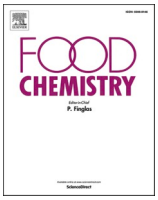
Food Chemistry 428 (2023) 136817

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem



Application of segmented analysis via multivariate curve resolution with alternating least squares to ^{1}H-nuclear magnetic resonance spectroscopy to identify different sugar sources

Cristian A. Fuentes [a,b], Mecit Halil Öztop [c], Macarena Rojas-Rioseco [a,b], Martín Bravo [a,b], Aylin Özgür Göksu [d], Marena Manley [e,*], Rosario del P. Castillo [a,b]

[a] Departamento de Análisis Instrumental, Facultad de Farmacia, Universidad de Concepción, Concepción 4070386, Chile
[b] Laboratorio de Bioespectroscopia y Quimiometría (BioSpeQ), Centro de Biotecnología, Universidad de Concepción, Concepción 4070386, Chile
[c] Department of Food Engineering, Middle East Technical University, Ankara 06800, Turkey
[d] Kayseri Sugar R&D Center, Kayseri Sugar Factory, Kayseri 38070, Turkey
[e] Deparment of Food Science, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa

ARTICLE INFO

Keywords:
Sugar beet
Brown sugar
Multivariate analysis
^{1}H NMR spectroscopy
Preprocessing

ABSTRACT

The different types of sugar employed in the food industry exhibit chemical similarity and are mostly dominated by sucrose. Owing to the sugar origin of and differences in production, the presence of certain minor organic compounds differs. To differentiate between sugars based on their botanical source, geographical origin, or storage conditions, commercial brown sugars and sugar beet extracts were analyzed by ^{1}H NMR spectroscopy applying a segmented analysis by means of multivariate curve resolution-alternating least squares (MCR–ALS). Principal component analysis and partial least squares-discriminant analysis yielded excellent differentiation between sugars from different sources after the application of this preprocessing strategy; without loss of chemical information and with direct interpretation of the results. By applying a segmented analysis via MCR–ALS to ^{1}H NMR sugar data, similar spectroscopic profiles could be differentiated. This improved the selectivity of ^{1}H NMR spectroscopy for sugar source differentiation which can be useful for industrial sugar authentication purposes.

1. Introduction

Sugar constitutes one of the main ingredients in food formulations. Sucrose is the most commonly used form of sugar. It is a nonreducing sugar comprising glucose and fructose monomers (Queneau et al., 2007). The anomeric carbons of both monomers are involved in glycosidic linkages, which prevents their mutarotation (Kamerling & Vliegenthart, 2021). Having only one form in the aqueous solution results in easier crystallization of sucrose compared to other sugars. Sucrose is primarily produced from sugar beet (*Beta vulgaris)* or sugar cane (*Saccharum officinarum)*. Worldwide, 80% of the annual sugar production is from sugar cane while 20% is from sugar beet (Arro et al., 2016; Ribeiro et al., 2016).

In sugar beet, sodium, potassium, betaine, amino acids, and nitrate affect the quality of sugar as they cannot be eliminated from the juice of sugar beet during the purification process (Kenter & Hoffmann, 2009). Moreover, factors such as storage conditions of the raw material, climatic temperature during the harvest, and soil composition of the sugar beet growing region can affect the efficiency of sugar extraction (Masetti et al., 2021; Vukov & Hangypal, 1985). Furthermore, the quality of sugar depends on the composition of sugar beet and, therefore, on the growing area and environmental conditions. Hence, the control and characterization of sugar beet, based on geographical region or cultivation area, are of great importance to the food industry to achieve high product quality. Several approaches using proton nuclear magnetic resonance (^{1}H NMR) spectroscopy-based metabolic profiling, have been described. These include evaluation of differences in leaves and roots of sugar beets subjected to different durations of intermittent drought (Wedeking et al., 2018), different degrees of resistance to *Cercospora* leaf spot in sugar beets (Sekiyama et al., 2017), and differences in the phytochemical profile of red beetroot from three different harvests (Giampaoli et al., 2021).

* Corresponding author.
E-mail address: mman@sun.ac.za (M. Manley).

https://doi.org/10.1016/j.foodchem.2023.136817
Received 11 April 2023; Received in revised form 19 June 2023; Accepted 3 July 2023
Available online 4 July 2023

Sucrose is mainly produced as white crystals; however, brown crystalline sugar (referred to as brown sugar hereafter) is an important commercial sugar that is widely produced and consumed globally (Chen et al., 2021a).

Brown sugar is either prepared from sugar cane juice by means of thermal processing or by mixing refined beet sugar with cane molasses (Chen et al., 2021a). Different types of edible brown sugar, including non-centrifugal cane sugar (NCS), muscovado sugar (MS), or brown granulated sugar (BGS) can be obtained depending on the manufacturing process (Chen et al., 2021). In addition, coconut sugar is similar in appearance and aromatic profile to that of brown sugar produced from either beet or cane sugar (Bachmann et al., 2022). Coconut sugar is produced by heating the coconut sap till the juices are saturated and sugar crystals are obtained (Nurhadi et al., 2018). Owing to the similarity in color and flavor, it is difficult to distinguish the origin of these brown sugars by means of only sensory evaluation. Furthermore, because brown sugar is usually sold at a higher price than white sugar, the former is susceptible to adulteration. For example, white sugar can be coated with caramelized sugar or synthetic dyes and sold as brown sugar.

Recently, gas chromatography–olfactometry-mass spectrometry has been used to differentiate between different types of brown cane sugars (NCS, MS, and BGS) as well as brown sugars produced in three provinces of China by comparing the constituent odor compounds (Chen et al., 2021, 2022). Also, three-dimensional fluorescence spectroscopy was used to identify and differentiate between natural and commercial brown cane sugars (Chen et al., 2021b). Similarly, the metabolic profile of brown beet, unrefined cane, and coconut blossom sugar was analyzed with 1H NMR (Bachmann et al., 2022).

In food analysis, NMR spectroscopy is considered similar to a fingerprinting technique (Caligiani et al., 2007) because it provides chemical information regarding the composition of the samples. However, elucidating the information in an NMR spectrum can be challenging given the strong overlap of the signals and complexity of its spectral interpretation (Masetti et al., 2021). To overcome these problems, multivariate analysis techniques can be applied to ^{1}H NMR data to reduce its dimensionality and enable the extraction of relevant information to identify similarities or differences between groups of samples. This has been illustrated by classifying samples from a number of different commodities based on geographical origin (Dimitrakopoulou et al., 2021; Longobardi et al., 2017; Schmitt et al., 2020; Wang et al., 2021), harvest time (Giampaoli et al., 2021), temporary stress (Munyai et al., 2022; Saviano et al., 2019; Wedeking et al., 2018) and adulterations (Bachmann et al., 2022; Ravaglia et al., 2019).

Generally, the chemometric models applied to ^{1}H NMR spectroscopy data use '*data binning*' to eliminate small variations between different samples due to chemical shifts produced by fluctuations in the pH, temperature, or concentration (Alam & Alam, 2004). However, binning the data across a defined frequency width decreases the spectral resolution. Also, grouping of overlapping signals in the same bin makes the direct interpretation of the results challenging. To improve spectral interpretation, Pérez et al. (2020) developed a methodology based on multivariate curve resolution-alternating least squares (MCR–ALS) (De Juan et al., 2014). This method was derived from the Decision Tree Correlation methodology reported by Puig-Castellví et al. (2017). The latter authors applied it as an independent preprocessing method to resolve the concentration (**C**) and spectral ($\mathbf{S^T}$) profiles in an ^{1}H NMR dataset of zebrafish samples. Furthermore, this method improves sample clustering when using PCA. In addition, Cavallini et al. (2021) proposed a strategy based on evaluating intervals via MCR–ALS as a resolution technique to characterize beer. The authors highlighted that the resolution of $\mathbf{S^T}$ profiles simplified the information, making it easily interpretable.

To the best of our knowledge, the application of segmented analysis via MCR–ALS to the ^{1}H NMR spectra for the differentiation based on geographical origin and storage conditions of sugar beet, and the botanical source of commercial sugars has not been explored before. Herein, segmented analysis via MCR–ALS was applied to ^{1}H NMR spectroscopy data to differentiate sugars based on their sources. MCR–ALS allow analysis of the ^{1}H NMR spectra of sugars, avoiding the issues associated with loss of chemical information that are observed when data binning preprocessing is used. It also removes, spectral noise and identifies the main chemical compounds responsible for the differences and the subsequent classification. Thus, segmented analysis via MCR–ALS was evaluated as a novel alternative method to data binning for classification or authentication purposes in the sugar industry.

2. Materials and methods

2.1. Reagents and chemicals

Deuterium oxide (D_2O, 99.9%), 3-(Trimethylsilyl)-propionic-2,2,3,3-d_4 acid sodium salt (TSP-d_4, 98%), potassium phosphate monobasic anhydrous (KH_2PO_4, >99%), potassium phosphate dibasic anhydrous (K_2HPO_4, >98%), and methanol (high-performance liquid chromatography grade) were purchased from Sigma-Aldrich (Darmstadt, Germany).

2.2. Sample collection and preparation

In 2022, twelve sugar beet samples, that were harvested from two geographical regions of the center-south of Chile (San Carlos and Los Ángeles) and three cultivation fields (Santa Isabel, Luciana, and Santa Laura) were kindly provided by the National Federation of Beet Growers (FENARE, Chile). Detailed information regarding the origin of the samples is provided in Table 1. From these samples, four fresh sugar

Table 1
Information of sugar beet and commercial sugar samples, including geographical region, country, cultivation field, brand, and code.

Source of sugar	Geographical region or country	Cultivation field or commercial brand	N° of samples	Code
Fresh sugar beet	San Carlos, Chile	Santa Laura (−36.460742°, −71.878242°)	4	F-SB1
	Los Ángeles, Chile	Santa Isabel (−37.43337°, −72.29983°)	2	F-SB2
	Los Ángeles, Chile	Luciana (−37.43191°, −72.31375°)	2	F-SB2
Stored sugar beet	San Carlos, Chile	Santa Laura (−36.460742°, −71.878242°)	4	S-SB1
Coconut sugar	Chile	NN	1	Coco-S
	Chile	NO	1	Coco-S
	Colombia	SL	1	Coco-S
	United States	MT	1	Coco-S
Brown Cane sugar	Türkiye	KY	1	Br-CaneS
	Chile	IA	1	Br-CaneS
	Colombia	DL	1	Br-CaneS
	Colombia	CO	1	Br-CaneS
	Perú	DU	1	Br-CaneS
Proportion mixtures between coconut and brown cane sugar (% w/w)		NN:KY 15:15	1	Mix-S
		NO:IA 20:10	1	Mix-S
		SL:DL 10:20	1	Mix-S
		MY:CO 25:5	1	Mix-S
		Total number of samples	**25**	Mix-S

beets from San Carlos (F-SB1) and Los Ángeles (F-SB2) were selected, washed, cut into slices, freeze-dried, pulverized in a mortar, and stored at −80 °C. In parallel, to establish the robustness of the segmented analysis and assess the discrimination capacity of the chemometric models by adding a new variable, four samples obtained from San Carlos (S-SB1) were stored in a growth chamber under controlled conditions at 20 °C (without humidity control) for 1 month. After this period, the samples were collected, washed, cut into slices, freeze-dried, pulverized in a mortar, and stored at −80 °C.

In addition, nine commercial sugars, four coconut sugars (Coco-S), and five brown cane sugars (Br-CaneS) were purchased from local markets in different countries. These commercial sugars were pulverized in a mortar and, subsequently, intentionally mixed in different proportions for simulated adulterations (Mix-S) (Table 1).

2.3. Extraction of carbohydrates from the sugar beet samples

The extraction of carbohydrates from F-SB1, F-SB2, and S-SB1 samples was conducted using the methanol/water protocol reported by Yang et al. (2012) with some modifications. Briefly, 150 mg freeze-dried sugar beet was suspended in 1000 μL cold (4 °C) MeOH:H_2O (1:1 v/v) solution, vortexed for 1 min, and sonicated for 15 min. Following centrifugation at 17,000 × *g* for 10 min, the supernatant was collected and transferred to a 2-mL Eppendorf tube. The extraction process for each sample was performed twice, and subsequently, both supernatants were combined, freeze-dried, and stored at −80 °C until analyzed.

2.4. NMR sample preparation

The sugar beet extracts were dissolved in 800 μL of a D_2O solution containing 100 mM K_2HPO_4/KH_2PO_4 pH 7.4 as a buffer and TSP-d_4 0.1% w/w as an internal standard. For Coco-S, Br-CaneS, and Mix-S, 300 mg of each sample was dissolved using the same solution. The samples were centrifuged at 17,000 × *g* for 5 min, and 600 μL supernatant was transferred to a 5-mm NMR tube.

2.5. NMR acquisition

All ^{1}H NMR spectra were acquired using a Bruker 400 MHz spectrometer model Ascend^{TM} (Bruker Biospin, Germany) operating at 400.13 MHz and equipped with a PABBI 1H/D-BB-Z-GRD liquid probe. For the ^{1}H NMR spectra, 96 number of scans were recorded for each sample with 65,536 data points over a spectral width of 6393.862 Hz, with an acquisition time of 1.26 s, receiver gain of 57.0, and relaxation delay of 1.00 s using the zg30 pulse sequence at 293 K. A representative ^{1}H NMR spectrum was obtained for each sample by averaging these 96 scans.

2.6. NMR data preprocessing using chemometric tools

2.6.1. Processing ^{1}H NMR dataset

All ^{1}H NMR spectra were manually phased, baseline corrected, and referenced to the TSP-d_4 resonance signal ($\delta_H = 0.00$) using MestreNova v.12.0 (MestreLab Research, Spain) software and imported into MATLAB R2021a (TheMathWorks Inc., Natick, MA, USA) as a spectral data matrix. The spectral regions of δ_H = 3.32–3.38 and δ_H = 4.68–5.17 caused by the presence of residual resonance signals from methanol and water, respectively, were excluded from the data. The spectral regions with chemical shifts $< \delta_H = 0.20$ and $> \delta_H = 9.50$ were also removed. The final ^{1}H NMR dataset comprised a data matrix with 25 spectra in rows with δ_H = 35,886 values in columns. To minimize any variability due to chemical shifts that may affect the chemometrics models, the resonance misalignments were corrected using the *icoshift* algorithm (Savorani et al., 2010), and rows of the dataset were normalized to the total area. The corrected and normalized ^{1}H NMR dataset (processed ^{1}H NMR dataset) was subjected to two independent preprocessing methods before the unsupervised and supervised analysis, i.e., ^{1}H NMR data binning and segmented analysis via MCR–ALS, respectively.

2.6.2. ^{1}H NMR data binning

The processed ^{1}H NMR dataset was used to make a binning dataset with a bucket width of 0.04 ppm using MestreNova v.12.0. The result consisted of a matrix with 25 spectra in rows and 219 buckets in columns.

2.6.3. Segmented analysis via MCR–ALS

Alternative to data binning, a segmented analysis via MCR–ALS was applied to the processed ^{1}H NMR dataset using the methodology developed by Pérez et al. (2020). The ^{1}H NMR dataset was manually divided into 29 subarrays (segments) to maintain the multiplicity of the resonance signals. Each subarray encompassed one or more resonance signals with distinct intensities in relation to the spectral noise. The dimensions of the 29 subarrays are reported in detail in Table S1. The 29 subarrays were independently analyzed using MCR–ALS GUI 2.0 Toolbox (Jaumot et al., 2015) via MATLAB R2021a. Every subarray was resolved via MCR–ALS using the single value decomposition (SVD) to find the optimal number of components. Non-negativity constraints were applied to the **C** and $\mathbf{S^T}$ profiles in the ALS optimization. The quality of all MCR–ALS models was determined by evaluating the values of lack of fit (LOF) and percentage of explained variance. The combination of the 29 concentration profile matrices with different number of components produced a combined **C** matrix with 25 samples in rows and 37 components in columns.

2.7. Unsupervised and supervised analyses

Exploratory analysis by PCA using data binning and combined **C** matrix as an independent **X** matrix, composed by the 25 samples (described in Table 1), was applied to evaluate the maximization sample clustering according to the intrinsic variance in the dataset. In addition, mean centering in combination with variance or *Pareto* scaling was used as a pretreatment for the different PCA models to evaluate the influence of the pretreatment on the sample clustering.

To improve the separation between the sample groups and evaluate the variables that produced considerable differences in the discrimination, two PLS–DA models were applied to the combined **C** matrix. The first PLS–DA model was used to discriminate between stored sugar beets and fresh sugar beets and between the different geographical regions of cultivation, considering 12 samples. The second PLS–DA model was applied to discriminate between coconut sugar, brown cane sugar, and their simulated adulterations, with a total of 13 samples. For each PLS–DA model, mean centering with variance scaling was used as a pretreatment.

All the PCA and PLS–DA models were constructed using PLS_Toolbox 9.0 (Eigenvector Research Inc.) via MATLAB R2021a. The quality of the PCA and PLS–DA models was evaluated using the Venetian blinds cross-validation method with 5 and 6 blinds, respectively. The optimal number of components (and latent variables [LVs]) was determined based on the percentage of explained variance. To identify the most discriminant variables from the PLS–DA models, the values from the variable importance in the projection (VIP) were analyzed. The variables with a VIP of ≥ 1 were considered relevant for discrimination. To assess the statistical significance of each relevant variable, a one-way analysis of variance (ANOVA) test with Bonferroni correction was applied to adjust the significance level, and for pairwise differences, t-tests were performed using OriginPro 2016 (OriginLab Corporation, Northampton, MA, USA). VIP values with $p < 0.05$ statistical significance were recognized as discriminant.

3. Results and discussion

The processed ^{1}H NMR dataset of the 25 samples is shown in Fig. 1A.

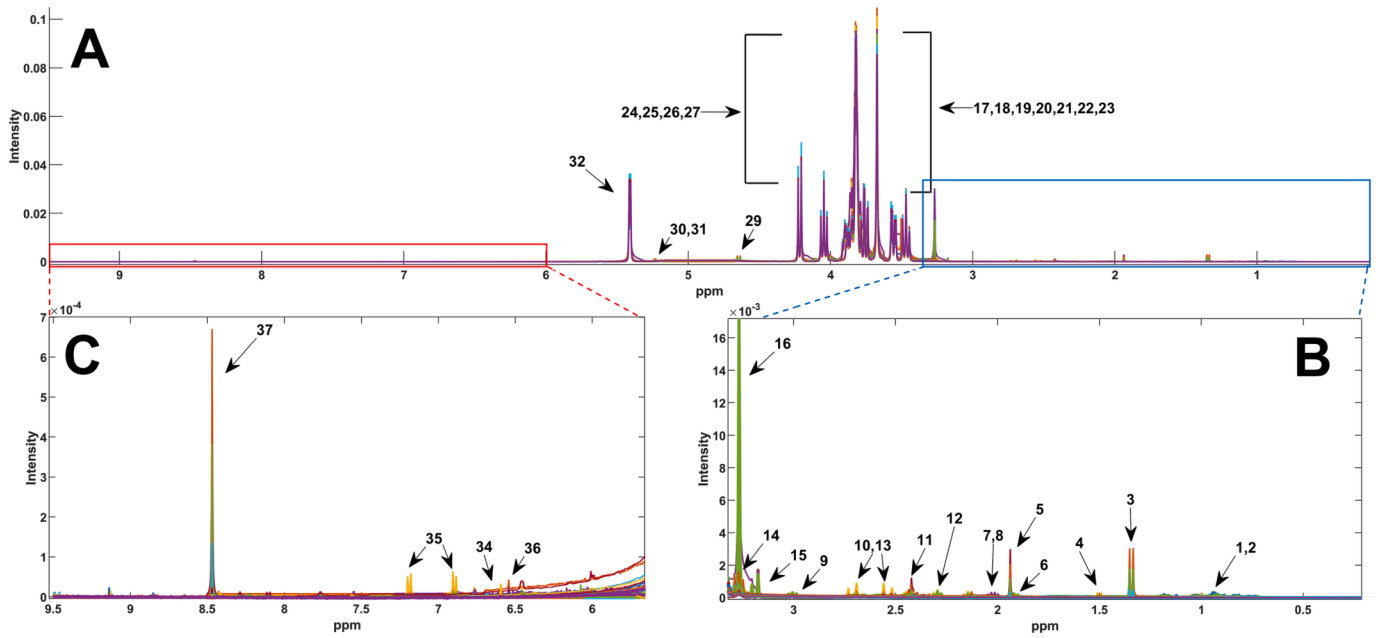


Fig. 1. Processed ^{1}H NMR dataset (400.13 MHz) from the 25 samples (carbohydrates extracted from eight fresh sugar beets, four storage sugar beets, four coconut sugars, five brown cane sugars, and four simulated adulterated samples) used for the segmented analysis via MCR–ALS. A) Complete range of resonance signals δ_H = 0.20–9.50. B) Expanded aliphatic spectral region δ_H = 0.20–3.50. C) Expanded aromatic/aldehyde spectral region δ_H = 6.00–9.50. Numbers 1–37 represent resonance signals (Table 3) associated to: (1): Leucine; (2): Valine + Isoleucine; (3): Lactate; (4): Alanine; (5): Acetate; (6–9-14): GABA; (7–8): Glutamate; (10–13): Acetate + Malate; (11): Pyroglutamate; (12): Glutamine; (15): Choline; (16): Betaine; (17–18-19–20-21–22-23–25-27–32): Sucrose; (24–26): Fructose; (29): β-glucose; (30): α-glucose; (31): α-xylose; (34): *trans*-aconitate; (35): Tyrosine; (36): Fumarate; (37): Formate.

Similarities were observed mainly in the spectral region of carbohydrates with δ_H = 3.50–6.00. In addition, lower resonance signals in the spectral regions of aliphatic compounds with δ_H = 0.20–3.50 and those of aromatic/aldehyde compounds with δ_H = 6.00–9.50 were detected in all the ^{1}H NMR spectra (Fig. 1B and 1C), suggesting similar spectroscopic profiles in all the samples. The identification of these resonance signals in the processed ^{1}H NMR dataset will be discussed on detail in section 3.4.

3.1. Resolution of resonance signals via MCR–ALS

The 29 manually selected subarrays across the full range of resonances (δ_H = 0.20–9.50) were independently resolved via MCR–ALS. For example, two different sub-arrays in the processed ^{1}H NMR dataset, containing resonance signals in the ranges of δ_H = 4.16–4.25 and δ_H = 1.80–1.96 are presented in Fig. S1A and S1B, respectively. The optimal number of components required to resolve **C** and $\mathbf{S}^{\mathbf{T}}$ profiles for each resonance signal was determined using the SVD method using MCR–ALS. The resonance signal of the first subarray was successfully resolved with only one component (Fig. S1A), whereas for the second subarray, two components were found to be optimal for accurately resolve resonance signals with good symmetry with a certain degree of overlap (Fig. S1B). For this preprocessing, 1–3 components were required to describe the **C** and $\mathbf{S}^{\mathbf{T}}$ profiles for each subarray, obtaining a total of 45 components with a convergence between 4 and 500 iterations. Of these 45 components, 8 were discarded because they contained only noise or were incorrectly resolved. The remaining 37 components had an LOF of ~ 4.46% and a total explained variance of 98.4%. Owing to their chemical shifts, components 1–17, 18–33, and 34–37 described resonance signals associated with the aliphatic, carbohydrate, and aromatic/aldehyde spectral regions, respectively. More precise representation of all the chemical shifts according to each **C** and $\mathbf{S}^{\mathbf{T}}$ profiles is shown in Table S2.

3.2. Effect of the pretreatment on the exploratory analysis using PCA

3.2.1. Binning data

The PCA score plot, obtained using mean centering with variance scaling as a pretreatment, is shown in Fig. 2A. The first two principal components (PCs) explained 45.29% of the total variance. No clear clustering trend was observed with respect to the sugar source (Fig. 2A); however, separations between S-SB1 and F-SB2 and Br-CaneS in PC1 (25.79% explained variance) were observed. PC2 (19.50%) showed lesser differences between Br-CaneS and Coco-S compared to PC1. In this case, the data binning did not allow a good differentiation between the sample groups. Using mean centering and *Pareto* scaling as a pretreatment, the total explained variance could be increased, based on the contribution of lower-intensity resonance signals; providing equal weighting to each variable (van den Berg et al., 2006). Thus, the differentiation was not biased toward variables with greater magnitude, as seen in the case of resonance signals present in the carbohydrate spectral region (Fig. 1A), which maximized the differentiation as in the case of the S-SB1 samples (Fig. 2B). In this case, the first two PCs explained 79.94% of the variance. The first source of variation was attributed to storage, in which S-SB1 differed from F-SB2, Coco-S, Br-CaneS, and Mix-S in PC1, (56.13% of the explained variance). PC2 (23.81%) demonstrated greater differentiation between F-SB1 and F-SB2 with respect to S-SB1, Coco-S, and Mix-S than PC1. Despite these observations, satisfactory results in differentiating sugar beet extracts and commercial sugars were not achieved.

3.2.2. Combined C matrix

PCA applied to the combined **C** matrix using mean centering and variance scaling as a pretreatment, is shown in Fig. 2C. The first two PCs explained 56.29% of the total variance. Variance scaling uses the standard deviation of each variable as a scaling factor (Ebrahimi et al., 2017), thereby rendering each variable equally important (same variance). However, erroneous estimations can be observed for those variables that do not contain chemical information or a certain degree of

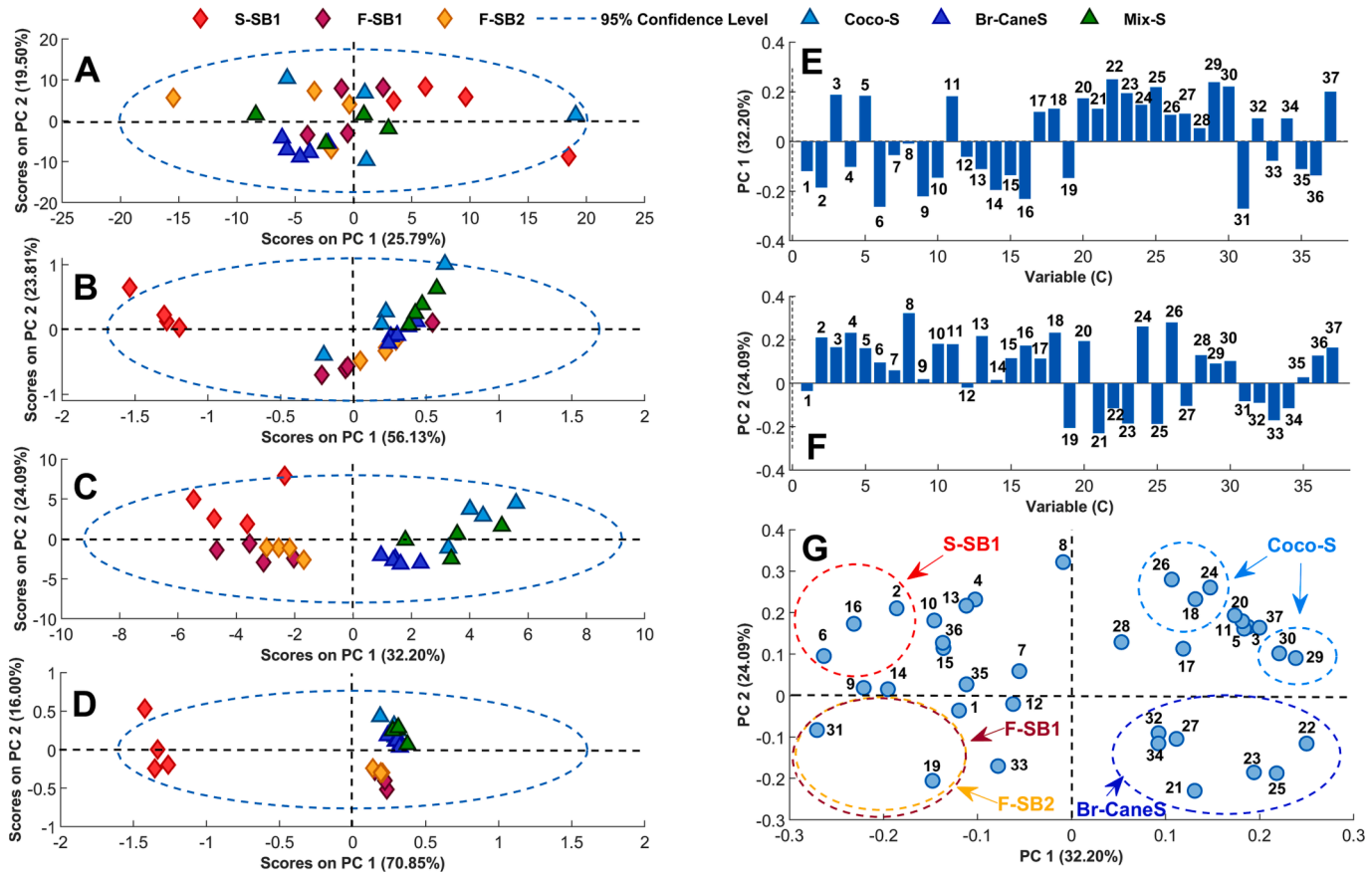


Fig. 2. PCA plots derived from the binning data (A, B) and the combined C matrix obtained via MCR–ALS (C, D) for the differentiation of sugar beet extracts and commercial sugars. A) PCA score plot (PC1 vs PC2) using mean centering and variance scaling on binning data. B) PCA score plot (PC1 vs PC2) using mean centering and *Pareto* scaling on binning data. C) PCA score plot (PC1 vs PC2) using mean centering and variance scaling on the combined C matrix. D) PCA score plot (PC1 vs PC2) using mean centering and *Pareto* scaling on the combined C matrix. E) and F): Loading bar plots of PC1 and PC2 from PCA described in C), respectively. G) Scatter loadings plot of PC1 vs PC2 of PCA described in C). Variables were represented as a bar plot and blue dots in which each number represented a concentration profile (C) that is associated to its respective spectral profile (S^T). S-SB1: storage sugar beet from San Carlos (red); F-SB1: fresh sugar beet from San Carlos (brown); F-SB2: fresh sugar beet from Los Ángeles (yellow); Coco-S: coconut sugar (light blue); Br-CaneS: brown cane sugar (blue); Mix-S (green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

overlap, as in the case of data binning**.** While binning data minimizes variations caused by chemical shifts, it also results in loss of spectral resolution and superposition of resonance signals in the same bin, making sample clustering challenging (Fig. 2A and 2B). However, this does not occur with the combined **C** matrix**,** because each **C** profile is described by an $\mathbf{S^T}$ profile regardless of the degree of overlap. The spectral noise is modelled in an independent matrix of residuals, and even small variations within the variables can be detected, thereby maximizing clustering in PCA (Fig. 2C). Thus, with mean centering and *Pareto* scaling used as a pretreatment (Fig. 2D), MCR–ALS preprocessing achieved clustering of similar samples. These results agreed with those reported by Pérez et al. (2020) and Khakimov et al. (2020) in zebrafish and human urine samples, respectively. MCR–ALS maximized the between-group variability and decreased the within-group variability. Therefore, different pretreatments can be applied depending on the objectives of the study.

3.2.3. Differentiation of sugar beet extracts and commercial sugars using the combined C matrix

The score plot in Fig. 2C shows that the first source of variation corresponded with sugar beet extracts or commercial sugar samples, in which separation between S-SB1, F-SB1, and F-SB2 and Coco-S, Br-CaneS, and Mix-S was achieved with 32.20% of the variance explained. The second source of variation allowed the differentiation between the stored and fresh sugar beet samples S-SB1, F-SB1, and F-SB2, as well as differentiation of Coco-S and Br-CaneS with 24.09% of variance explained. Notably, the clustering of samples based on the sugar source was maximized when using the combined **C** matrix**.** Also, the simulated adulteration samples (Mix-S) were separated from the Coco-S and Br-CaneS clusters, with one sample of Coco-S clustering with the group of Mix-S samples. Most samples remained inside the 95% confidence level, except for one sample from S-SB1. The samples that showed Q residuals over the 95% confidence level and did not exhibit significant improvement in the percentage of explained variance following their exclusion were not considered outliers.

The variables that contributed to the clustering according to the sugar source can be observed in the loading plots of both PCs in Fig. 2E and 2F. These loadings are shown as a bar plot, in which each component represents **C** that was described by its respective $\mathbf{S^T}$ profile. The variables 3, 5, 11, 17–32 (except 19 and 31), 34, and 37 were most relevant for the clustering of the commercial sugars Coco-S, Br-CaneS, and Mix-S (Fig. 2E). The variables 1–16 (except 3, 5, and 11), 19, and 31 contributed to the clustering of the stored and fresh sugar beet samples S-SB1, F-SB1, and F-SB2, followed by less relevant variables 33, 35, and 36. Regarding the spectral regions of each variable (Table S2), the carbohydrates region made the greatest contribution to the differentiation of commercial sugars from sugar beet extracts, whereas the aliphatic region exhibited a high relevance for differentiating sugar beet extracts under storage conditions, except for certain resonance signals. The spectral region of aromatic/aldehyde compounds also exhibited specific

resonance signals that contributed to the differentiation of sugar beet extracts and commercial sugars. The loading plot for PC2 (Fig. 2F) illustrated that the variables from the carbohydrate spectral region 18, 20, 24, and 26 were more relevant for clustering S-SB1, Coco-S, and some samples of Mix-S, whereas the variables 19, 21, 22, 23, 25, 27, 31, 32, 33 and 34 were relevant for F-SB1, F-SB2, Br-CaneS, Mix-S, and one sample of Coco-S, implying the possible influence of resonance signals associated with different sugars contributing to the differentiation. Similarly, the loading plot for PC2 shows that the variables 2–16 (except 1 and 12) of the aliphatic spectral region made a clear contribution to the clustering of S-SB1, Coco-S and some Mix-S samples. According to the scatter plot of the loadings (Fig. 2G), variables 18, 24, 26, 29, and 30 characterized the clustering of Coco-S samples, while stored samples were characterized by variables 2, 6, and 16. Variables 19 and 31 characterized the samples F-SB1 and F-SB2, while variables 21, 22, 23, 25, 27, 32, and 34 had a high contribution for Br-CaneS.

3.3. Discrimination of a group of samples using PLS–DA

3.3.1. Sugar beet under storage conditions and geographical origin

Samples coming from sugar beet were analyzed by PLS–DA to evaluate differences by storage conditions and geographical origin. The first two LVs explained 55.62% of the total variance. The discrimination of S-SB1 from F-SB1 and F-SB2 was achieved with the first LV (LV1), with 40.27% of the variance explained. This discrimination is depicted in the score plot of the PLS–DA model in Fig. 3A. The PCA score plot revealed that the same differences were obtained with PC1, with 41.47% of the variance explained (Fig. S3A). Discrimination based on geographical origin (F-SB1 from F-SB2) was possible with LV2 (15.35% of explained variance). In the PCA score plot, a similar differentiation was observed but to a lesser extent in the fourth PC, which explained 7.64% of the variance (Fig. S3B).

The VIP plot (Fig. 3B) revealed that the highest contributions associated with the discrimination between S-SB1 and F-SB1 and F-SB2 were from the variables 2, 3, 8, 10, 16, 18, 21, 22, 23, 25 and 36. After the analysis of the statistical significance of each relevant variable (VIP of $\geq$ 1), only variables 2, 8, 10, 16, and 36 were observed to contribute to differentiate S-SB1 from the fresh samples. This was expected due to the other variables being related to the resonance signals of the sucrose molecule. Table 2 summarized the VIP values obtained by the PLS–DA model for these vaiables. Based on the identification of resonances signals (Table 3), these variables were associated with valine, isoleucine, glutamate, citrate, malate, betaine, and fumarate. The score plot for LV1

Table 2
Variables that contribute to the discrimination of sugar beet extracts and commercial sugars associated with the VIP values ($p < 0.05$) obtained by PLS-DA.

Source of sugar	Variable (C) and (S^T)	Compound	VIP value
S-SB1	2	Valine-Isoleucine	1.50
	8	Glutamate	1.41
	10	Citrate-Malate	1.30
	16	Betaine	1.30
	36	Fumarate	1.18
F-SB1	1	Leucine	1.82
Coco-S	3	Lactate	1.05
	5	Acetate	1.03
	11	Pyroglutamate	1.03
	24	Fructose	1.31
	26	Fructose	1.26
	29	β-glucose	1.42
	30	α-glucose	1.24
	37	Formate	1.07
Br-CaneS	31	α-xylose	1.22
	32	Sucrose	1.03

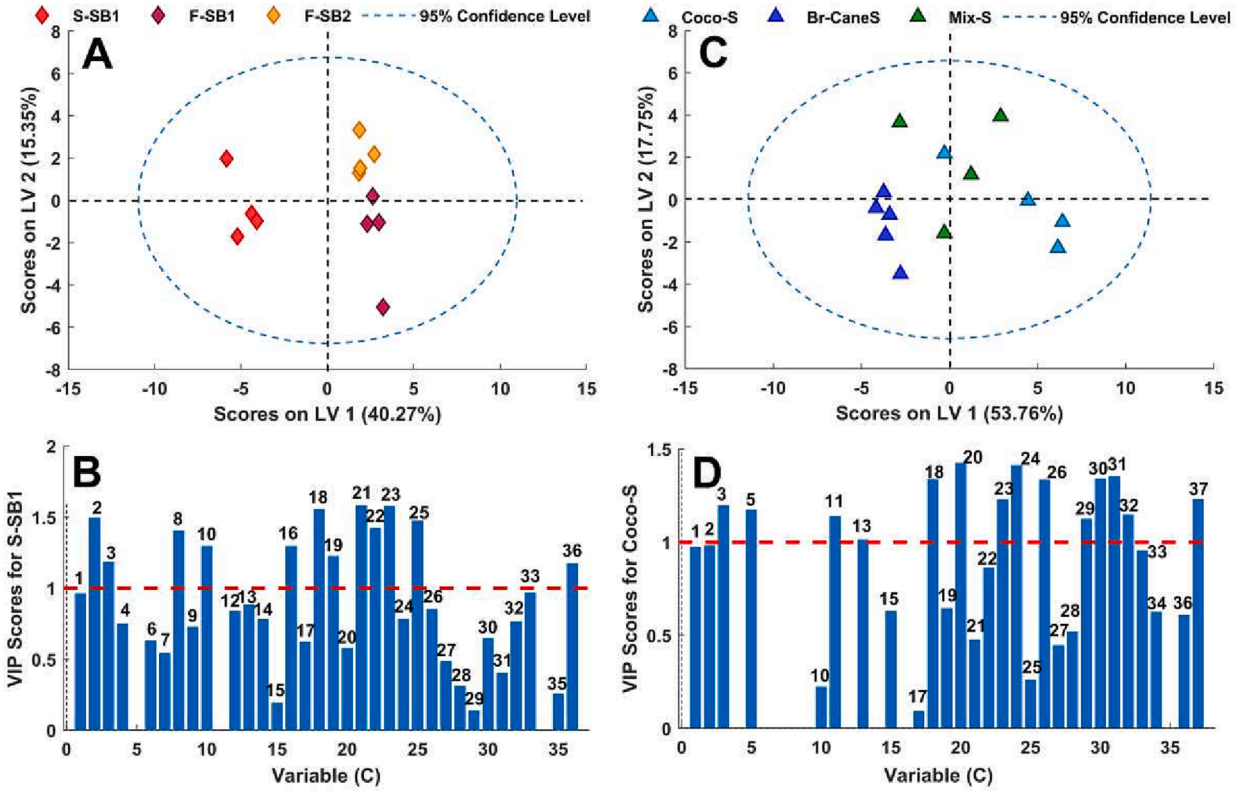


Fig. 3. PLS–DA plots derived from the combined C matrix obtained via MCR–ALS for the discrimination of sugar beets and commercial sugars. A) PLS–DA score plot (LV1 vs LV2). B) VIP plot for S-SB1. C) PLS–DA score plot (LV1 vs LV2). D) VIP plot for Coco-S. Variables with VIP $\geq$ 1 were considered relevant for discrimination. S-SB1: storage sugar beet from San Carlos (red); F-SB1: fresh sugar beet from San Carlos (brown); F-SB2: fresh sugar beet from Los Ángeles (yellow). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Identification and assignments of each spectral profile (S^T) obtained by MCR-ALS for ^{1}H NMR dataset from sugar beets and commercial sugars. S-SB1: Storage sugar beet from San Carlos; F-SB1: Fresh sugar beet from San Carlos; F-SB2: Fresh sugar beet from Los Ángeles; Coco-S: Coconut sugar; Br-CaneS: Brown cane sugar; s: singlet; d: doublet; t: triplet; dd: doublet of doublets; dq: doublet of quartets; m: multiplet.

Variable (S^T)	Compound	δ_H in ppm (mult, *J* in Hz)	Group	Detected in sugar beet extracts	Detected in commercial sugars
Amino acids and amino compounds					
2	Valine	1.00 (d, $J = 6.64$)	γCH_3	+	Only in 3 samples of Coco-S
		1.04 (d, $J = 6.86$)	$\gamma' CH_3$		
1	Leucine	0.93 (t, $J = 6.80$-6.30)	δCH_3	+	Only in 3 samples of Coco-S
2	Isoleucine	1.07 (d, $J = 7.06$)	γCH_3	+	Only in 3 samples of Coco-S
4	Alanine	1.50 (d, $J = 7.22$)	βCH_3	+	-
6-9-14	GABA	1.91 (m)	βCH_2	+	-
		2.29 (t, $J = 7.42$)	αCH_2		
		3.01 (t, $J = 7.95$-7.84)	γCH_2		
8-7	Glutamate	2.04 (m)	βCH_2	Only in S-SB1	-
		2.14 (m)	$\beta' CH_2$		
11	Pyroglutamate	2.42 (m)	γCH_2	-	Only in 3 samples of Coco-S
		2.52 (m)	βCH_2		
12	Glutamine	2.44 (m)	γCH_2	Only in F-SB1 and F-SB2	-
15	Choline	3.17 (s)	$N(CH_3)_3$	Only in S-SB1	Only in Coco-S
16	Betaine	3.27 (s)	$N(CH_3)_3$	+	-
35	Tyrosine	6.89 (d, $J = 8.73$)	3,5-CH	Only in 3 samples	-
		7.19 (d, $J = 8.88$)	2,6-CH		
Organic acids					
10-13	Malate	2.37 (dd, $J = 15.1$-9.53)	βCH	Only in S-SB1	Only in 3 samples of Coco-S
		2.68 (dd, $J = 15.4$-3.53)	$\beta' CH$		
5	Acetate	1.94 (s)	CH_3	-	+
10-13	Citrate	2.54 (d, $J = 16.2$)	CH_2	Only in S-SB1	Only in 3 samples of Coco-S
		2.71 (d, $J = 16.3$)	CH_2		
3	Lactate	1.34 (d, $J = 6.93$)	βCH_3	+	+
36	Fumarate	6.54 (s)	$(CH=)_2$	+	Only in 3 samples of Coco-S
37	Formate	8.47 (s)	CH	-	+
34	*trans*-aconitate	6.59 (s)	CH=	-	Only in Br-CaneS
Carbohydrates					
17	Sucrose	3.47 (t, $J = 9.38$)	G_4H	+	+
18		3.56 (dd, $J = 9.97$-3.85)	G_2H	+	+
19		3.67 (s)	F_1H	+	+
20		3.76 (t, $J = 9.56$)	G_3H	+	+
21		3.82 (m)	F_6H	+	+
22		3.85 (m)	G_5H	+	+
23		3.89 (dq, $J = 8.30$-4.01)	F_5H	+	+
25		4.05 (t, $J = 8.76$)	F_4H	+	+
27		4.22 (d, $J = 8.54$)	F_3H	+	+
32		5.41 (d, $J = 3.87$)	G_1H	+	+
24	Fructose	4.01 (m)	C_6H_2	+	+
26		4.11 (m)	C_3H+C_4H	+	+
31	α-xylose	5.20 (d, $J = 3.82$)	CH	+	Only in Br-CaneS
30	α-glucose	5.23 (d, $J = 3.76$)	CH	+	+
29	β-glucose	4.65 (d, $J = 7.93$)	CH	+	+
Not identified					
28	Unknown	4.39 (d, $J = 8.77$)	-	+	+
33	Unknown	5.62 (d, $J = 3.82$)	-	+	+

(Fig. 3A) revealed that storage time affected sugar beet, characterized by an increase in the intensity of concentration profiles associated with amino acids and nitrogenous compounds (Figure S6). This could be explained by a possible tendency toward the hydrolysis of proteins to amino acids and their catabolism via the tricarboxylic acid (TCA) cycle (citrate, malate, and fumarate), thereby increasing the concentration of drought stress–induced *N*-amino compounds (Wedeking et al., 2018). Furthermore, an increase in the intensity of betaine concentration profiles was observed (Figure S6E). This, could be related to a decrease in the intensity of resonance signals associated with sucrose, because its biosynthesis requires energy from sucrose. Regarding discrimination according to geographical region, the VIP plots (Figure S4) indicate the variables with more relevance are associated with the concentration profiles of amino acids. Based on this, only leucine showed significant discrimination for F-SB1 (Figure S6A), suggesting possible differences in soil type per cultivation field, generating variations in the concentration of free amino acids.

3.3.2. *Coconut sugar, brown cane sugar, and simulated adulterations*

A separated analysis was performed for the of commercial brown to identify compounds responsible for discrimination based on botanical source using PLS–DA. The PLS–DA model enabled the discrimination between Coco-S and Br-CaneS, with the first two LVs explaining 71.51% of the total variance. The PLS–DA score plot (Fig. 3C) revealed that the discrimination between Coco-S and Br-CaneS was achieved with LV1, explaining 53.76% of the variance. The same differentiation was possible with PCA, with 54.34% of the explained variance in PC1 (Fig. S3C). LV2 showed that discrimination of Mix-S from Coco-S and Br-CaneS was possible to a certain extent, with 17.75% of the variance explained (Fig. 3C), whereas Mix-S was mostly differentiated from both commercial sugars with PC1 (Fig. S3C). Comparing these results with the PCA score plot of Fig. 2C, it was observed that in all models, the Mix-

S samples were clustered with Coco-S and Br-CaneS, with one sample of Coco-S considered as Mix-S, suggesting a slight difference from the other Coco-S samples. The identification of resonance signals (Table 3) revealed that only the variables 15 (choline), 5 (acetate), 3 (lactate), 37 (formate), 17–32 (sucrose), 24–26 (fructose), 30 (α-glucose), and 29 (β-glucose) were detected in all the samples of Coco-S, whereas some resonance signals associated with amino/organic acids were only detected in three samples of Coco-S. This indicated that the use of the combined **C** matrix obtained from segmented analysis via MCR–ALS represents the pure resonance signals resolved in the processed ^{1}H NMR dataset. This allows, supervised or unsupervised models to present a less biased contribution to each variable and, therefore, be able to differentiate the one sample of Coco-S from the other three Coco-S samples.

The VIP plot (Fig. 3D and Table 2) showed that the contribution associated with the discrimination of Coco-S and Br-CaneS were from the variables 3, 5, 11, 24, 26, 29, 30, 31, 32, and 37. Regarding the identification of resonance signals (Table 3), these variables corresponded to lactate, acetate, pyroglutamate, fructose, β-glucose, α-glucose, α-xylose, sucrose, and formate. The high intensities of variables 3 (lactate), 5 (acetate), and 37 (formate) in Coco-S (Figure S7A, S7B, and S7H) may be partly due to fermentation after harvesting and the Maillard reaction (Bachmann et al., 2022; E. Chen et al., 2021), contributing to its discrimination. The formation of pyroglutamate through the intramolecular cyclization of glutamate or glutamine (Gazme et al., 2019) due to high temperatures and pressure during the production of coconut sugar contributed to its discrimination. In addition, the **C** of the variables 24–26 (fructose), 29 (β-glucose), and 30 (α-glucose) exhibited higher intensities in Coco-S than in Br-CaneS (Figure S7D, S7E, S7F, and S7G). Conversely, the high intensities of **C** associated with resonance signals of sucrose and possibly related to α-xylose contributed to the differentiation of Br-CaneS (Figure S8A and S8B).

3.4. Tentative identification of resonance signals and assignment for sugar beet extracts and commercial sugars through the spectral profiles (S^T) obtained via MCR–ALS

The 37 $\mathbf{S}^T$ profiles obtained via MCR–ALS according to each subarray were evaluated to determine the spectroscopic parameters as chemical shifts (δ_H), multiplicity, and coupling constants (J) using MestreNova v.12.0. All the resolved $\mathbf{S}^T$ profiles are presented in Figure S2. The spectral regions of aliphatic, carbohydrate, and aromatic/aldehyde compounds were determined by comparing the spectroscopic parameters with values from previously reported studies (Bachmann et al., 2022; Boffo et al., 2012; Dimitrakopoulou et al., 2021; Giampaoli et al., 2021; Saviano et al., 2019; Wedeking et al., 2018; Yang et al., 2012), using reference library Chenomx NMR suite 8.3, and the online available Human Metabolome Database (HMDB) (Wishart et al., 2009). The latter was used for some primary compounds present also in plant metabolome. Overall, 22 polar compounds were identified in the sugar beet extracts and commercial sugar samples, i.e., 9 amino acids, 2 amino compounds, 7 organic acids, and 4 sugars. The identification, assignment, chemical shifts, coupling constants, and multiplicities of all the compounds are presented in detail in Table 3. The identification of the resonance signals through $\mathbf{S}^T$ profiles was consistent with the results of the scores and loading plots obtained with PCA (Fig. 2C, 2E, and 2F). The specific resonance signals for sugar beet extracts and commercial sugars, i.e., variables 4 (alanine), 6, 9, 14 (gamma-aminobutyric acid [GABA]), 8, 7, (glutamate), 11 (pyroglutamate), 12 (glutamine), 16 (betaine), 35 (tyrosine), 5 (acetate), 37 (formate), and 34 (*trans*-aconitate), demonstrated accurate contribution to the clustering and differentiation according to the sugar source.

3.4.1. Carbohydrates

As anticipated, most of the intense resonance signals present in the ^{1}H NMR dataset (Fig. 1A) corresponded to the resonance signals of sucrose in the carbohydrate spectral region. Glucose anomers (α and β), α-xylose, and sucrose were represented by the doublets of protons bonded to the anomeric carbon at $\delta_H = 5.23$ (d, $J = 3.76$ Hz), $\delta_H = 4.65$ (d, $J = 7.93$ Hz), $\delta_H = 5.20$ (d, $J = 3.82$ Hz), and $\delta_H = 5.41$ (d, $J = 3.87$ Hz), respectively. Resonance signals from fructose were detected as multiplets at $\delta_H = 4.01$ and 4.11.

3.4.2. Amino acids and amino compounds

The compounds in the aliphatic region associated with the branched-chain amino acid signals of valine $\delta_H = 1.00$, 1.04 (d, $J = 6.64$, 6.86 Hz), leucine $\delta_H = 0.93$ (t, $J = 6.80$–6.30 Hz), and isoleucine $\delta_H = 1.07$ (d, $J = 7.06$ Hz) were identified in S-SB1, F-SB1, F-SB2, and only in three Coco-S samples. Alanine $\delta_H = 1.50$ (d, $J = 7.22$ Hz) and GABA $\delta_H = 1.94$ (m), 2.29 (t, $J = 7.42$ Hz), and 3.01 (t, $J = 7.95$–7.84 Hz) were detected only in sugar beet extracts. The resonance signal associated with betaine $\delta_H = 3.27$ (s) was found only in the sugar beet extracts, as betaine is a specific molecule of sugar beet (Bachmann et al., 2022; Palmonari et al., 2020). The multiplets of glutamate at $\delta_H = 2.04$, 2.14 were detected only in S-SB1, and the multiplet of glutamine $\delta_H = 2.44$ was found in F-SB1 and F-SB2. Pyroglutamate $\delta_H = 2.42$, 2.52 (m) was present in only three samples of Coco-S, possibly due to the heat-degradation of glutamate or glutamine during sugar production (Wedeking et al., 2018). The resonance signal belonging to choline $\delta_H = 3.17$ (s) was detected in S-SB1 and Coco-S.

3.4.3. Organic acids and derivates

Malate $\delta_H = 2.37$, 2.68 (dd, $J = 15.1$–9.53, 15.4–3.53 Hz), acetate $\delta_H = 1.94$ (s), citrate $\delta_H = 2.54$, 2.71 (d, $J = 16.2$–16.3 Hz), and lactate $\delta_H = 1.34$ (d, $J = 6.93$ Hz) were identified as the most prominent organic acids in Coco-S (except for lactate), suggesting the occurrence of mixed acid fermentation between the harvesting of the sugar juice and heating (Bachmann et al., 2022). In sugar beet extracts, only malate and citrate were detected in S-SB1, whereas lactate was identified in F-SB1, F-SB2, and S-SB1. In the aromatic/aldehyde spectral region, tyrosine at $\delta_H = 6.89$–7.19 (d, $J = 8.73$, 8.88 Hz) was detected in only three samples of the sugar beet extracts, i.e., one fresh and two stored samples. Furthermore, the resonance signal of fumarate $\delta_H = 6.54$ (s), an intermediate of the TCA cycle with citrate and malate (Wedeking et al., 2018), was detected in S-SB1, F-SB1, F-SB2, and three samples of Coco-S. Formate $\delta_H = 8.47$ (s) was identified only in Coco-S and Br-CaneS samples. The identification of formate and acetate only in Coco-S and Br-CaneS could be related to the formation of these compounds due to the degradation of sugars at high temperatures and the Maillard reaction (Chen et al., 2021). Finally, the presence of *trans*-aconitate $\delta_H = 6.59$ (s) was detected only in Br-CaneS, consistent with the results of Bachmann et al. (2022) and Palmonari et al. (2020). Aconitate is a specific acid produced by sugar cane, and its predominant form is *trans*-aconitate (Montoya, Londono, Cortes, & Izquierdo, 2014), formed by the isomerization of *cis*-aconitate.

4. Conclusion

The application of a segmented analysis via MCR–ALS to ^{1}H NMR spectral data was effective in discriminating sugar beet extracts and commercial sugars with respect to botanical source, geographical region, and storage. This had not been possible using other preprocessing methods such as data binning. In addition, the discrimination was successful without the loss of chemical information required for identifying compounds. The spectral assignment, based on $\mathbf{S}^T$ profiles using MCR–ALS, allowed the identification of 22 compounds present in the sugar beet extracts and commercial sugars. Valine, isoleucine, glutamate, betaine, acetate, malate, and fumarate were responsible for discriminating stored from fresh sugar beet, whereas lactate, acetate, pyroglutamate, fructose, α-glucose, β-glucose, and formate were responsible for discriminating coconut sugar from brown cane sugar. Finally, this study demonstrated that sugar from different sources, with

similar spectroscopic profiles and which are dominated by sucrose, can be differentiated independent of their degree of spectral overlap. The use of MCR–ALS enables greater selectivity in chemometric models. It is important to note that the number of samples used for the analysis was small, and more samples are required in both groups (sugar beet and commercial sugars) to improve robustness of the models and ensure that the methodology is more generalizable in future.

CRediT authorship contribution statement

Cristian A. Fuentes: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft. **Mecit Halil Öztop:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. **Macarena Rojas-Rioseco:** Formal analysis, Methodology, Writing – review & editing. **Martín Bravo:** Conceptualization, Methodology, Writing – review & editing. **Aylin Özgür Göksu:** Conceptualization, Methodology, Writing – review & editing. **Marena Manley:** Methodology, Supervision, Writing – review & editing. **Rosario del P. Castillo:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be accesible from SuChAQuality web site by giving the data repository link.

Acknowledgments

The authors thank FONDECYT 1221387 project (ANID, Chile), National Federation of Beet Growers (FENARE, Chile) for supplying the Chilean sugar beet samples. Cristian Fuentes thanks the Postgraduate Direction of the University of Concepcion PhD scholarship.

Funding

This study received funding from the European Union's Horizon 2020 Research and Innovation program–MSCA RISE under grant agreement #101008228 and FONDECYT 1221387 (ANID, Chile).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2023.136817.

References

Alam, T. M., & Alam, M. K. (2004). Chemometric analysis of NMR spectroscopy data: A review. *Annual Reports on NMR Spectroscopy, 54*, 41–80. https://doi.org/10.1016/S0066-4103(04)54002-4

Arro, J., Park, J. W., Wai, C. M., VanBuren, R., Pan, Y. B., Nagai, C., da Silva, J., & Ming, R. (2016). Balancing selection contributed to domestication of autopolyploid sugarcane (Saccharum officinarum L.). *Euphytica, 209*(2), 477–493. https://doi.org/10.1007/s10681-016-1672-8

Bachmann, R., Horns, A. L., Paasch, N., Schrieck, R., Weidner, M., Fransson, I., & Schrör, J. P. (2022). Minor metabolites as chemical marker for the differentiation of cane, beet and coconut blossom sugar. From profiling towards identification of adulterations. *Food Control, 135*(December 2021). https://doi.org/10.1016/j.foodcont.2022.108832

Boffo, E. F., Tavares, L. A., Tobias, A. C. T., Ferreira, M. M. C., & Ferreira, A. G. (2012). Identification of components of Brazilian honey by 1H NMR and classification of its botanical origin by chemometric methods. *LWT, 49*(1), 55–63. https://doi.org/10.1016/j.lwt.2012.04.024

Caligiani, A., Acquotti, D., Palla, G., & Bocchi, V. (2007). Identification and quantification of the main organic components of vinegars by high resolution 1H NMR spectroscopy. *Analytica Chimica Acta, 585*(1), 110–119. https://doi.org/10.1016/j.aca.2006.12.016

Cavallini, N., Savorani, F., Bro, R., & Cocchi, M. (2021). A Metabolomic Approach to Beer Characterization. *Molecules, 26*(5), 1–15. https://doi.org/10.3390/molecules26051472

Chen, E., Song, H., Zhao, S., Liu, C., Tang, L., & Zhang, Y. (2021). Comparison of odor compounds of brown sugar, muscovado sugar, and brown granulated sugar using GC-O-MS. *LWT, 142*(January), Article 111002. https://doi.org/10.1016/j.lwt.2021.111002

Chen, E., Zhao, S., Song, H., Zhang, Y., & Lu, W. (2022). Analysis and Comparison of Aroma Compounds of Brown Sugar in Guangdong, Guangxi and Yunnan Using GC-O-MS. *Molecules, 27*(18). https://doi.org/10.3390/molecules27185878

Chen, J. Y., Chen, X. W., Lin, Y. Y., Yen, G. C., & Lin, J. A. (2021a). Authentication of dark brown sugars from different processing using three-dimensional fluorescence spectroscopy. *LWT, 150*(December 2020), Article 111959. https://doi.org/10.1016/j.lwt.2021.111959

Chen, J. Y., Chen, X. W., Lin, Y. Y., Yen, G. C., & Lin, J. A. (2021b). Authentication of dark brown sugars from different processing using three-dimensional fluorescence spectroscopy. *LWT, 150*(June), Article 111959. https://doi.org/10.1016/j.lwt.2021.111959

De Juan, A., Jaumot, J., & Tauler, R. (2014). Multivariate Curve Resolution (MCR). Solving the mixture analysis problem. *Analytical Methods, 6*(14), 4964–4976. https://doi.org/10.1039/c4ay00571f

Dimitrakopoulou, M. E., Matzarapi, K., Chasapi, S., Vantarakis, A., & Spyroulias, G. A. (2021). Nontargeted 1H NMR fingerprinting and multivariate statistical analysis for traceability of Greek PDO Vostizza currants. *Journal of Food Science, 86*(10), 4417–4429. https://doi.org/10.1111/1750-3841.15873

Ebrahimi, P., Viereck, N., Bro, R., & Engelsen, S. B. (2017). Modern Magnetic Resonance. *Modern Magnetic Resonance, 1–20*. https://doi.org/10.1007/978-3-319-28275-6

Gazme, B., Boachie, R. T., Tsopmo, A., & Udenigwe, C. C. (2019). Occurrence, properties and biological significance of pyroglutamyl peptides derived from different food sources. *Food Science and Human Wellness, 8*(3), 268–274. https://doi.org/10.1016/j.fshw.2019.05.002

Giampaoli, O., Sciubba, F., Conta, G., Capuani, G., Tomassini, A., Giorgi, G., Brasili, E., Aureli, W., & Miccheli, A. (2021). Red beetroot's nmr-based metabolomics: Phytochemical profile related to development time and production year. *Foods, 10*(8), 1–12. https://doi.org/10.3390/foods10081887

Jaumot, J., de Juan, A., & Tauler, R. (2015). MCR-ALS GUI 2.0: New features and applications. *Chemometrics and Intelligent Laboratory Systems, 140*, 1–12. https://doi.org/10.1016/j.chemolab.2014.10.003

Kamerling, J. P., & Vliegenthart, J. F. G. (2021). Chapter 4. Carbohydrates. *Mass Spectrometry*, 175–264. https://doi.org/10.1515/9783112418123-007

Kenter, C., & Hoffmann, C. M. (2009). Changes in the processing quality of sugar beet (Beta vulgaris L.) during long-term storage under controlled conditions. *International Journal of Food Science and Technology, 44*(5), 910–917. https://doi.org/10.1111/j.1365-2621.2007.01641.x

Khakimov, B., Mobaraki, N., Trimigno, A., Aru, V., & Engelsen, S. B. (2020). Signature Mapping (SigMa): An efficient approach for processing complex human urine 1H NMR metabolomics data. *Analytica Chimica Acta, 1108*, 142–151. https://doi.org/10.1016/j.aca.2020.02.025

Longobardi, F., Innamorato, V., Di Gioia, A., Ventrella, A., Lippolis, V., Logrieco, A. F., Catucci, L., & Agostiano, A. (2017). Geographical origin discrimination of lentils (Lens culinaris Medik.) using1H NMR fingerprinting and multivariate statistical analyses. *Food Chemistry, 237*, 443–448. https://doi.org/10.1016/j.foodchem.2017.05.159

Masetti, O., Sorbo, A., & Nisini, L. (2021). Nmr tracing of food geographical origin: The impact of seasonality, cultivar and production year on data analysis. *Separations, 8*(12). https://doi.org/10.3390/separations8120230

Montoya, G., Londono, J., Cortes, P., & Izquierdo, O. (2014). Quantitation of trans -Aconitic Acid in Different Stages of the Sugar-Manufacturing Process. *Journal of Agricultural and Food Chemistry, 62*(33), 8314–8318. https://doi.org/10.1021/jf5008874

Munyai, R., Raletsena, M. V., & Modise, D. M. (2022). LC-MS Based Metabolomics Analysis of Potato (Solanum tuberosum L.) Cultivars Irrigated with Quicklime Treated Acid Mine Drainage Water. *Metabolites, 12*(3). https://doi.org/10.3390/metabo12030221

Nurhadi, B., Sukri, N., Sugandi, W. K., Widanti, A. P., Restiani, R., Noflianrini, Z., Rezaharsamto, B., & Herudiyanto, M. (2018). Comparison of crystallized coconut sugar produced by traditional method and amorphous coconut sugar formed by two drying methods: Vacuum drying and spray drying. *International Journal of Food Properties, 21*(1), 2339–2354. https://doi.org/10.1080/10942912.2018.1517781

Palmonari, A., Cavallini, D., Sniffen, C. J., Fernandes, L., Holder, P., Fagioli, L., Fusaro, I., Biagi, G., Formigoni, A., & Mammi, L. (2020). Short communication: Characterization of molasses chemical composition. *Journal of Dairy Science, 103*(7), 6244–6249. https://doi.org/10.3168/jds.2019-17644

Pérez, Y., Casado, M., Raldúa, D., Prats, E., Piña, B., Tauler, R., Alfonso, I., & Puig-Castellví, F. (2020). MCR-ALS analysis of 1H NMR spectra by segments to study the zebrafish exposure to acrylamide. *Analytical and Bioanalytical Chemistry, 412*(23), 5695–5706. https://doi.org/10.1007/s00216-020-02789-0

Puig-Castellví, F., Alfonso, I., & Tauler, R. (2017). Untargeted assignment and automatic integration of 1H NMR metabolomic datasets using a multivariate curve resolution approach. *Analytica Chimica Acta, 964*, 55–66. https://doi.org/10.1016/j.aca.2017.02.010

Queneau, Y., Jarosz, S., Lewandowski, B., & Fitremann, J. (2007). Sucrose Chemistry and Applications of Sucrochemicals. *Advances in Carbohydrate Chemistry and Biochemistry, 61*(07). https://doi.org/10.1016/S0065-2318(07)61005-1

Ravaglia, L. M., Pizzotti, A. B. C., & Alcantara, G. B. (2019). NMR-based and chemometric approaches applicable to adulteration studies for assessment of the botanical origin of edible oils. *Journal of Food Science and Technology, 56*(1), 507–511. https://doi.org/10.1007/s13197-018-3485-3

Ribeiro, I. C., Pinheiro, C., Ribeiro, C. M., Veloso, M. M., Simões-Costa, M. C., Evaristo, I., Paulo, O. S., & Ricardo, C. P. (2016). Genetic diversity and physiological performance of portuguese wild beet (beta vulgaris spp. maritima) from three contrasting habitats. *Frontiers. Plant Science, 7*(AUG2016), 1–14. https://doi.org/10.3389/fpls.2016.01293

Saviano, G., Paris, D., Melck, D., Fantasma, F., Motta, A., & Iorizzi, M. (2019). Metabolite variation in three edible Italian Allium cepa L. by NMR-based metabolomics: A comparative study in fresh and stored bulbs. *Metabolomics, 15*(8). https://doi.org/10.1007/s11306-019-1566-6

Savorani, F., Tomasi, G., & Engelsen, S. B. (2010). icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *Journal of Magnetic Resonance, 202*(2), 190–202. https://doi.org/10.1016/j.jmr.2009.11.012

Schmitt, C., Schneider, T., Rumask, L., Fischer, M., & Hackl, T. (2020). Food Profiling: Determination of the Geographical Origin of Walnuts by 1H NMR Spectroscopy Using the Polar Extract. *Journal of Agricultural and Food Chemistry, 68*(52), 15526–15534. https://doi.org/10.1021/acs.jafc.0c05827

Sekiyama, Y., Okazaki, K., Kikuchi, J., & Ikeda, S. (2017). NMR-based metabolic profiling of field-grown leaves from sugar beet plants harbouring different levels of resistance to Cercospora leaf spot disease. *Metabolites, 7*(1). https://doi.org/10.3390/metabo7010004

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J. (2006). Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics, 7*, 1–15. https://doi.org/10.1186/1471-2164-7-142

Vukov, K., & Hangypal, K. (1985). Sugar beet storage. *Sugar. Technology Review, 12*, 143–265.

Wang, Q., Wang, X., Wu, X., Wang, Y., Zhang, Y., Jiang, Y., Zhang, C., Huang, X., An, L., Ma, H., & Xu, K. (2021). 1H NMR-based metabolic profiling approach to identify the geo-authentic Chinese yam (Dioscorea polystachya Turczaninow cv. Tiegun). *Journal of Food Composition and Analysis, 98*(January). https://doi.org/10.1016/j.jfca.2021.103805

Wedeking, R., Maucourt, M., Deborde, C., Moing, A., Gibon, Y., Goldbach, H. E., & Wimmer, M. A. (2018). 1H-NMR metabolomic profiling reveals a distinct metabolic recovery response in shoots and roots of temporarily drought-stressed sugar beets. *PLoS One, 13*(5), 1–21. https://doi.org/10.1371/journal.pone.0196102

Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., … Forsythei, I. (2009). HMDB: A knowledgebase for the human metabolome. *Nucleic Acids Research, 37*(Suppl. 1), 603–610. https://doi.org/10.1093/nar/gkn810

Yang, S. O., Shin, Y. S., Hyun, S. H., Cho, S., Bang, K. H., Lee, D., Choi, S. P., & Choi, H. K. (2012). NMR-based metabolic profiling and differentiation of ginseng roots according to cultivation ages. *Journal of Pharmaceutical and Biomedical Analysis, 58*(1), 19–26. https://doi.org/10.1016/j.jpba.2011.09.016