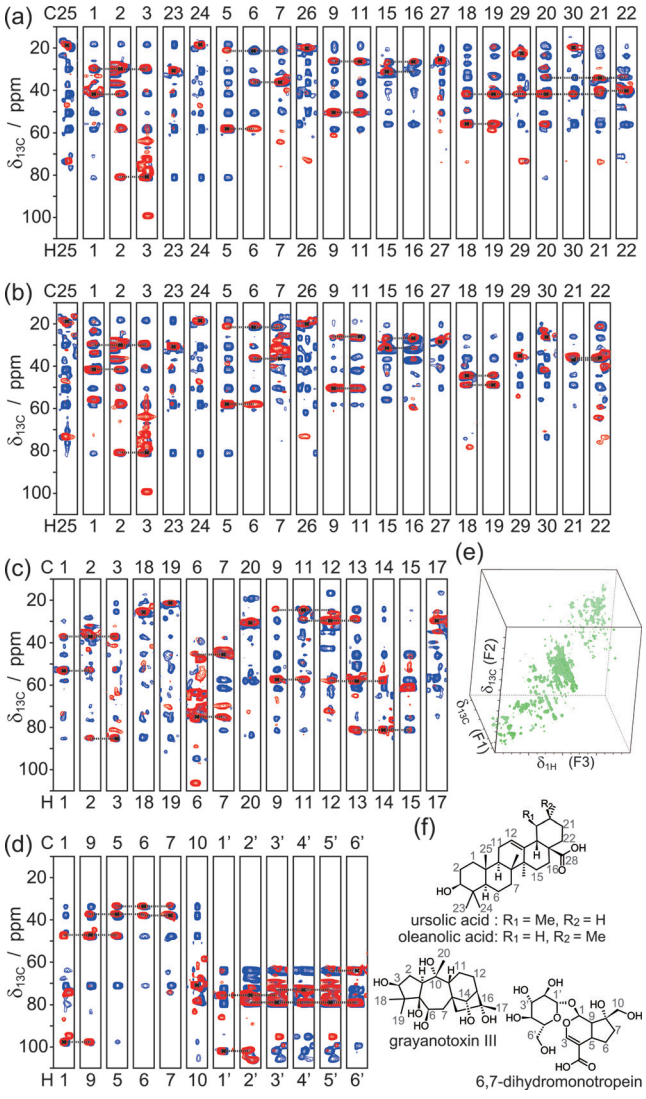

Figure 2. Sequential assignment of unexpected BSMs using (H)CCH-TOCSY and COSY spectra of ^{13}C-labeled BSM mixtures in MeOD/HEPES. Stripe plots are shown of the (H)CCH-TOCSY and COSY spectra corresponding to ursolic acid (a), oleanolic acid (b), grayanotoxin III (c), and 6,7-dihydromonotropein (d). e) Cubic view of the (H)CCH-TOCSY spectra. f) Chemical structures of the identified compounds.

Further NMR analyses were required to obtain all of the ^{13}C correlations in the BSMs. ^{13}C-^{13}C COSY spectra were used to obtain ^{13}C-^{13}C correlations barely occurred in (H)CCH-TOCSY spectrum (Figure S4 a). For example, in 6,7-dihydromonotropein, correlations between C4/C5 and between C4/C11, and those between C17/C28, C18/C13, and C11/C12 were difficult to obtain in (H)CCH-TOCSY; however, they were obtained by ^{13}C-^{13}C COSY. Refocused heteronuclear multiple bond COSY with ^{13}C decoupling (D-HMBC) was used to obtain long-range correlations crossing over oxygen nuclei (Figure S4 b). In 6,7-dihydromonotropein, H1′/C1, H1/C1′, H1/C3, and H5/C1 three-bond correlations crossing over the oxygen nuclei between ^{1}H and ^{13}C were present in the obtained spectrum. One-bond ^{13}C spin systems are frequently separated by oxygen in BSMs, for instance, in an ester or ether linkage such as the glycoside linkage in oligosaccharides and glycosides.

Seven unlisted BSMs (ursolic acid, oleanolic acid, grayanotoxin III, monotropein, 6,7-dihydromonotropein, avicularin, and 3,7-*O*-dimethylquercetin) were elucidated by multidimensional NMR analyses by following this approach. Maslinic acid, which is a 2-hydroxylated form of oleanolic acid, was also identified in the spectra. Structures and chemical shifts of these BSMs are shown in Figure S3 and Table S2, respectively. In summary, at least 66.8% of the correlations in the ^{1}H-^{13}C HSQC spectrum were assigned to BSMs (Figure S5).

Quantum chemistry calculation techniques were used to verify the identified BSMs by calculating chemical shifts from first principles (Figure 3, see details in the Supporting Information).[8] The root mean squared errors (RMSEs) between experimental and theoretical ^{13}C chemical shifts were 2.89 (ursolic acid), 1.87 (oleanolic acid), 2.36 (maslinic acid), 3.75 (grayanotoxin III), 2.79 (monotropein), 2.69 (6,7-dihydromonotropein), 3.32 (avicularin), and 3.10 ppm (3,7-*O*-dimethylquercetin).

To evaluate whether such errors support the validity of the identified structures, RMSEs of the ^{13}C chemical shifts in 43 common BSMs (including sugars, amino acids, short-chain fatty acids, alcohols, flavonoids, and steroids) were calculated by using the same methods (Figure S6). The obtained mean and standard deviation (SD) values of the RMSEs were 2.61 and 1.22 ppm, respectively. Therefore, the RMSEs of the ^{13}C chemical shifts in the eight identified BSMs were within mean ± SD RMSE of the 43 BSMs. These results support the validity of the identification, because larger RMSE values were often obtained for the incorrect structures of BSMs with known spectra (Figure S7).

The RMSEs of the ^{1}H chemical shifts were 0.29 (ursolic acid), 0.14 (oleanolic acid), 0.22 (maslinic acid), 0.44 (grayanotoxin III), 0.26 (monotropein), 0.36 (6,7-dihydromonotropein), 0.37 (avicularin), and 0.33 (3,7-*O*-dimethylquercetin) ppm, while the mean and SD values for the RMSEs in the 43 common BSMs were 0.28 and 0.11 ppm, respectively.

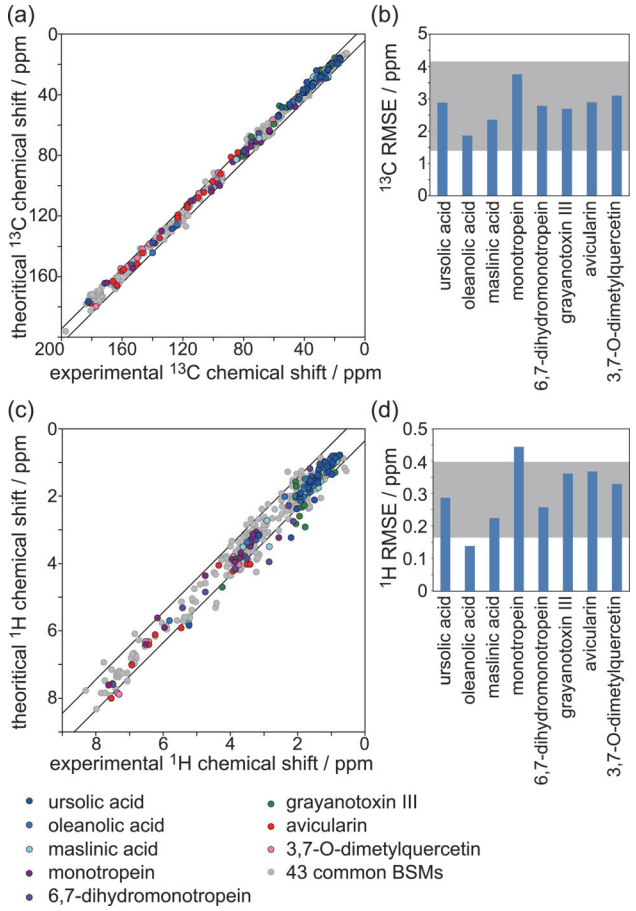

Figure 3. Comparisons between experimental and theoretical chemical shifts in the elucidated eight BSMs and 43 common BSMs. Experimental chemical shifts were compared with theoretical chemical shifts for ^{13}C (a) and ^{1}H (c) NMR. Each chemical shift is shown in Table S2. Drawn lines show the mean $\pm 2\times$SD chemical-shift error for the 43 common BSMs. RMSEs for each elucidated BSM are shown for ^{13}C (b) and ^{1}H (d) NMR. The gray domain denotes the mean $\pm$ SD values for RMSEs in the 43 common BSMs (^{13}C: 2.61 $\pm$ 1.22 ppm, ^{1}H: 0.28 $\pm$ 0.11 ppm, Figure S5). All ^{13}C RMSEs in the elucidated BSMs were within the range of mean $\pm$ SD.

The RMSE of the ^{1}H chemical shift in oleanolic acid was thus smaller than the mean−SD, and that of grayanotoxin III was larger than the mean + SD, the limit of the RMSE range of the 43 common BSMs. Since the calculations were performed in vacuum, these errors are partially due to neglect of solvent effects. While the NMR spiking experiments also contributed to verify the elucidated structures, they required actual molecules; therefore, these experiments were applicable only for some molecules (Figure S8).

A literature search of studies discussing the identified BSMs and *R. japonicum* was performed. Monotropein and 6,7-dihydromonotropein were found in juice of cranberries and blueberries.[9] Similar to *R. japonicum*, these are ericaceous plants. Chemotaxonomic studies have revealed that the chemical and molecular characteristics of plant metabolites are related to phylogenetic lineage.[10] Ursolic acid, oleanolic acid, and maslinic acid were also found in various berries.[11] Grayanotoxins were previously found in ericaceous plants including *R. japonicum*.

Describing the metabolic state is one of the goals of metabolome analysis, and structural elucidation for BSM identification is one of the means used for this purpose. Turnover ratio in the organism is closely associated with the functions of BSMs and metabolic state, and metabolic-flow analysis is a powerful analytical tool in biochemistry.[12] Metabolic-flow analysis using a time series of zero-quantum-filtered (ZQF)-TOCSY provided insight into the overall metabolic dynamic state in *R. japonicum*.[13]

^{13}C isotopic ratios of the entire plant decreased with time in the 15th, 16th, 17th, 18th, 20th, 23th, and 27th weeks because the supply of $^{13}CO_2$ was stopped after the 14th week (Figure 1b). Isotopic analysis, in which the ^{13}C isotopic ratios for each carbon atom in each BSM were calculated using integral ratios of ^{13}C-coupled and non-coupled ^{1}H signals (Figure S9),[13] provided a time series from ^{13}C isotopic ratios in the BSMs with atomic resolution (Figure 4a). ^{13}C enrichments of BSMs in the 14th week appeared uniform (one distribution with small dispersion, mean: 73.3 %, SD: 4.98 %), however the histograms obtained in weeks 15, 16, 17, and 18 were not uniform (Figure 4b). Interestingly, some BSMs exhibited high turnover rates and some BSMs exhibited smaller turnover rates in weeks 15, 16, and 17. Sucrose, glucose, γ-aminobutyric acid, and malic acid located near the central metabolism were in the former group, while epicatechin, quercetrin, 4-coumaric acid, 6,7-dihydromonotropein, monotropein, and ursolic acid, which belong to so-called secondary metabolism, were in the latter group (Figure 4c).

BSMs involved in secondary metabolism exhibited slower turnover rates than those involved in central metabolism, and their biological functions involve strategies for survival in harsh environments. Massad et al. compared the turnover time for phenolic glycosides, which are defensive secondary metabolites, and metabolites involved in plant growth through ^{13}C-labeling of *Populus trichocarpa* with $^{13}CO_2$.[14] The results indicated that the faster turnover ratios for sucrose and the production of phenolic glycosides were in accordance with the optimal-defense hypothesis. Grayanotoxins are a type of toxin and act as repellents for herbivorous animals. Similarly, sparassol, an aromatic antifungal compound, was identified in NMR spectra of the root (Figure S10); it was also found in roots in previous studies.[15] Ursolic acid, oleanolic acid, and maslinic acid, which are triterpenes, were found in waxes in the cuticular layer; they act as a barrier to water and provide resistance against pathogens.[16] These acids were more prevalent in the aerial part than in the root (Figure S10), which is consistent with the well-developed cuticle in the aerial part of higher plants.[17] Flavonoids are generally considered to act as antioxidants in plants.

In summary, the identification of BSMs and analysis of their metabolic flow provided insight into the reaction dynamics of *R. japonicum* metabolism. While the development of standard spectral databases is clearly very important for comprehensive identification of BSMs, in practice, it is difficult to collect all of them in nature. The method proposed here is a practical solution for improved BSM identification.

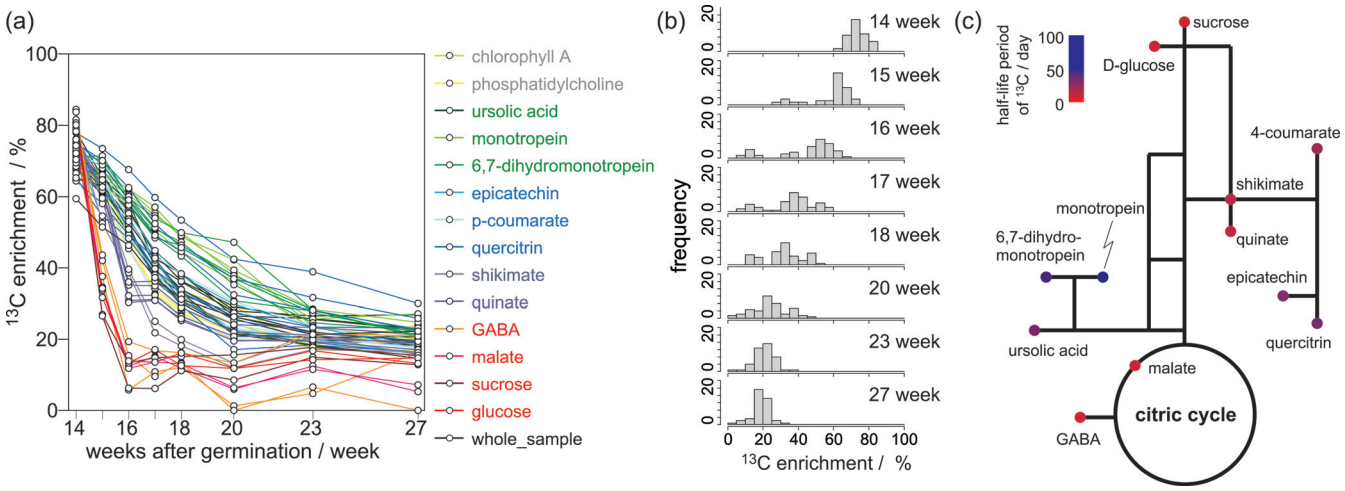

Figure 4. Metabolic-flow analysis provided insight into the metabolic state in the plant. Atomic-resolution ^{13}C enrichment values for the BSMs are shown as functions of time (a) and as histograms (b). These decreased over time owing to the incorporation of naturally abundant carbon dioxide. c) Each BSM is plotted in a schematic of a common metabolic pathway, with colors according to the averaged half-life ^{13}C to ^{12}C replacement. The half-life was calculated from ^{13}C enrichments of each BSM at 14th and 16th week.

Acknowledgements

The authors acknowledged Dr. Mitsuru Hirota (Tsukuba Univ.) for providing *R. japonium* seeds, Mr. Kenji Sakata (RIKEN) for IR-MS measurements. This research was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Keywords: density functional theory · isotopic labeling · metabolism · NMR spectroscopy · structure elucidation

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 6000–6003
Angew. Chem. **2016**, *128*, 6104–6107

[1] a) S. Ravanbakhsh et al., *PLoS One* **2015**, *10*, e0124219; b) J. K. Nicholson, J. Connelly, J. C. Lindon, E. Holmes, *Nat. Rev. Drug Discovery* **2002**, *1*, 153–161.
[2] D. S. Wishart et al., *Nucleic Acids Res.* **2013**, *41*, D801–D807.
[3] Q. Cui et al., *Nat. Biotechnol.* **2008**, *26*, 162–164.
[4] E. Chikayama, Y. Sekiyama, M. Okamoto, Y. Nakanishi, Y. Tsuboi, K. Akiyama, K. Saito, K. Shinozaki, J. Kikuchi, *Anal. Chem.* **2010**, *82*, 1653–1658.
[5] J. J. Ellinger, R. A. Chylla, E. L. Ulrich, J. L. Markley, *Curr. Metabolomics* **2013**, *1*, 28–40.
[6] a) C. S. Clendinen, G. S. Stupp, R. Ajredini, B. Lee-McMullen, C. Beecher, A. S. Edison, *Front. Plant Sci.* **2015**, *6*, 611; b) K. Bingol, F. Zhang, L. Bruschweiler-Li, R. Brueschweiler, *J. Am. Chem. Soc.* **2012**, *134*, 9006–9011; c) Y. J. An, W. J. Xu, X. Jin, H. Wen, H. Kim, J. Lee, S. Park, *ACS Chem. Biol.* **2012**, *7*, 2012–2018; d) Y. Sekiyama, E. Chikayama, J. Kikuchi, *Anal. Chem.* **2011**, *83*, 719–726.
[7] a) J. Kikuchi, T. Hirayama, *Methods Mol. Biol.* **2007**, *358*, 273–286; b) T. Mori, Y. Tsuboi, N. Ishida, N. Nishikubo, T. Demura, J. Kikuchi, *Sci. Rep.* **2015**, *5*, 11848.
[8] a) T. Misawa, T. Komatsu, Y. Date, J. Kikuchi, *Chem. Commun.* **2016**, *52*, 2964–2967; b) K. Ito, Y. Tsutsumi, Y. Date, J. Kikuchi, *ACS Chem. Biol.* **2016,** DOI: 10.1021/acschembio.5b00894.
[9] H. D. Jensen, K. A. Krogfelt, C. Cornett, S. H. Hansen, S. B. Christensen, *J. Agric. Food Chem.* **2002**, *50*, 6871–6874.
[10] R. J. Grayer, M. W. Chase, M. S. J. Simmonds, *Biochem. Syst. Ecol.* **1999**, *27*, 369–393.
[11] C. Neto, in *Berries and Cancer Prevention* (Eds.: N. P. Seeram, G. D. Stoner), Springer, New York, **2011**, pp. 41–49.
[12] I. A. Lewis, M. E. Campanella, J. L. Markley, P. S. Low, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18515–18520.
[13] T. Komatsu, R. Ohishi, A. Shino, K. Akashi, J. Kikuchi, *Metabolites* **2014**, *4*, 1018–1033.
[14] T. J. Massad, S. E. Trumbore, G. Ganbat, M. Reichelt, S. Unsicker, A. Boeckler, G. Gleixner, J. Gershenzon, S. Ruehlow, *New Phytol.* **2014**, *203*, 607–619.
[15] S. Woodward, H. Y. Sultan, D. K. Barrett, R. B. Pearce, *J. Gen. Microbiol.* **1993**, *139*, 153–159.
[16] I. Lara, B. Belge, L. F. Goulao, *J. Agric. Food Chem.* **2015**, *63*, 4005–4019.
[17] E. Domínguez, J. Cuartero, A. Heredia, *Plant Sci.* **2011**, *181*, 77–84.

Received: January 12, 2016
Published online: April 6, 2016

