Evolution of Quantitative Measures in NMR: Quantum Mechanical qHNMR Advances Chemical Standardization of a Red Clover (Trifolium pratense) Extract

Rasika S. Phansalkar,† Charlotte Simmler,† Jonathan Bisson,† Shao-Nong Chen,† David C. Lankin,†,‡ James B. McAlpine,† Matthias Niemitz,§ and Guido F. Pauli*†,

†UIC/NIH Center for Botanical Dietary Supplements Research and ‡Center for Natural Product Technologies, Department of Medicinal Chemistry and Pharmacognosy, University of Illinois at Chicago, College of Pharmacy, 833 S. Wood Street, Chicago, Illinois 60612, United States
§PERCH Solutions Ltd., Puijonkatu 24 B 5, 70110 Kuopio, Finland

ABSTRACT: Chemical standardization, along with morphological and DNA analysis ensures the authenticity and advances the integrity evaluation of botanical preparations. Achievement of a more comprehensive, metabolomic standardization requires simultaneous quantitation of multiple marker compounds. Employing quantitative 1H NMR (qHNMR), this study determined the total isoflavone content (TIfCo; 34.5−36.5% w/w) via multimarker standardization and assessed the stability of a 10-year-old isoflavone-enriched red clover extract (RCE). Eleven markers (nine isoflavones, two flavonols) were targeted simultaneously, and outcomes were compared with LC-based standardization. Two advanced quantitative measures in qHNMR were applied to derive quantities from complex and/or overlapping resonances: a quantum mechanical (QM) method (QM-qHNMR) that employs 1H iterative full spin analysis, and a non-QM method that uses linear peak fitting algorithms (PF-qHNMR). A 10 min UHPLC-UV method provided auxiliary orthogonal quantitation. This is the first systematic evaluation of QM and non-QM deconvolution as qHNMR quantitation measures. It demonstrates that QM-qHNMR can account successfully for the complexity of 1H NMR spectra of individual analytes and how QM-qHNMR can be built for mixtures such as botanical extracts. The contents of the main bioactive markers were in good agreement with earlier HPLC-UV results, demonstrating the chemical stability of the RCE. QM-qHNMR advances chemical standardization by its inherent QM accuracy and the use of universal calibrants, avoiding the impractical need for identical reference materials.

Red clover (Trifolium pratense L., Fabaceae) is a common botanical dietary supplement used by postmenopausal women.1−3 It contains mildly estrogenic isoflavones, including compounds 1−9, which are purported to alleviate hot flashes as well as other peri- and postmenopausal symptoms. Genistein (3) and daidzein (4) are known to be directly estrogenic, whereas biochanin A (1) and formononetin (2) are pro-estrogens that are metabolically activated in vivo to form 3 and 4, respectively, via demethylation.4−6 Cancer risks associated with hormone replacement therapy have encouraged women to look for safer botanical alternatives.7−9 Although not as rigorously regulated as drugs, dietary supplements still are expected to be safe for human consumption and, at the very least, must be true to their label claims. The recent “cease and desist” notices to dietary supplement manufacturers once more revealed the complexity of analytical challenges associated with the apparently simple question of botanical identity, authenticity, and overall botanical integrity.10,11 Moreover, even for products with fully established botanical integrity, phytochemical stability of their constituents is crucial for their safety and becomes a part of continued quality control measures.
In 2006, the University of Illinois at Chicago (UIC) Botanical Center performed phase I and II clinical trials that involved a custom-made isoflavone-enriched red clover extract (RCE). As a part of its chemical standardization, qualitative as well as quantitative analyses were used to determine the content of individual isoflavones using HPLC-UV and LC-MS methods. Today, 10 years later, the extract continues to be evaluated in various bioassays as well as drug safety trials as a part of collaborative research. As variations in the isoflavone patterns and content potentially can result in altered biological effects, assessment of the chemical stability of this extract was performed concurrently with implementing quantitative $^1$H NMR (qHNMR) methodology for multimarker metabolomic standardization. In addition to developing a qHNMR standardization procedure for RCE, the present study implemented advanced quantitative measures in qHNMR and, thereby, established a qHNMR-based metabolomic standardization protocol for the following 11 isoflavone and flavonol marker compounds in red clover preparations: biochanin A (1), formononetin (2), genistein (3), daidzein (4), calycosin (5), prunetin (6), irilone (7), pratensein (8), pseudobaptigenin (9), quercetin (10), and kaempferol (11).

### RESULTS AND DISCUSSION

#### Recent Advances in qNMR Applications

Very recently, the *Journal of Medicinal Chemistry* embraced the qHNMR concept by including absolute qHNMR as an acceptable and established scientific method for purity analysis. As a quantitative method, qHNMR continues to gain popularity in the biopharmaceutical industry and acceptance by the International Conference on Harmonization (ICH). Notably, qHNMR is a (relative) primary analytical method and is part of the general chapter $<$761$>$ of the *United States Pharmacopeia* (USP) as well as an official method in the *Japanese Pharmacopoeia* (JP17). With respect to the development of quantitative methodology for botanical standardization, the application potential of qHNMR for purity analysis, residual complexity studies, metabolomics, multimarker standardization, and orthogonal quantitation has been demonstrated.

Due to the presence of very many components, possible excipients, and the resulting intense signal overlap, the $^1$H NMR spectra of botanicals are typically very complex. Thus, while established qHNMR methods work well for a single component and relatively simple mixtures, analysis of such complex samples demands an expanded qHNMR tool box. In general, the choice of appropriate quantitative measures in qHNMR depends on the type of sample and the required accuracy. In classical qNMR, both relative and absolute quantitation methods rely on signal integration (INT). As INT-qHNMR is by far the most practiced form of qNMR, both accuracy and precision of most contemporary qNMR methods depend on the means of the integral measurement. Presently, three mechanistically different quantitative measures exist apart from classical INT-qHNMR (Figure 1). One entails peak height measurements, but due to limitations related to resolution of resonances and signal multiplicity, it was rendered impractical for this, given the complex peak overlap in the RCE qHNMR spectrum. Before delving into the mechanistic differences of the two spectral deconvolution (SD)-based approaches used in this study, it is important to differentiate the often interchangeably used terms "line", "peak", and "resonance". Multiple lines constitute an NMR resonance (or a signal that is the digital visual output of a

---

**Figure 1.** In qHNMR, aside from peak height quantitation (see text), three principal methods with different characteristics (Yes/No) exist for the extraction of quantitative measures ("numbers") that reflect the molarity of the analytical response to be used for the quantitative calculation. (i) Classical integration (INT-qHNMR) calculates the areas-under-the-curve of discrete regions of the spectra. (ii) Spectral deconvolution (SD) that employs peak fitting methods (PF-qHNMR), such as global SD (GSD), are limited by line and peak assignment, especially in the case of non-first-order effects. (iii) Quantum-mechanics (QM)-based total SD. QM-qHNMR derives exact measures from the spin populations of all target analytes. It has the ability to account only for relevant individual line intensities, thereby excluding "impurities" from the quantification.
resonance) due to the presence of scalar couplings (giving rise to signal multiplicities). The number of observable lines depends on the digital resolution as well as the natural line width of the signals. Peaks refer to only the visual "top" of each line or resonance. Linear SD is the second and more viable alternative to integration. It involves SD of select lines or entire spectra via mathematical peak fitting (PF) algorithms. PF-qHNMR methods take a one-by-one approach to the fitting of visible peaks (not spectral lines) and do not involve quantum mechanical (QM) calculations of the underlying spin systems. The third alternative to integration, builds on the method of $^1H$ iterative full spin analysis (HiFSA)$^{18-20}$ and consists of the QM-driven generation of a simulated qHNMR spectrum. This spectrum accumulates the subspectra of all individual components, and the underlying QM simulation includes their individual line shape characteristics as well as the relative intensities as present in the mixture. Intrinsically, the QM simulations yield precise intensities for all resonances, including those belonging to non-first-order (i.e., higher order) spin systems. The presence of non-first-order situations is nearly inevitable in $^1H$ NMR, even at very high magnetic field strengths (equivalent to $\geq 700 \text{ MHz} \cdot ^1H$), and its impact on qHNMR quantitation continues to be underrated despite the availability of well-established spin simulation tools.

During the initial stage of this study, it became evident that both SD methods (QM-based and non-QM/PF) outperform the classical region-based integration (INT-qHNMR) in terms of achievable accuracy and precision, especially with complex samples such as the RCE or similar natural products. With deconvolution processing being the common denominator in SD, one major aim of the study was a systematic comparison and evaluation of the capabilities of the two methods for covering a wide dynamic range of analytes. Additional study goals related to the multidisciplinary study context, supporting the continued biological studies of the RCE, are the establishment of an orthogonal analytical means of chemical standardization, as a means of advancing the accuracy of standardization, and the need to (re)assess the stability of the extract, using these new methods.

The recently developed HiFSA fingerprinting$^7$ and HiFSA-qHNMR$^{18,20,21}$ methodological frameworks were enabling technologies for these aims: they could be applied to the relatively complex RCE sample and results could be compared with those from previously published HPLC-UV and LC-MS standardization.$^7$ Moreover, to generate an independent data set that is mechanistically congruent with the studies performed some 10 years ago, a fast 10 min UHPLC-UV method was developed that allowed both chromatographic fingerprinting and quantitation.

**Advantages of qNMR over LC/MS-Based Methods.**

LC-UV/MS are the most commonly used analytical methods worldwide, partly as a result of the widely abundant instrumentation and the relatively high sensitivity of UV and MS detection. While often categorized as a relatively insensitive method, qNMR analysis typically requires only 15–30 min of instrument time, especially when samples are not mass limited and/or when the mass sensitivity of the instrumentation is high (cryoprobes, microprobes). Importantly, the identical and costly reference materials required for any LC-based quantitation are entirely dispensable in qNMR. Another intrinsic advantage is that multtargeted analysis can be performed on a single NMR sample, using standard qHNMR conditions, and can be repeated as needed.$^{22,23}$ This largely reduces tedious sample preparation protocols required for building LC calibration curves. Furthermore, qHNMR provides valuable spin parameter information simultaneously with quantitation, thereby enabling a thorough structural characterization of the constituents. The use of well-established quantitative conditions assures that signal intensities in qHNMR spectra are directly proportional to the molarity of hydrogen atoms giving rise to them. Notably, qHNMR is unaffected by response factors, which are inevitable in LC-based quantitation.

Considering these main advantages of qHNMR, including its time efficiency, simplicity of sample preparation, linearity of signal response, independence of sample intrinsic factors, and even cost considerations, qHNMR offers a unique set of attractive properties, making it a standardization method that deserves broader consideration in the botanical research community and industry. Furthermore, qHNMR can recognize dynamic properties such as chemical exchange (e.g., keto–enol tautomerism) or rotational/conformational isomerism, along with the identification of UV-transparent and/or poorly ionizable molecules. While such phenomena often evade LC/MS-based analysis, they can be important for the explanation of biological outcomes.

$^1H$ NMR Signal Assignment and Total Isoflavone Content (TIfCo). In light of the natural variation of phytoconstituents in source plants and the potential overlap with (chemo-)taxonomically related and unrelated species (both authentic species and adulterants), the specificity of botanical standardization increases with the number of markers used and their individual chemotaxonomic (sub)species specificity. Compared to the quantitation of a single marker, multimarker schemes enhance the significance of the standardization method by capturing a wider chemical window of the botanical metabolome. Accordingly, over the years, the development of new standardization protocols at the UIC Botanical Center has progressed from oligo- to multimarker/metabolomic schemes, in order to better capture the metabolic complexity of botanical preparations. The underlying hypothesis of this approach is that metabolomic standardization enhances the measure of botanical integrity and is the key to better reproducibility in botanical research.$^{10,11}$ As the number of targeted markers increases, the experimental effort for LC-based methods increases overproportionally due to the requirement of having high-quality (purity) reference materials available for each of the target compounds. As the chemotaxonomic significance and abundance (% content) are often mutually exclusive properties, this leads to substantial isolation efforts associated with the purification of milligram amounts of minor constituents. This is often rendered impractical and the main contributing factor why multimarker standardization schemes are not used more widely despite their acknowledged significant advantage. In contrast, the requirement for identical reference materials does not exist in qNMR. While there might be a need for the use of such compounds in the initial establishment of a qHNMR method (e.g., for the confirmation of peak assignments and assessment of method specificity, see discussion below), all subsequent qHNMR analyses can be performed without these reference materials. In addition, qHNMR is fully capable of multtargeted quantitation, as shown recently for preparations from *Ginkgo biloba* and *Glycyrrhiza* species.$^{20,21}$ Achieving a targeted multimarker standardization of the RCE using qHNMR was one important methodological aim of this study. The first step toward a qHNMR-driven multimarker standardization is the identification and unambiguous assignment of the signals of the target components in the qHNMR spectrum of a mixture. Isoflavones 1–9 (Chart 1) and
the flavonols 10 and 11 were identified in the RCE spectrum. The isoflavone signals were divided into four key regions (Figure 3): region 1 consisting of the methoxy proton signals from the B-ring (3.5–4.0 ppm); region 2 with the signals of the A-ring protons H-6/H-8 (6.0–6.7 ppm); region 3 with the signals of the B-ring AA′XX′ or AMX spin systems (6.7–8.0 ppm); and region 4 containing the C-ring H-2 singlets (8.2–8.5 ppm). Signal assignment in region 3 was the most challenging due to a combination of intense overlaps and higher order effects. The characteristic H-2 singlets in region 4 enabled the determination of the total isoflavone content of the extract. The H-2 resonances maintained their singlet property even after the application of very strong Lorentzian–Gaussian resolution enhancement window functions, proving their “pure singlet” characteristics. This was in line with the predictable lack of observable long-range couplings for these protons: the only potential coupling partners are located in the B-ring, giving rise to $^5J$ and $^6J$ couplings that evidently are well below the natural line width of the H-2 signal (∼1 Hz) and even lower than the achievable signal splitting (∼0.4 Hz).

The signals of biochanin A (1), formononetin (2), irilone (7), and quercetin (10) were assigned unambiguously by stacking of reference spectra and, in the case of 7, based on the data in the literature.

To demonstrate the specificity of close resonances, the assignments for genistein (3), daidzein (2), calycosin (5), and prunetin (6) were confirmed using spiking experiments. The resonances of the isoflavones pratensein (8) and pseudobaptigenin (9) were initially assigned based on previously published results.

Except for the methoxy protons, all isoflavone $^1H$ NMR signals were concentrated in the aromatic region. This naturally gave rise to severe overlap in the spectrum of an extract sample. For example, the 7.36 to 7.40 ppm portion of region 3 of the RCE spectrum contained a set of two protons from an AA′XX′ system belonging to four minor isoflavones, which altogether represented eight individual protons (Figure 4B, blue region). Such a high level of complexity precluded the use of classical integration and posed a significant challenge for an exact assignment. Moreover, field-dependence resulting from the involvement of non-first-order spin particles intrinsically limits the transmission of classical integration and non-QM linear SD

---

**Figure 2.** Non-QM deconvolution of region 4 of the RCE qHNMR spectrum using the global spectral deconvolution approach. The deconvoluted signals of the internal calibrant (IC: 3,5-dinitrobenzoic acid), the isoflavones 1–7, and flavonoid 10 are labeled. The colored lines represent the sum (blue), the fitted peaks (green), and the residual (red) of the deconvolution process.
Figure 3. Stacked qHNMR spectra (DMSO-d$_6$, 600 MHz) of the red clover extract and the isoflavones were divided into four key regions representing different spin systems present in the isoflavones. 6, prunetin; 5, calycosin; 4, daidzein; 3, genistein; 2, formononetin; 1, biochanin A; RCE, red clover extract. The NMR spin parameters for PS, PB, and IR were obtained from MetIDB. Depending on the type of the spin system, the A-ring protons fall into either region 2 (AAXX') or 3 (AMX).

Figure 4. Regions 3 and 4 of the qHNMR spectrum of RCE were the most relevant for the qHNMR analyses. Panel A shows the signal assignments of the two major isoflavones 1 and 2. Magnified expansions of the colored regions in panel B show the assignments of the seven minor isoflavones 3–9 and the flavonols 10 and 11 to demonstrate the importance of quantitative measures derived from QM-based deconvolution in strongly overlapping spectra (IC, internal calibrant; 3,5-dinitrobenzoic acid).
outcomes between laboratories. Even when considering a region with relatively simple resonances, such as region 4, the observed signal resolution still does not allow sufficiently broad or even uniform integral ranges. For example, the H-2 singlet of genistein (3) at 8.3083 ppm overlapped with the corresponding H-2 singlet of pseudobaptigenin (9) at 8.3132 ppm, and the signals of both the minor components were riding on the feet of the H-2 singlet of the major isoflavone foromononetin (2) (Figure 4B, green region). Another close overlap occurred between the H-2 singlets of daidzein (1) and calycosin (5) at 8.2743 and 8.2770 ppm, respectively. In this case, spiking experiments combined with HiFSA were necessary to ensure unambiguous assignment of the minor constituents.

Whereas signal overlap can be resolved for the purpose of analyte assignment, the complexity of the situation also shows that resolution of the underlying individual peak areas is not straightforward and cannot be achieved by simple definition of (integral) ranges. This underscores again that classical integration (iteration), respectively. Panels C and D demonstrate the higher order effects observed in the B-ring AMX system of calycosin (5) in the pure isoflavone (C) and the signal assignments in the RCE (D) (* indicates the assignment of H-6′ of 5 in the RCE).

Figure 5. Representative portions of the QM-qHNMR spectra of RCE showing the high congruence of the experimental (red) and QM-simulated (blue) spectra. Panels A and B show the assignments of the B-ring AMX system of the minor isoflavone pratensein (8) before and after the QM-based optimization (iteration), respectively. Panels C and D demonstrate the higher order effects observed in the B-ring AMX system of calycosin (5) in the pure isoflavone (C) and the signal assignments in the RCE (D) (* indicates the assignment of H-6′ of 5 in the RCE).
compounds in the extract can be affected by the temperature, sample matrix, pH, and/or water content. Notably, the $^1$H NMR spectrum of RCE without 3,5-dinitrobenzoic acid (3,5-DNB) added also showed the signal for H-2' in 8 as a doublet, ruling out the possibility of a drastic chemical shift perturbation due to the addition of an internal calibrant (IC). HiFSA was capable of resolving the higher order effects observed in the case of calycosin ($s$, Figure SC/D), as well as achieved assignments of the signals of 5 and 8 and in the spectral region with the most intense overlap. Using an algorithm that sets strong probability distribution (prior) on metabolite resonance patterns, the recently developed Bayesian SD approach allows automated peak (re)assignment under conditions of sample-dependent resonance shifts ($\Delta\nu$),30 but by default cannot overcome the intrinsic limitations of non-QM-based methods. Therefore, it was not included in the present study.

Along with the quantitation of individual components, qHNMR also enabled the close approximation of the total iso flavone content of the extract. The presence of a deshielded C-ring H-2 singlet resonance in the 8.20 to 8.55 ppm region is highly characteristic of all isoflavones and leads to the formation of “chromatogram-like” iso flavone fingerprints in the spectra of iso flavone-containing extracts. Nine of the ten iso flavone singlets observed in RCE could be identified, and seven could be quantified using the QM-based HiFSA-qHNMR method. In this way, integration of H-2 resonances in the entire region 4 against the calibrant signals was used to estimate the TIfCo of the RCE. The inter-isoflavone signal overlap was not critical, as region 4 was well isolated from the signals of the other regions in the spectrum (Figure 3). If interfering signals were present, a more elaborate SD approach would be necessary. Linear SD performed with the Total Line Shape (TLS) module of the PERCH software was used as a second means of estimating the TIfCo. Quantitation was done with reference to the IC signals, and a weighted average of the individual molecular weights of the iso flavones was used to determine the TIfCo as a total average iso flavone content. Using the integration and the TLS methods, the TIfCo in the RCE was estimated to be 34.5% and 36.5% w/w, respectively (Table 2). Given preference to the deconvolution method, this outcome showed that it is feasible to standardize an RCE to its TIfCo, instead of, or in addition to, the individual iso flavone percentages, in a single-step qHNMR analysis.

### Table 1. Comparative Results of the Quantitation (% w/w ± SD) of the RCE Iso flavones Using the QM-qHNMR, Non-QM (GSD) qHNMR, and UHPLC-UV Methods, Along with the Previously Published HPLC-UV Results2

| compound          | % w/w (mean ± std) in RCE | QM-qHNMR (2009) | GSD (non-QM) | UHPLC-UV | HPLC-UV (2009) |
|-------------------|----------------------------|----------------|--------------|----------|----------------|
| biochanin A (1)   | 15.2 ± 0.60                | 14.5 ± 0.83    | 13.04 ± 0.85 | 14.4     |                |
| foronmenhin (2)   | 16.0 ± 0.50                | 15.2 ± 1.16    | 16.01 ± 0.69 | 14.2     |                |
| genistein (3)     | 0.64 ± 0.04                | 0.64 ± 0.04    | 0.43 ± 1.07  | 0.41     |                |
| daidzein (4)      | 0.40 ± 0.04                | 0.35 ± 0.02    | 0.25 ± 1.19  | 0.23     |                |
| calycosin (5)     | 0.53 ± 0.11                | 0.53 ± 0.04    | 0.44         |          |                |
| prunetin (6)      | 0.71 ± 0.04                | 0.71 ± 0.05    | 0.59         |          |                |
| irilone (7)       | 2.21 ± 0.11                | 2.13 ± 0.09    | 3.20         |          |                |
| quercetin (10)    | 1.43 ± 0.11                | 1.40 ± 0.04    | 1.00         |          |                |
| total             | 37.3 ± 1.80                | 35.5 ± 2.19    | 34.6         |          |                |

### Advancement of Quantitative Measures in qHNMR

The 100% normalization, internal calibrant (IC), and external calibrant (EC) methods are the three established modes of quantitation in qHNMR, differing in the way calibration is achieved. Additionally, an EC can be used to calibrate the (residual) solvent signal, which can then be used as an IC for an extended series of qHNMR analyses (ECIC method),31 provided the same batch of solvent is used throughout and the exact volume is known. Traditionally, these quantitation modes have employed manual integration of the NMR signals as a quantitative measure. Classical integration is most commonly used for the purity evaluation of single chemical entities such as isolated or synthesized compounds and can be used successfully for slightly complex mixtures. However, the classical integration approaches are limited or even inadequate for the analysis of more complex samples such as plant extracts, foods, or biological fluids. This is a result of the discrepancy between the widths of resonances and the chemical shift dispersion required for clean, nonoverlapped signals for integration. The former depends on the natural spectral line width ($\Delta\nu_l$ ~ 1 Hz or slightly below in well-shimmered spectra) and the total width of the individual resonances, which is the sum of all $J$ values plus 2$\Delta\nu$, but notably only as long as first-order assumptions apply. Moreover, reliable integration requires relatively wide integration ranges (5~10$\Delta\nu$) in order to capture the full resonance, depending on the accuracy requirements. Another confining factor is the presence of $^13$C satellite signals, which limits the width of integral ranges and produces an additional level of signal overlap with minor constituents. Collectively, these challenges explain why advancement of quantitative measures is required to make 1D qHNMR fit for the purpose of metabolomic standardization. Whereas chemometric approaches employing multivariate analysis can be used for untargeted NMR metabolomics, some form of deconvolution (peak fitting) is needed to achieve a targeted quantitative analysis of known metabolites, but also of unknowns when assuming molecular structures and weights. Reported NMR peak fitting methods employ manual deconvolution.32,33 Moreover, generic software tools exist for highly flexible curve fitting, such as fityk (http://fityk.nieto.pl/) and general commercial statistics tools. One of the recent developments in this area geared toward NMR spectra is the Bayesian deconvolution and quantitation of metabolites using the automated BATMAN software. This R package enables automated peak assignment and quantitation based on a list of chemical shifts of the metabolites entered by the user.34,35

The present study entailed a targeted qHNMR multimarker standardization using two methods: a QM-based approach, which utilized the previously developed $^1$H full spin analysis to deconvolute the entire qHNMR spectrum, and a non-QM SD

### Table 2. Total Isoflavone Content (TIfCo) Calculated with Reference to the Internal Calibrant, 3,5-DNB, Using the 8.25 to 9.00 ppm $\delta$ Window

| signal                      | methods (integral/1H) | molecular weight (%) w/w |
|-----------------------------|-----------------------|--------------------------|
| IC (3,5-DNB)                | 6112.6                | 212.12                   |
| isoflavone region           | 9381.9                | 277.81                   |
| TIfCo estimate (% w/w)      | 34.53                 | 36.54                    |

$^a$Weight of IC = 1.38 mg; weight of RCE = 8.05 mg. $^b$Peak picking and deconvolution used the PAC and TLS modules of the PERCH software tool, respectively. The weight percentages of isoflavones obtained from the QM-qHNMR analyses were used to calculate the weighted average of the molecular weight of the isoflavones as a whole.

DOI: 10.1021/acs.jnatprod.6b00923

J. Nat. Prod. 2017, 80, 634–647
Spectral analysis (qQMSA). In this serum metabolomics study, QM aspects, the present study used QM-qHNMR (quantum mechanical quantitative 1H NMR) to highlight the benefits of using quantum mechanical calculations to derive the quantifiable measures, while continuing to refer to HiFSA for the actual SD process, involving full spin analysis through iterative calculations.

Comparison of QM vs Non-QM Deconvolution in qHNMR. Even under very careful experimental conditions, the analysis of qHNMR spectra is limited by a variety of factors such as incomplete signal assignments, extensive signal overlap, and imperfect baselines. These confounding factors are encountered commonly in the spectra of complex natural product samples such as RCE. Collectively, they influence the choice of the spectral window that is amenable to iteration or integration and, thereby, impact the achievable quantitative results. However, when compared to non-QM deconvolution, QM-based HiFSA SD yields a much enhanced resolution of signal overlap by virtue of the QM-based calculation of individual spectral lines. This includes all interdependencies of the individual lines, especially in non-first-order situations, which are very common. For instance, the H-2′/6′ signals of irilone (7) and prunetin (8) almost perfectly coincided with each other, as well as overlapped with those of the AA′XX′ spin systems of genistein (3) and daidzein (4) (Figure 4B).

The ability to assess the contribution of each component in such highly overlapped resonances and correctly assign peaks to individual components has to take into consideration the system of all spin particles (i.e., atoms) involved. Thus, the resolution of signals where multiple degenerated lines from both the individual and one or several other species overlap inevitably requires a QM approach (Figure 1). In contrast, non-QM, linear SD based approaches fail at this level of complexity, leaving the selection of the cleanest region with the least overlap as the only and predictably error-prone option for quantitation. Collectively, this explains the intrinsic limitation of non-QM-based deconvolution as a quantitative measure in qHNMR. Despite its advantage over classical integration, SD is flawed systematically and, thus, less rigorous for building comprehensive analyses of multiple components in complex mixtures. Analogous to chromophores in UV- and molecular/fragment masses in MS-detected quantitation, QM is the key to the definition of the spin parameters of the target compounds and ultimately guides the prerequisites of QM-qHNMR analyses. Additionally, non-QM, peak-fitted lines of unknown components can be added to improve the overall fit, their quantitation requires the assignment of partial structures and amu values. While any assumptions made in this process become part of the systematic error of the initial quantitation, additional knowledge gained afterward can be used to eliminate such errors due to the inherent calibration characteristics of qHNMR.

QM-qHNMR does not face this limitation for the development of multimarker standardization approaches. Using the HiFSA deconvolution mechanism, QM-qHNMR employs quantum mechanical rules and is, hence, “NMR-aware”. In contrast, other non-QM SD methods are “NMR-blind” and limited to modeling certain peak shape assumptions such as Lorentzian or Gaussian line shapes. As couplings and higher order effects involve multiple individual spins, non-first-order effects encode spectral information into the NMR resonances of coupled spin particles in other regions of the spectrum that might be subject to less overlap. As QM assures the integrity of these interdependencies, the degree of freedom and the number of iterative solutions that are consistent with the QM rules are reduced considerably, because only chemical shifts and coupling constants need to be optimized rather than frequencies and intensities for each single line. Non-QM approaches cannot take advantage of this mechanism. Thus, an important thing to note is that application of any QM method does require at least partial characterization of an analyte of interest in order to enable the use of accurately predicted and simulated spin parameters. While fully defined spin systems of target compounds with known molecular weight are desirable, the QM definition of partial spin systems (“spin cages”; e.g., sugars, amino acids) still enables targeted quantification of full or partial structures. QM-qHNMR utilizes the quantum mechanically simulated spectra, which represent the most accurate achievable replicas of the experimental spectra of the target analytes and can be derived from predicted initial spin parameters via an iterative QM calculation.

Importantly, because it employs the same mechanisms as HiFSA profiling does for pure compounds, QM-qNMR performs a QM-based SD of all spectral lines, not just of the (visible) peaks, of a given species in a mixture. Considering that individual peaks often consist of more than one line and that multiple peaks constitute the composite 1H signals (i.e., resonances) of a given proton, this has far reaching consequences for both the qualitative and quantitative interpretation of HNMR spectra: the distinction of lines and consideration of their intensities is fundamental to the accurate interpretation of all HNMR signals. The quantum mechanical description of a spin system correlates frequencies and intensities of the corresponding lines not only for each spin particle but via spin—spin coupling also through the complete spin system. These constraints dramatically reduce the number of variables and allow the complete SD of even completely overlapping lines also in the presence of higher order effects. In the relatively rare case of a pure first-order spin system these constraints degenerate to the simple first-order rules. In practice, significant proportions of the intensity of 1H NMR signals can reside outside the Δ “peak range” assigned to the respective proton (regardless of its assigned multiplicity), explaining at least in part the widely experienced relative weakness of INT-qHNMR. Taking into account the observed and/or achievable natural line width and signal-to-noise of the experimental spectra, it becomes evident that these effects can add considerable error to the accuracy and precision of SD methods in PF-qHNMR and, even more so, to classical INT-qHNMR.

In contrast, QM-based SD is less limited when analyzing the same experimental data. Due to the ability to correlate spin particles through the underlying QM rules, the intensity information on individual lines is encoded in the intensity of other lines as well. As a result, QM-qHNMR has the ability to automatically exclude line(s) (intensity) that cannot be explained by the given spin system(s) of the target analyte(s) from the quantification. In other words, QM-based SD methods...
have superior accuracy by virtue of their inherent capability to account only for the intensity of those lines (consequently also peaks and signals) that truly belong to the analyte.

In addition to its higher intrinsic accuracy, the HiFSA deconvolution process imparted in QM-qHNMR enabled a thorough qualitative analysis of the spectra and validated the identity of the target compounds simultaneously with their quantitation.

**Multimarker Standardization by Spectral Deconvolution: PF-qHNMR and QM-QHNMRI.** Once a library of HiFSA spin parameter sets has been created for the individual compounds, a multitarget quantitation can be performed for any mixture containing these components. HiFSA has been shown to be fit for the purpose of obtaining the exact chemical shifts up to ppb level accuracy, as well as coupling constants with ~10 MHz precision.\(^ {19, 28}\) This level of detail is particularly critical in the case of regio-/stereoisomers, higher order spectra, and when determining small coupling constants that cannot readily be detected visually. Especially when considered as precise relative values (\(\Delta \delta\) and \(\Delta \omega\)),\(^ {29}\) this helps in structural dereplication, confirmation, and differentiation of closely related analogues. Practical considerations of HiFSA processing involved the use of the integral optimization module (D-mode, coarse adjustment) of the PERCH iterator, which helped with the optimization of the chemical shifts, coupling constants, and signal intensities, thereby optimizing the relative ratios of the target components. The total line shape module (T-mode, fine adjustment) was used subsequently to refine the simulation of signal overlaps, deconvolute fine line shapes, and achieve the best overall match of the line characteristics of the spectrum. In this study, the iterations were performed until the observed and calculated spectrum were visually identical as indicated by a RMS value of \(\leq 0.1\). As the \(J\) coupling constants are largely unaffected by extrinsic interactions in the sample, they were kept fixed during the HiFSA iterative processes, typically after performing an initial optimization. Also, a “fit and lock” method\(^ {37}\) was utilized for certain spin parameters of the more problematic regions of the spectrum. Alternatively, in principle it is possible to mask and, thereby, exclude a certain region that still cannot be assigned reliably due to excessive overlap or unknown impurities, in order to achieve a converging iteration for the rest of the spectrum.

In this scenario, HiFSA SD retains its QM advantage via conservation of the underlying spin parameters, as long as the spin particles belonging to the molecule present in the overlapped region are also present elsewhere in the spectrum. Additionally, overlap with unknown components can be mimicked by adding simulated (non-QM) lines that represent these unknowns, improving the overall fit and giving lower global RMS values. A local RMS value of 0.030 for the most relevant region of interest (7.6 to 9.0 ppm) indicated an excellent match between the calculated and experimental spectrum (Figure 6).

The percent w/w content of the isoflavones biochanin A (1), formononetin (2), genistin (3), daidzein (4), calycosin (5), prunetin (6), irilone (7), and the flavonol quercetin (10) in the RCE was determined as 15.2, 16.0, 0.64, 0.40, 0.53, 0.71, 2.21, and 1.43, respectively, using the QM-qHNMR method (Table 1).

The alternative quantitative measure for multitarget qHNMR explored in this study involved the non-QM SD of the relevant portions of the spectrum. The global spectral deconvolution (GSD) module of the MestReNova software tool was used for both automated peak picking and SD. GSD was primarily developed to improve the quality of an NMR spectrum by deconvoluting and performing deliberate subtraction of solvent and/or impurity peaks from the spectrum.\(^ {36}\) Unlike the HiFSA-based QM-qHNMR approach, the initial time investment for GSD spectral processing is small, due to automation and the yield of acceptable deconvolution results. However, it must be kept in mind that GSD does not establish linkages between structural and spectroscopic features, nor can it verify any assignment and/or the recognition of (hidden) lines. This also brings up the need to distinguish QM-based lines (resonance frequencies) from experimentally distinguishable peaks (including peak shoulders). In the case of higher (non-first) order effects or excessive overlap, and due to its non-QM nature, GSD cannot assign peaks and/or achieve resolution of peak overlap, leaving the choice of a simpler, cleaner region of the spectrum as the only option for accurate analysis.

Based on these considerations, a region inclusive of the IC signals and the C-ring H-2 singlets was used for GSD-assisted quantitation. After deconvolution, the resulting spectrum consisted of the original signal, the deconvoluted lines, the

---

Figure 6. Observed (top/blue), residual (middle/green), and calculated (bottom/red) qHNMR spectra of RCE resulting from the QM-qHNMR analysis. The residual represents an RMS value of 0.031 for the 7.44 to 9.00 ppm region. The more complex part of region 3 (7.05 to 6.79 ppm; see main text) was optimized and locked. The final iteration was performed on the marked region, followed by simulation of the entire spectrum.
sum/fit peaks, and a residual (Figure 2). The resulting percent w/w contents of each isoflavone, thus determined, were 14.5, 15.2, 0.64, 0.35, 0.53, 0.71, 2.13, and 1.40 for biochanin A, formononetin, genistein, daidzein, calycosin, prunetin, iridone, and quercetin, respectively (Table 1; Table S4). These values are relatively close to those obtained with HiFSA-qHNMR, but show more substantial differences when compared to the published HPLC quantitation results, despite the similar TiCo.2

Table 1 compares the quantitation results obtained by QM-qHNMR, non-QM qHNMR using GSD, and the newly developed UHPLC-UV analysis with the 2009 HPLC results.2 The goal of achieving a more comprehensive, metabolomic description of the sample calls for the most inclusive quantitative measure, which currently is offered by QM-qHNMR. Although the initial data analysis and preparation of the HiFSA profile (including all the relevant NMR parameters) can require significant effort for complex samples, the established profile can be readily transferred within and between laboratories. The generated spin parameter files of the individual compounds can serve as libraries for future analysis of mixtures containing any or all of these compounds. As the QM basis of NMR ensures their validity across the various magnetic field strengths, QM-qHNMR is a scalable technology. Moreover, it can potentially be adopted for industrial quality control applications, e.g., to monitor batch-to-batch variations of known samples. More broadly, the present study showed that QM-qHNMR can be a valuable tool for the analyses of complex natural product matrices, including but not limited to botanical extracts.

One of the continuing challenges that are inherent to qNMR and affect QM-qHNMR as well is baseline “imperfections”. While they can represent measurement artifacts, they can also be real and due to signal overlapping of a myriad of minor constituents that fall well below the limit of detection (LOD), particularly in metabolomic samples.37 While masking of problematic regions from the iterations can improve the iteration and PF of resonances assigned to components that are identified by other means, these approaches may not be feasible with every sample and/or may require major additional analytical effort, at least for an exemplary sample. Furthermore, as iterators typically give more statistical weight to major signals, optimization of the line characteristics of signals from minor components is often limited when samples exhibit a very broad dynamic range. Certain weighting functions such as the “intensity weight parameter” may be used to give higher significance to minor components during iterations. In the experience of the authors, careful selection of processing and quantitation measure in SD-based qNMR outperforms automated approaches and justifies the required extra effort.

**Accuracy of Quantitative Measures in qHNMR.** The fundamental QM theory that underlies the observation of NMR infers that QM-based spectral deconvolution and quantitation like QM-qHNMR has a greater intrinsic accuracy compared to non-QM methods such as GSD (Figure 1). The importance and achievable level of accuracy of a method depends on the intended application and the type of sample, respectively. For example, in the case of a pharmaceutical formulation that contains a specific active ingredient, inert excipients, and a well-defined pharmacological target, high levels of quantitative accuracy and precision are desired and justified, as both are intertwined with the efficacy and potency of the material. This situation differs for complex plant extracts: large numbers and dynamic range of the phytochemicals, including multiple unknowns, are natural limits of achievable accuracy. Under such circumstances, the simultaneous identification and quantitation of as many components as possible, with a reasonable overall accuracy, becomes the primary goal.

Accuracy of both the QM-qHNMR and GSD methods was determined by spiking the IC-qHNMR analysis with a known amount of genistein (3). The spectra thus acquired were subjected to quantitative processing, and the amount of 3 before and after spiking was determined using these methods as described further in the Experimental Section. The accuracy of the QM-qHNMR and GSD methods for the multimarker quantitation of RCE was determined to be associated with 2.3% and 6.5% relative errors, respectively (S5 and Table S7). This outcome demonstrated that the QM-qHNMR approach holds its theoretical promise of higher accuracy in 1D qHNMR analysis. A 97% recovery associated with the quantitation accuracy for the spiked GE was determined by QM-qHNMR. A comprehensive qHNMR study had reported previously that an S/N ratio of ≥150 leads to more accurate qHNMR results.13 This suggests that all components with abundance levels of >0.9% w/w (S/N = 206; see also the Supporting Information) in the RCE were quantified with high accuracy.

**Orthogonal UHPLC UV Quantitation.** A new, 10 min, UHPLC method was developed for the quantitation of the four bioactive isoflavones biochanin A (1), formononetin (2), genistein (3), and daidzein (4). Their percent w/w content was determined to be 13.0, 16.0, 0.43, and 0.25, respectively. Compared to the qHNMR results, the differences between the quantities of the major isoflavones 1 and 2 could result from the observed different solubilities of the reference materials that likely affected the standard solutions for LC calibration and led to a slight overestimation of the quantities. The exploration of polymorphism and other batch-to-batch variations of reference standards was beyond the scope of the present qNMR study. As 1 exhibited lower solubility in MeOH compared to 2, extended sonication with mild heating was required to solubilize 1. However, the UHPLC-UV method was found to be concordant with the LC-UV/MS methods previously established for this sample2 and served as an orthogonal approach to the current study.

**Stability of a 10-Year-Old Red Clover Extract.** Over a period of 10 years, the clinical RCE had been stored in a sealed container at −20 °C. The orthogonal analyses used in the present study confirmed its stability over the prolonged storage. A UHPLC-UV profile (Figure 7) and quantitation showed that 1 and 2 were found consistently to be the major constituents, whereas 3 and 4 were present as minor components. This was confirmed by the two qHNMR methods, and the present outcome was consistent with the original HPLC standardization results. Despite the 10-year consistency of the findings, it is important to realize that repetition of any LC-based analysis requires revalidation due to its dependence on numerous instrument-related factors, particularly those related to (non)-linearity of the detectors (UV, MS), the injector system, and the characteristics of the columns. This does not apply to qHNMR methods: once established, they can be rerun anytime later, provided the NMR instrument is validated per se,35 and the qHNMR acquisition parameters are identical or demonstrated to be congruent between the different hardware and magnetic fields used. Another unique advantage of qHNMR is that additional analytes can be included even years later when peak assignments have become available, to study the same sample or the original FID of the qHNMR measurement. This emphasizes the immense value of original NMR data.40 The data can be used indefinitely not only for qualitative comparison but also for...
full-fledged quantitative measurements, provided IC is performed or EC/ECIC calibration data are archived concurrently.

In the present study, both qHNMR methods enabled the parallel quantitation of eight marker compounds: seven isoflavones, as well as quercetin. From a practical perspective, qHNMR proved to be the more time- and labor-efficient method in the long term and for large sample sets. The observed differences in the percentages of the markers determined using the three different methods are tolerable and can be attributed to the different intrinsic properties of the orthogonal analytical methods. As observed differences were within the combined error parameters of the methods, chemical degradation or interconversion being the cause of these variances seems unlikely. Accordingly, the clinical RCE was concluded to be phytochemically stable over the 10-year period and can be used confidently for future biological evaluations.

### CONCLUSIONS

The study identified quantum-mechanics-driven quantitative $^1$H NMR (QM-qHNMR) as currently the most capable and flexible method to achieve the targeted multimarker standardization of a complex botanical extract. Using a 10-year-old clinical RCE as study material, qHNMR allowed not only the quantitation of a total of nine biologically relevant and two additional markers but also the determination of the total isoflavone content. The TIFCo represents a unique reference point for botanical standardization, as it captures the entire group of bioactive isoflavones and their clinical relevance. The ability of the HiFSA SD process to account for peak overlap and higher order effects of the target markers led to what can be considered the most accurate signal assignment and quantitative measurement currently achievable in qHNMR by available established methods. The calculated recovery rate following spiking with genistein (GE) (Figure S5 and Table S6) in RCE was found to be 97.0%, demonstrating that the developed QM-QHNMR method is reasonably accurate and fit for the intended purpose of multimarker quantitation. Additionally, considering the complexity of the RCE $^1$H NMR spectrum, the extensive signal overlap, the large dynamic range of the major/minor components, the unavoidable presence of multiple unknowns, and the general limitations of reasonable experimental effort, the 2.3% relative error in quantitation observed for the QM-qHNMR method can be considered highly acceptable for the intended purpose of botanical standardization.

Comparing the two studied SD methods used to derive the quantitation measures, QM-based HiFSA vs non-QM-based GSD, the QM-qHNMR method exhibited a significantly lower percent relative quantitation error. This is in line with the theoretically predictable higher accuracy of QM-based approaches to $^1$H NMR analysis. A particularly notable strength of QM-qHNMR lies in the wealth of qualitative information that can be obtained in conjunction with the quantitative data. Due to its QM nature, QM-qHNMR produces a true representation of the experimental qHNMR data. Therefore, QM-qHNMR is a bivalent standardization method, which combines the identification of multiple marker compounds with their simultaneous quantitation (i.e., normalization), all in a single analysis. On a general note, (q)NMR analysis enables a highly efficient use of the information produced when made available in the scientific literature in its original form. Accordingly, the FIDs underlying the present work are made available for future use. Finally, the orthogonal UHPLC-UV and qHNMR analyses confirmed the stability of RCE over a period of 10 years. Good overall consistency with previous standardization data was observed, and the extract was hence deemed fit for further biological studies.

### EXPERIMENTAL SECTION

#### General Experimental Procedures.

Deuterated dimethyl sulfoxide (DMSO-$d_6$, 99.9% D) was obtained from Cambridge Isotope Laboratories Inc. (Andover, MA, USA). Calycosin and prunetin for spiking experiments came from Sigma-Aldrich Inc. (St. Louis, MO, USA). Biochanin A, daidzein, formononetin, and genistein were purchased from IndoFine Chemical Company Inc. (Hillsborough, NJ, USA). The internal calibrant, 3,5-dinitrobenzoic acid, was purchased from Fluka Analytical (Buchs, Switzerland). A gastight (1 mL) syringe, purchased from Pressure-Lok, Precision Sampling (Baton Rouge, LA, USA), was used for volumetric NMR sample preparation. The qHNMR experiments were performed on a Bruker (Karlsruhe, Germany) Avance 600 NMR spectrometer equipped with a 5 mm TXI cryoprobe. NMR data were analyzed and processed with Mnova 10.0.2 software from Mestrelab Research S. L. (Santiago de Compostela, Spain). PERCH NMR software (v.2013.1) from PERCH Solutions Ltd. (Kuopio, Finland) was used for all QM-based NMR spectroscopic analysis including iteration, simulation, and HiFSA, as described previously. A Shimadzu (Kyoto, Japan) Nexera UHPLC system equipped with a DAD detector was used for the UHPLC analysis of the extract. Quantitation was performed on a Kinetex 1.7 μm XB-C18 100 Å column (50 mm × 2.1 mm, S/N: 619546-18). Data analysis was performed using the Shimadzu LabSolutions software package.

#### Plant Material.

The red clover extract prepared previously for clinical trials (as documented previously) was used in this study. Acquisition of qHNMR Spectra. The qHNMR spectrum of the RCE was acquired using standard qHNMR conditions, which included a relaxation delay (D1) of 60 s, a calibrated 90° pulse (P1), and tuning and matching of the probe immediately preceding acquisition.

![Figure 7. UHPLC-UV chromatogram (254 nm) of the red clover extract, showing the quantified major bioactive isoflavones 1–4.](image-url)
For sample preparation, RCE and the IC (3,5-DNB) were weighed accurately as 8.05 mg and 1.39 mg, respectively, into one glass vial and dissolved in a total of 200 μL of DMSO-<i>d</i><sub>6</sub>. The solution containing the extract and the calibrant was transferred into a 3 mm NMR tube using 200 μL Drummond calibrated pipets. The qHNMR experiment consisted of 16 scans acquired with an RG (receiver gain) value of 64 and an acquisition time of 5.98 s.

**General NMR Data Processing.** Postacquisition processing included zero-filling of the 256k FID to 512k real data, a mild Lorentzian-Gaussian window function (exponential factor 0.3, Gaussian factor 0.05 in GF mode), and baseline correction (fifth-order polynomial). The residual DMSO-<i>d</i><sub>6</sub> signal at 2.500 ppm was used for chemical shift referencing. The NMR spectrum was exported as a jdx (JCAMP) file for QM-qHNMR processing. Spectra of the pure isolavones either had been acquired previously or were obtained from the Met-IDB database. Spiking experiments were performed for the minor components, genistein (3), daidzein (4), calycosin (5), and prunetin (6), in order to assign their NMR signals unambiguously.

**HiFSA and QM-qHNMR Processing.** First, 1H iterative full spin analysis<sup>1</sup> using the PERCH NMR software tool was performed for all isolavones (S1, Supporting Information) identified in the extract, except for the isolavones 7-9. The analysis involved initial prediction and subsequent iterative refinement of the spin parameters until the simulated spectrum matched the experimental spectrum. As a result, all chemical shifts, coupling constants, and individual line widths (δ, J, and <i>ω</i><sub>1,2</sub>) were determined for each compound, compiled into a spin parameter set (pms) file, one for each individual compound. Once HiFSA profiles for all the target components had been generated, the parameter sets for the individual compounds were copied and pasted to produce a combined pms file that will be utilized for the analysis of the RCE (S2, Supporting Information). In this combined pms file the coupling constants previously determined for each individual compound were kept fixed (unalterable) before starting the iteration process against the original 1H NMR spectrum of the RCE. A simulated spectrum of the combined pms file was then iterated against the experimental RCE to produce a fully matched calculated qHNMR spectrum. The iterations were performed until a close visual spectral match, also determined by an RMS value ≤0.1, was observed between the experimental and the calculated spectrum. Along with the refined spin parameters for the mixture, the process generated relative populations (as % mol/mol) of the individual species in the extract. Considering the exact weights of the extract and the IC, the absolute quantities of each individual component could be calculated using these populations (Scheme 1).

**Non-QM-Based Spectral Deconvolution qHNMR.** After post-acquisition processing of the qHNMR spectrum, SD was performed for region 4 (7.6-9.00 ppm) containing the signals of the IC, H-2 singlets of all isolavones, and the B-ring H-2’ doublet of quercetin (10) (Figure 2). The SD process was initiated by performing a peak picking of the entire spectrum using the global SD function of the MestReNova software tool with five fitting cycles. This yielded all detectable individual lines and their respective shapes and areas. Subsequently, the deconvoluted lines and their peak areas were assigned to the H-2 singlet resonances of the individual isolavones. Similarly, for composite resonances (doublets, triplets, etc.), the total area of their individual deconvoluted peaks additively constituted the area of the resonance, e.g., by adding the two individual peak areas of a doublet. Alternatively, or in addition to the automated GSD-driven method, peak picking can be performed manually, followed by individual peak fitting and using the “edit fit” option, as required. This generates a peak table, which contains the areas of the deconvoluted signals as qHNMR quantitation measures (S4, Supporting Information). The absolute weights of the target analytes can then be calculated using the IC method.<sup>12</sup>

**Accuracy, Recovery, and Limits of Detection and Quantitation (LOD/LOQ).** Accuracy was determined using the spike-recovery method. For the spiking solution, 5.10 mg of genistein (3) was accurately weighed into a glass vial and dissolved in 100.0 μL of DMSO-<i>d</i><sub>6</sub>. Three RCE samples containing a known amount of 3,5-DNB (internal calibrant) were quantified by both the QM-qHNMR and GSD methods as presented above, both before and after spiking the samples with a 1.00 μL aliquot of the genistein (3) standard solution amounting to 0.051 mg of 3. A Hamilton gastight GC manual syringe (1% accuracy) was used for measuring the spiking solution. The HiFSA profiles of the three samples used for accuracy determination are presented in S5.

---

Scheme 1. The Five Steps of the QM-qHNMR Protocol, Starting with the Acquired qHNMR Spectrum and Eventually Leading to the Quantitation of Individual Target Compounds<sup>14</sup>

---

"The resulting pms file contains all spin parameters (δ, J, and ω) and the individual populations of the target compounds (highlighted in red); t, target analyte, IC, internal calibrant; W, weight; M, molecular weight; P, purity; Po, population."
(Supporting Information) in the PERCH.pms text file format. From the percent w/w GE content of the three spiked samples, the percent relative error was calculated as follows.

First, the amount of GE in each unspiked sample was determined by the QM (HiFSAs)-qHNMR method (Table 1). On the basis of the known amount of GE in the spiked sample, the true value for each sample was determined as the sum of percent w/w GE in the unspiked sample and the spiked amount, 0.0510 mg (Table S6, Supporting Information). The percent recovery was calculated using the following formula:

\[
\% \text{ recovery} = \left( \frac{\text{conc spiked sample} - \text{conc unspiked sample}}{\text{spiking concentration}} \right) \times 100
\]

The LOD and LOQ thresholds were determined using the signal-to-noise ratio (S/N) method (Table S7, Supporting Information). In order to keep the matrix effect consistent, another calibrant, caffeine, was added to both the RCE and the 3,5-DNB-containing samples at seven concentrations ranging from 0.030% to 1.90% w/w. According to the ICH guidelines, the LOD at an S/N of 3:1 and LOQ at the S/N of 10:1 is commonly used in LC-based applications. A ratio of LOQ = 3.3 × LOD was used for the present study according to a previous qHNMR validation report.23 The S/N ratio calculator function in MestReNova was used to determine the S/N of individual signals. The methine proton singlet of caffeine at 7.98 ppm was used for the S/N determination, as it shares the ¹H relative integral value (1.000) with those of the target signals and falls into a relatively flat baseline region compared to the range of the upfront methyl singlets.

**UHPLC-UV Analysis.** A 10 min UHPLC-UV method was developed for analysis of the RCE. Solvents A and B were 0.1% formic acid and acetonitrile, respectively. The mobile phase gradient used was as follows: 20–26% B in 0.5 min, 26–31.3% B at 6.5 min, held isocratic at 31.3% B up to 6.7 min; re-equilibration from 31.3–20% B at 7 min; and reconditioning at 20% B up to 10 min. The flow rate was 0.6 mL/min, and the injection volume was 1.0 μL. The column oven, detector cell, and autosampler temperatures were maintained at 40, 40, and 4 °C, respectively, throughout the analysis. UHPLC-UV quantitation targeted the following four markers: biochanin A (1), formononetin (2), genistein (3), and daidzein (4). Calibration curves with nine concentrations ranging from 1.00 to 500 μg/mL and 0.25 to 125 μg/mL were generated for the major (1 and 2) and minor (3 and 4) isoflavones, respectively (Table S8, Supporting Information). Calculations considered the qHNMR purities of the reference materials used for calibration. All samples including the calibrants and the extract were run in triplicate with a blank injection in between each triplicate set. Stock solutions of 2 (in MeOH), 1, 3, and 4 (in EtOH) were prepared at 0.50, 0.50, 0.125, and 0.125 mg/mL, respectively. Aliquots of stock solutions were diluted with MeOH to produce the calibration solutions. RCE was dissolved in MeOH at 3.00 mg/mL. The PDA UV chromatograms were extracted at 254 nm for quantitative analysis.

**ACKNOWLEDGMENTS**

We are thankful to Dr. B. Ramirez for his expert NMR support at the UIC Center for Structural Biology (CSB). This research was supported by NIH grants P50AT000155 and U41AT008706 from NCCIH (previously NCCAM) and ODS. The construction of the UIC CSB facility and NMR instrument purchase were generously funded by a grant to Dr. P. Gettins from NIGMS Grant No. P41 GM068944.

**DEDICATION**

Dedicated to Professor Phil Crews, of the University of California, Santa Cruz, for his pioneering work on bioactive natural products.

**REFERENCES**

(1) Booth, N. L.; Piersen, C. E.; Banuvar, S.; Geller, S. E.; Shulman, L. P.; Farnsworth, N. R. *Menopause* 2006, 13 (2), 251–264.
(2) Booth, N. L.; Overk, C. R.; Yao, P.; Burdette, J. E.; Nikolic, D.; Chen, S.-N.; Bolton, J. L.; van Bremen, R. B.; Pauli, G. F.; Farnsworth, N. R. *J. Altern. Complement. Med.* 2006, 12 (2), 133–139.
(3) Low Dog, T. *Am. J. Med.* 2005, 118 (12B), 98–108.
(4) Wang, S. W. J.; Chen, Y.; Joseph, T.; Hu, M. *J. Altern. Complement Med. 2009*, 14 (3), 287–297.
(5) Heinonen, S.-M.; Wahlå, K.; Adlercreutz, H. *J. Agric. Food Chem.* 2004, 52 (22), 6802–6809.
(6) Tolleson, W. H.; Doerge, D. R.; Churchwell, M. L.; Marques, M. M.; Roberts, D. W. *J. Agric. Food Chem.* 2002, 50 (17), 4783–4790.
(7) Shumaker, S. A.; Legault, C.; Rapp, S. R.; Thal, L.; Wallace, R. B.; Ockene, J. K.; Hendrix, S. L.; Iii, B. N. J.; Jackson, R. D.; Kotchen, J. M. *JAMA* 2003, 289 (20), 2651–2662.
(8) Wasserman-Smoller, S.; Hendrix, S.; Limacher, M.; Geiss, G.; Kooperberg, C.; Baird, A.; Kotchen, T.; Curb, D. J.; Black, H. *JAMA* 2003, 289 (20), 2673–2684.
(9) Trial, R. C. *World Health* 2002, 288 (3), 321–333.
(10) Simmler, C.; Chen, S.-N.; Anderson, J.; Larkin, D. C.; Phansalkar, R. S.; Krause, E.; Dietz, B.; Bolton, J. L.; Nikolic, D.; van Bremen, R. B.; Pauli, G. F. *Herb. Gram* 2015, No. 106, 56–58.
(11) Simmler, C.; Chen, S.-N.; Anderson, J.; Larkin, D. C.; Krause, E.; Dietz, B.; Bolton, J. L.; Nikolic, D.; van Bremen, R. B.; Pauli, G. F.; Phansalkar, R.; Soejarto, D. D.; Larkin, D. C.; Krause, E.; Nikolic, D.; van Bremen, R. B.; Dietz, B.; Bolton, J. L.; Pauli, G. F. *Herb. Gram* 2016, No. 109, 60–64.
(12) Pauli, G. F.; Chen, S.-N.; Simmler, C.; Larkin, D. C.; Godecke, T.; Jaki, B. U.; Friesen, J. B.; McAlpine, J. B.; Napoliando, L. *J. Med. Chem.* 2014, 57 (22), 9220–9231.
(13) Malz, F.; Jancke, H. *J. Pharm. Biomed. Anal.* 2005, 38 (5), 813–823.
(14) In *United States Pharmacopeia and National Formulary* (USP31-NF26); 2011; p 298.
(15) In *Japanese Pharmacopeia* (JP17); 2016; p 2519.
