Electronic Supplementary Material (ESI) for Analytical Methods.

Supplementary Information

Detection of Honey Adulteration Using Benchtop ^{1}H NMR Spectroscopy

Yuki Rhee[1], Ella R. Shilliday[1], Yevgen Matviychuk[2], Thien Nguyen[1], Neil Robinson[1], Daniel J. Holland[2], Paul R. J. Connolly[1] and Michael L. Johns[1*]

[1] Department of Chemical Engineering, The University of Western Australia, Perth, WA 6009, Australia
[2] Department of Chemical and Process Engineering, University of Canterbury, Christchurch 8140, New Zealand

1. Experimental: Table of Honey Samples

Table S1: Primary saccharide content of un-adulterated honey and adulterants determined by benchtop NMR. H_2O contents of honeys A-K were determined gravimetrically.

	Primary sugars (%)				
Sample	**Maltose**	**Glucose**	**Sucrose**	**Fructose**	**H_2O (wt%)**
A	4	39	1	56	16.3
B	4	37	1	58	15.1
C	4	38	0	57	14.8
D	3	38	0	60	15.1
E	3	34	0	64	14.7
F	2	36	0	62	14.8
G	3	34	1	62	-
H	2	35	3	61	13.4
I	2	38	0	60	16.7
J	2	32	2	64	13.6
K	1	34	0	64	16
L	6.22	36.4	0	57.4	-
M	5.58	37.7	0	56.7	-
N	4.88	34.8	0	60.3	-
O	6.67	31.3	0	62	-
P	3.69	36.3	0	60	-
Q	5.49	38.1	0	56.4	-
R	7.54	37.5	0	55	-
S	5.63	34.2	0	60.2	-
T	5.76	37	0	57.2	-
U	5.55	35.3	0	59.2	-
V	5.83	37.3	0	56.9	-
W	5.78	34	0	60.2	-
X	4.46	36.4	0	59.2	-
Y	5.11	32.9	0	62	-
Z	6.61	34.9	0	58.4	-
Brown Rice Syrup (Batch 1)	31.9	54.9	5.07	8.05	-
Brown Rice Syrup (Batch 2)	21	49.7	9.26	20.1	-
Corn Syrup	49	12.5	15.1	23.4	-
Glucose Syrup	22.5	36.4	20.5	20.5	-
Sugar Cane Syrup	0	27.7	34.3	38	-
Wheat Syrup	25.9	41.2	26.6	6.25	-

2. Experimental: Spectral Analysis

We treat each NMR spectrum x as a superposition of K signatures u_k related to chemical species present in the sample, and weighted by their corresponding amounts, c_k:

$$x = \sum_{k=1}^{K} c_k u_k(\theta_k). \#(S1)$$

The sets of model parameters θ_k determine the appearance of each reference signal; they are unique for each compound and may include, for example, chemical shifts, relative peak intensities and widths, as well as J-coupling constants, if relevant. In general, the resulting model x is a complex-valued vector of length *N*. Although we work in the frequency domain here, the generalized formulation of Eq. S1 can also be applied to modelling FID signals directly in the time domain [1, 2].

Given an experimental NMR signal y**,** we aim to estimate a suitable set of parameters by fitting the model to the data in the least-squares sense:

$$\min_{\{\theta_k\}_{k=1}^{K},\varphi,c} \|\varphi(y) - Zc\|, \#(S2)$$

where, in the compact matrix notation, $Z \in C^{N \times K}$ is a model matrix with columns $Z_{:,k} = u_k(\theta_k)$ for $k = 1,\ldots,K$, $c = [c_1,\ldots,c_k]^T$ is a column vector of intensities, and $\|\cdot\|$ denotes the Euclidean norm. The function φ in the above equation denotes phasing and baseline correction applied to the experimental spectrum. Thus, each u_k eventually forms a column of Z and represents an elemental quantified entity, where the primary goal is to estimate the corresponding intensity c_k. We note that Z implicitly depends on the model parameters θ_k, whose optimal values are found by solving the non-linear problem of fitting the model to the data. Likewise, zero- and first-order phasing parameters along with a polynomial baseline that define φ are estimated during the same model fitting procedure and the experimental data are adjusted accordingly. Thus, given a fixed model matrix evaluated with the optimal parameters θ_k and the optimally phased and baseline-corrected spectrum, Eq. S2 leads to an ordinary least squares problem with respect to the intensities c_k.

The observed NMR peaks for a given chemical species are the result of quantum transitions undergone by the spins system. The linewidth and frequency result from the separation of energy levels, and the rate of signal decay is related to the relaxation or exchange [3]. The number and nature of *P* (Eq. S3) transition peaks are defined by the quantum mechanical (QM) properties of the spin system and the chemical structure, and in general can be expressed as some non-explicit function:

$$\{\omega_p b_p \alpha_p\}_{P=1}^{P} = f^{QM}(\delta, J, r). \#(S3)$$

This function accounts for the shielding, and hence chemical shift of different nuclei δ, the set of mutual J-coupling constants J, and possibly a relaxation model with rates r [4, 5]. Here we apply a simple isotropic damping model for relaxation and assume that all nuclei in the same spin system relax with equal relaxation rates leading to peaks with the same width. For smaller spin systems with no more than 12-13 coupled spins, evaluating f^{QM} is achieved by diagonalization of the Hamiltonian operator [6]. With knowledge of the aforementioned quantum properties of the spin system, peak positions ω, peak intensities b and relaxation rates α, completely define the reference component spectra in Eq. S2 as collections of corresponding Lorentzian peaks.

Although the peaks of an NMR spectrum relate to a chemical species, variations in experimental conditions, such as pH, can cause peaks to shift in frequency. Fortunately, these shifts will often act uniformly for groups of molecules, as was found to be the case for sugars [6], and as such movements can be accounted for in the QM model by shifting ω by the same amount. Hence, the QM models can be computed for fixed values of chemical shifts and the resulting response is shifted accordingly to fit the observations. Furthermore, any increase in linewidths due to magnetic field inhomogeneity can also be corrected with high level parameter adjustment of the Lorentzian broadening rate α.

The proposed method, including the evaluation and fitting of QM models, was implemented in Python 3.5. Model parameters were optimized using the SciPy implementation of the L-BFGS-B algorithm with basin-hopping [7]. Component intensities are reported as their respective mole fractions $\hat{\chi}_k = \hat{c}_k \Big/ \sum_k \hat{c}_k$. Uncertainty in each concentration is estimated to be ± 0.1 mol/mol based on previous measurements of sugar solutions. We report results as mass fractions, determined using the molecular mass of each fraction. This is trivial in most instances but for maltose and higher weight isomers this result is likely marginally inaccurate.

References

1. Matviychuk, Y.; von Harbou, E.; Holland, D. J., An experimental validation of a Bayesian model for quantification in NMR spectroscopy. *Journal of Magnetic Resonance* **2017,** *285*, 86-100.
2. Rubtsov, D. V.; Griffin, J. L., Time-domain Bayesian detection and estimation of noisy damped sinusoidal signals applied to NMR spectroscopy. *J Magn Reson* **2007,** *188* (2), 367-79.
3. Bain, A. D.; Fletcher, D. A.; Hazendonk, P., What is a transition? *Concepts in Magnetic Resonance* **1998,** *10* (2), 85-98.
4. Castillo, A. M.; Patiny, L.; Wist, J., Fast and accurate algorithm for the simulation of NMR spectra of large spin systems. *J Magn Reson* **2011,** *209* (2), 123-30.
5. Tiainen, M.; Soininen, P.; Laatikainen, R., Quantitative Quantum Mechanical Spectral Analysis (qQMSA) of (1)H NMR spectra of complex mixtures and biofluids. *J Magn Reson* **2014,** *242*, 67-78.
6. Matviychuk, Y.; Yeo, J.; Holland, D. J., A field-invariant method for quantitative analysis with benchtop NMR. *Journal of Magnetic Resonance* **2019,** *298*, 35-47.

7. Jones, E.; Oliphant, T.; Peterson, P., SciPy: Open Source Scientific Tools for Python. **2001**.