J Sci Food Agric 1991, **56**, 419–434

Natural Isotope Fractionation in the Discrimination of Sugar Origins

Gérard J Martin,* Doubou Danho‡ and Claude Vallet

Laboratoire de RMN et Réactivité Chimique, UA CNRS 472—Université de Nantes 2, rue de la Houssinière, 44072 Nantes Cedex, France

(Received 1 November 1990; revised version received 2 April 1991; accepted 18 April 1991)

ABSTRACT

More than 500 carbohydrate samples have been characterised by hydrogen, carbon and oxygen isotopic parameters measured on ethanol and water resulting from controlled fermentation. Different chemical states of the carbohydrate pool have been considered: polysaccharides from cereals, tubers and leguminosae (maize, sorghum, rice, wheat, barley, potato, bean), glucose, fructose and sucrose from fruits (pineapple, citrus, apple, soft fruits), sucrose from sugar plants (beet, cane). The overall carbon-13 content of ethanol and the deuterium and oxygen-18 contents of water before fermentation were determined by mass spectrometry of isotope ratios whereas the investigation of site-specific natural isotope fractionation by deuterium NMR has provided access to the deuterium contents in the methyl (I) and methylene (II) sites of ethanol. The results have been analysed in the multidimensional space of these isotopic variables. Hierarchical clustering, principal component and discriminant analyses have been performed. Among fruits, for example, the pineapple group exhibits a 100% discrimination with respect to the apple and citrus groups which are themselves well distinguished. A still higher discrimination is reached between the three groups, citrus, cane and beet, and the addition of 10% exogenous cane or beet sucrose to citrus juice is unambiguously detected.

Key words: NMR, MS, stable isotope, sugars, juice.

* Author to whom correspondence should be addressed.
‡ Permanent address: Department of Chemistry, University of Abidjan, BP 852 Abidjan 22, Ivory Coast.

J Sci Food Agric 0022-5142/91/$03.50

INTRODUCTION

Isotope fractionation occurring in plants under the influence of biochemical or environmental factors is the source of relatively large ranges of isotopic composition in carbohydrates, lipids or other metabolites extracted from different plant species (O'Leary 1981). This phenomenon is of fundamental interest in biochemistry and in plant physiology. Thus the overall deuterium, carbon-13 and oxygen-18 contents determined by mass spectrometry have been shown to depend on various characteristics of the plant and on the climatic conditions which have prevailed during photosynthesis (Park and Epstein 1960; Whelan *et al* 1970; Rauschenbach *et al* 1979; Ziegler 1979; Sternberg *et al* 1984). In spite of the large number of factors that are likely to contribute, the total isotopic ratios of a given molecular species constitute good criteria for characterising the metabolic pathway and, in some cases, the botanical nature of the plant precursor. Thus sugars from plants with a C_3 (Calvin) metabolism are unambiguously distinguished from sugars with a C_4 (Hatch and Slack) origin on the basis of their $^{13}C/^{12}C$ isotope ratios (Smith and Epstein 1971). Similarly cellulose from plants with a CAM metabolism exhibits a higher total deuterium content than cellulose from C_3 or C_4 plants (Sternberg *et al* 1984). This type of behaviour can be exploited for origin authentication of food products, and isotope ratio mass spectrometry (IRMS) has been widely applied to the characterisation of fruit juices, alcoholic beverages, sugars, etc (Krueger and Reesman 1982; Schmidt 1986).

In this context we have shown that site-specific natural isotope fractionation can be studied by nuclear magnetic resonance (SNIF–NMR) and that the deuterium contents in the different sites of a molecule may deviate strongly from a statistical distribution (Martin and Martin 1981a,b). In particular, fermentation ethanol has been proved to constitute a reliable probe for characterising precursor sugars (Martin *et al* 1986). In contrast to mass spectrometry measurements, in which the selective information transmitted by the isotopomers of glucose and water is lost during the combustion of ethanol, site-specific ^{2}H-NMR determinations simultaneously provide several complementary isotopic criteria. Thus the hydrogen isotope ratio of the methyl site of ethanol, $(D/H)_I$, is strongly related to the isotopic behaviour of glucose whereas the isotope contents of the methylene site, $(D/H)_{II}$, and of the fermentation water, $(D/H)_w^Q$, are connected with that of must water via biochemical and exchange processes (Martin *et al* 1986). Having already demonstrated that the SNIF–NMR method greatly enhances the analytical potential of the isotopic methods (Martin *et al* 1982; Martin and Martin 1988) we have now developed a multi-site and multi-isotope approach for investigating the relative roles of the metabolic and physiological factors intervening in an isotopic discrimination of sugars from various fruits, cereals and sugar plants. By considering the alcoholic media resulting from standardised fermentation of the different sugars, overall and site-specific isotopic ratios have been obtained in reliable conditions for more than 500 samples. Since every sample is characterised by several parameters which are at least partly independent, the results are more conveniently represented in multidimensional space. The whole set of isotopic data has therefore been subjected to multivariate analyses in order

to check the analytical significance in terms of botanical criteria and to state precisely the discriminating power of the method.

MATERIALS AND METHODS

Sample treatments

Extraction of water and fermentable sugars

The different materials investigated (juices, grains, syrups) were processed in order to extract the chemical species (sugars and water) which were used as probes for subsequent isotopic analyses. Water was separated from the starting material by toluene extraction, and great care was taken to avoid any contamination by the organic solvent. In some typical cases, pure glucose was extracted either directly from the genuine juice or after hydrolysis of starch and elimination of proteins. Since isotope fractionation is likely to occur at any extraction step as a consequence of isotope effects, a careful mass balance was established for each extraction.

Fermentation of juice or water solutions of sugars into ethanol

Samples (400 or 500 ml) of juice were fermented in a 1 litre flask topped by a bubble dropper. Five grams of fresh reconstituted yeast belonging to the species *Saccharomyces cerevisiae* were directly added *in situ*. The fermentation was conducted (continuous stirring in a thermostated bath at 28°C) in order to achieve complete transformation of the sugars into ethanol. After 5 to 7 days, the fermentation is complete and the concentration of the remaining sugars is always lower than 0.5 g litre^{-1}.

The sugar concentrations were determined by HPLC on a Lichrosorb NH_2 (5-μm) column with an ethyl acetate (30%)/ethanol (55%)/water (15%) mixture as eluent. The results were calibrated against the data obtained in a conventional enzymic procedure, and the agreement was better than ± 1 g litre^{-1}. At the end of the fermentation the alcoholic grade of the fermented medium was determined by chromic oxidation, and azeotropic ethanol was recovered from a distillation with a Cadiot column fitted with a Teflon spinning band. The accurate percentage of water in ethanol was obtained by the Karl-Fischer method ($\pm 0.05\%$). Most often the yield of ethanol extraction was higher than 97%.

Concentrates were diluted with a sample of standard Nantes tap water (NTW) of a known isotopic content ($\delta D = -37.6‰$ (149.9 ppm) and $\delta^{18}O = -5.6‰$). The concentration of sugars and the isotope ratios of the medium were determined before and after the fermentation in order to compute the conversion rates which are always greater than 98%.

Isotopic determinations

The isotopic parameters are defined either on an absolute scale as the ratio, H/L, of the number of heavy isotopes, H, to the number of light isotopes, L, or on the relative scale $\delta‰$ in which the isotope ratio of the sample $(H/L)_i$, is referred to

that of the reference, $(H/L)_{ref}$

$$\delta‰ = \frac{(H/L)_i - (H/L)_{ref}}{(H/L)_{ref}} \quad (1)$$

Isotope ratio mass spectrometry (IRMS)

Deuterium. The isotope ratios of water extracted from juices or from the fermented mixtures were measured with a SIRA 9 VG Instruments mass spectrometer fitted with the on-line reduction assembly ISOPREP (uranium furnace). In some cases the fermentable sugars (glucose, fructose and sucrose) were converted into nitrate (Dunbar *et al* 1983) and combusted into CO_2 and H_2O for subsequent reduction with Zn.

The results are expressed in ‰ (1 ppm $\simeq$ 6‰) with respect to V.SMOW standard (Gonfiantini 1978). The overall precision of the different analytical steps going from the sample treatment to the IRMS determination is 1‰ (0.15 ppm) for water and 2‰ (0.35 ppm) for ethanols.

Carbon. The ^{13}C isotope ratios of ethanols and sugars were obtained with a Finnigan Delta E mass spectrometer equipped with a Carlo Erba micro-analyser. They are expressed on the PDB scale (Craig 1957). The precision of the whole analytical line, including ethanol distillation with a 98% yield, or sugar extraction from the juices, is equal to 0.2‰.

Oxygen. In some cases the $\delta^{18}O$ value of the water extracted from the juice or from the fermented mixture was determined with the Delta E instrument after equilibration with CO_2. The precision of the determination is equal to 0.1‰ and the results are expressed on the V.SMOW scale (Gonfiantini 1978).

NMR spectrometry (SNIF–NMR)

The site-specific isotope ratios of ethanols were determined at 61.4 MHz using a AM 400 Bruker spectrometer equipped with a $^2H\{^1H\}$ specific probe, an internal ^{19}F lock system and an automatic sample changer (Martin *et al* 1985). An appropriate computing system allows the FID recorded to be processed continuously and provides the values of the isotope ratios directly. The acquisition conditions were, respectively: sweep width, 1200 Hz; acquisition time, 6.8 s; broad-band decoupling, WALTZ (Shaka *et al* 1983); numbers of scans, 200. Eight different spectra were recorded for each sample. The overall precision is respectively equal to 1.5‰ (0.25 ppm) or 2‰ (0.35 ppm) for sites I (CH_2D) or site II (CHD) of ethanol.

Data analysis

Up to five different isotopic variables were observed for each fermentation mixture (565 samples). Such a large body of multivariate data needs to be reduced and clarified. The overall information represented by the isotopic parameters of the individuals contains some noise which should be cleared. The purpose of *principal*

component analysis is to reduce the original multidimensional space into a smaller number of uncorrelated factors. In the present work, up to 85% of the overall information is contained in the two main factors (principal components). A drawback of the method is that it transforms a complex hyper-space with a physical meaning into a simpler but abstract plane. A number of methods have been proposed for interpreting the principal components on a physical basis (Harman 1976; Malinowski and Howery 1980). One basic approach rotates the axes in order to find a new set of factors closely correlated with the original isotopic variables. Since we are dealing with a rather limited number of physical variables, the quartimax rotation (Harman 1976), which maximises the contributions of the whole set of variables to a restricted number of factors, is preferable. Examination of the reduced form of the original data then allows analogies or correspondences between the individuals or between an individual and the variables to be inferred.

The scores [**C**] (coordinates) of the individuals in the abstract plane (Fig 2) are readily computed from the post-multiplication of the autoscaled data (X_{ij}) by the loadings [**L**] (eigenvectors). Since the original data (x_{ij}) have different variation ranges, they are standardised according to:

$$X_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \tag{2}$$

where $\bar{x}_j$ and S_j are the mean and standard deviation of row j of the original data matrix [X].

The coordinates C_{i1} and C_{i2} of an individual i are computed from:

$$C_{i1} = [\mathbf{X}_i][\mathbf{L}_1] \tag{3a}$$

$$C_{i2} = [\mathbf{X}_i][\mathbf{L}_2] \tag{3b}$$

where $\mathbf{X}_i$ is the vector of the autoscaled observations ($(D/H)_I$, $(D/H)_{II}$, $\delta^{13}C$, $(D/H)^S_W$, for instance) of the individual i and [$\mathbf{L}_1$], [$\mathbf{L}_2$] are the eigenvectors corresponding to the first two components (Table 2).

Cluster analysis enables the similarity or dissimilarity between samples to be quantified. The similarity may be expressed in terms of the distance between two individuals in the representation space. This distance can be calculated in a number of ways (Neath and Sokal 1973). In this study, the euclidian distance is appropriate and the dissimilarity index of a node, d% (Fig 3), provides a quantitative estimation of the similarity between two groups or two samples. For example, the node between the group of the C_4-CAM and whey samples on one hand and of the C_3 samples on the other hand is characterised by a high degree of dissimilarity (35%), whereas the cane and maize botanical species exhibit a low dissimilarity (2%). The dissimilarity between the citrus and beet groups is relatively high (11%), and this behaviour may be the basis for discriminating pure citrus juices and juices sweetened by beet invert syrup.

Discriminant analysis is a classification technique which tries to maximise the distances between groups of individuals arranged beforehand. In contrast to principal component analysis, which does not require *a priori* arrangement of the individuals, discriminant analysis introduces a qualitative parameter, the

assignment of a given individual to a well defined group. From a mathematical point of view (Volle 1985), the number of discriminant components is equal to the number of groups minus one. For example, three groups of beet, citrus and cane samples are represented in a plane (Fig 4). In this representation, mixtures of the components lie between the groups.

In order to perform the appropriate statistical analyses we have developed a dedicated program which connects a spreadsheet with matrix operation facilities and a graphics module. Alternatively, commercial packages for multivariate analyses can be used (eg Einsight Pattern Recognition Software from Infometrix, Seattle, WA, USA; and Analys from Systèmes Informatiques de Gestion, Jouy en Josas).

RESULTS AND DISCUSSION

Sampling and isotopic characterisation

We have considered the different chemical states which exist in the carbohydrate pool of plants. In fruits, the carbohydrates are in the form of glucose, fructose and sucrose. With the exception of grapes which have already been considered (Martin *et al* 1988) the two main sources of C_3 fruit juices are apple (*Malus*) and citrus, including sweet orange (*Citrus sinensis* Osbeck) and grapefruit (*C paradisi* Macf). We have also exhaustively investigated the case of a CAM fruit, pineapple (*Ananassa sativa*), and have explored the isotopic behaviour of different soft fruits, strawberry (*Fragaria ananassa*), raspberry (*Rubus idaeus*), blackberry (*Rubus fructicosus*), blackcurrant (*Ribes nigrum*) and blueberry (*Vaccinium myrtillis*). One hundred and eighty-seven samples of fruits have been analysed.

Sugar plants such as sugar beet (*Beta vulgaris*) and sugar cane (*Saccharum officinarum*) contain mainly sucrose. Moreover the molasses of these plants are prolific sources of bio-ethanol which can subsequently be extracted from beverages or industrial materials. Altogether 197 samples have been investigated.

In cereals the carbohydrate pool is formed of starch polysaccharides. Five different botanical species belonging to the Gramineae have been considered: *Hordeum vulgare*, *Triticum sativum*, *Oryza sativa*, *Zea mays* and *Sorghum vulgare*. Two of them have a C_4 metabolism (maize, sorghum) whereas the other three are C_3 plants. One hundred and forty-one ethanol samples obtained from cereals have been considered.

Usually tubers and members of the Leguminosae also have a high starch content. We have investigated a limited number (20) of samples extracted from potato (*Solanum tuberosum*), manioc (*Manihot utilissima*), bean (*Vicia faba major*) and lentil (*Lens esculenta*).

For the sake of comparison, ethanols from animal (whey, milk) and fossil sources have been included in the study.

All the carbohydrate samples were fermented into ethanol and water in standardised conditions. We have shown previously, on the basis of model experiments (Martin *et al* 1986), that the deuterium content in the methyl site of ethanol is closely related to that of sites 1 and 6,6′ of glucose and also depends

on the isotope content of the aqueous medium. On the other hand, the proportion of methylenic isotopomers mainly depends on the isotope ratio of water, which transfers the hydrogens to site II under enzymic control, and on the concentration of fermentable sugars. The ethanolic mixture therefore constitutes a good, common probe for representing the isotopic behaviour of different molecular species of carbohydrates in their natural aqueous environment.

A given carbohydrate sample can be represented by a vector in a five-dimensional isotopic space described by three coordinates for ethanol, $(D/H)_I$, $(D/H)_{II}$ and $\delta^{13}C$, and two coordinates for water, $(D/H)^S_W$ and $\delta^{18}O^S_W$. $(D/H)_I$ and $(D/H)_{II}$ are the site-specific isotope ratios of the methyl and methylene sites of ethanol, $\delta^{13}C$ represents the overall carbon isotope ratio, and $(D/H)^S_W$ and $\delta^{18}O^S_W$ represent the hydrogen and oxygen isotope ratios of the water of the juice before fermentation. The value of the $(D/H)^S_W$ parameter may be inferred from that of the fermented medium, $(D/H)^Q_W$, since the deuterium enrichment accompanying fermentation, $\Delta(D/H)^{QS}_W = (D/H)^Q_W - (D/H)^S_W$, is correlated, either with the concentration of the starting sugar, C^S, or with the alcoholic grade of the product, t^Q:

$$\Delta(D/H)^{QS}_W = k_1 t^Q \qquad \text{or} \qquad k_2 C^S \tag{4}$$

The proportionality constant varies with the nature of the plant since it depends on the specific deuterium contents in the sites of the glucose molecule which are transferred to the aqueous medium (Martin *et al* 1982, 1986). The ranges of values characterizing the k_1 and k_2 coefficients are 0.35 to 0.5 ppm/‰ v/v and 0.02 to 0.03 ppm/g litre^{-1} respectively. According to model experiments and literature data(Dunbar *et al* 1983), the value of $\delta^{18}O^S_W$ is not significantly modified by the fermentation.

The data are usually well centred, and the mean over the samples can be considered as a good estimate of the mean over the whole population of the group. The mean site-specific deuterium parameters and $\delta^{13}C$ values computed for relatively large numbers of ethanol samples from typical origins can be compared in Table 1 with literature data obtained by isotope ratio mass spectrometry. The overall $\overline{\delta D}$ parameters represent a mean value over the methyl (I), methylene (II) and hydroxyl (III) sites, and are therefore influenced by possible exchange of the hydroxylic hydrogen with the water medium. The hydroxylic content can be estimated from $\overline{\delta D}$, δD_I and δD_{II} by considering that the overall isotope ratio is a weighted mean over the site-specific parameters. The calculated δD_{III} value ranges from -60‰ for the synthetic samples to about $+60$‰ for the beet ethanols. This behaviour suggests that the deuterium contents of the fermentation media exhibit relatively large differences. Significant perturbations may be introduced, in particular, by the technological treatments. Since such variations must also be reflected in the isotope ratio of the methylene site, the overall $\overline{\delta D}$ parameter measured by mass spectrometry is not simply representative of the sugar precursor, and the site-specific isotope ratios are expected to provide more meaningful information.

Relations between the isotopic variables

As expected from mechanistic evidence and from analyses of model experiments (Martin *et al* 1986), the five isotopic variables considered in this approach are

TABLE 1

Site-specific and overall hydrogen and carbon isotope parameters of typical ethanols from different origins

Origin	n	δD_I	δD_{II}	$\delta^{13}C$	δD^{κ}_{w}	n	$\delta\bar{D}$	$\delta^{13}C$	*Reference*
Sugar cane	89	−281 (9)	−180 (12)	−12.2 (0.8)	−36.0 (5)	22	−202.7 (18)	−12.6 (0.7)	Misselhorn *et al* 1983
						11	−201.2[a]	−11.6 (0.5)	Bricout and Menoret 1975; Simpkins and Rigby 1982
Maize	76	−290 (6)	−211 (10)	−10.2 (0.8)	−48 (9)	10	−208.3 (12)	−10.4 (0.5)	Misselhorn *et al* 1983
						12[b]	−206[a]	−11.9 (1.2)	Bricout and Menoret 1975; Koziet and Bricout 1978
							−220[a]	−10.3[a]	Rauschenbach *et al* 1979; Koziet and Bricout 1978
Sugar beet	100	−405 (5)	−203 (8)	−26.7 (0.6)	−41 (6)	7	−262.9 (6)	−27.2 (0.8)	Misselhorn *et al* 1983
							257.2[a]	−28.0[a]	Bricout and Menoret 1975
Tubers	20	−376 (6)	−184 (8)	−26.9 (0.5)	−4.0 (4)	24[c]	−237.2 (14)	−26.3 (0.8)	Misselhorn *et al* 1983
						23	−243.6 (12)		Rauschenbach *et al* 1979
Synthesis	8[d]	−186 (25)	−75 (22)	−31.0 (3)	—	13	−128.8 (8)	−28.0 (2.5)	Misselhorn *et al* 1983
						10	−126.9 (7)	−27.4 (2)	Rauschenbach *et al* 1979
							−135.8[a]	−31.0[a]	Bricout *et al* 1975
Fruits	45[e]	−376 (4)	−196 (5)	−26.6 (0.6)	−35 (6)				

n is the number of investigated samples. The numbers in parentheses are the standard deviations. δD_I, δD_{II} are the site-specific parameters (‰) of the methyl and methylene sites respectively and $\overline{\delta D}$ is the mean value of the whole ethanol molecule (‰) measured by mass spectrometry (eqn **1**).

[a]The standard deviation is not specified.

[b]Number of carbon-13 determinations.

[c]These results are for potato samples.

[d]From ethylene.

[e]Mainly apple and soft fruits.

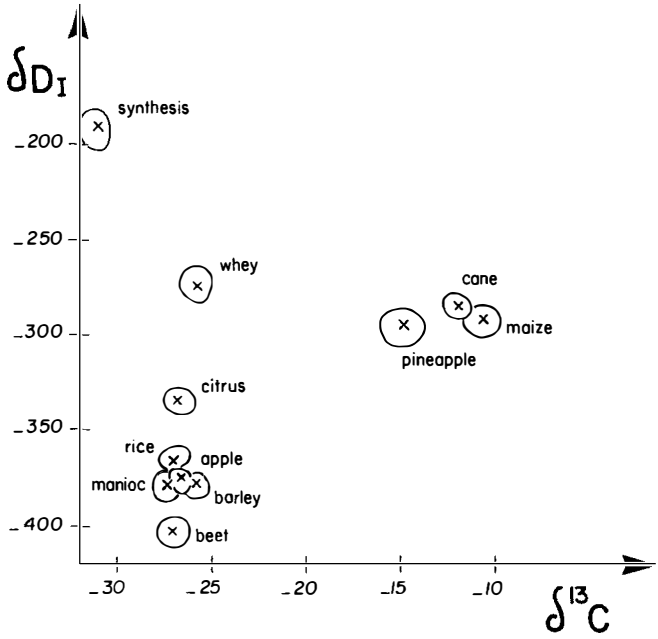

Fig. 1. Representation of groups of ethanol samples from different origins in the plane δD_I, $\delta^{13}C$. $\delta^{13}C$ is the overall carbon-13 content whereas δD_I represents the site-specific deuterium content associated with the methyl isotopomer of ethanol CH_2DCH_2OH. The data are expressed on the δ scale defined in the experimental section (eqn **1**).

only partly independent. Thus, if synthetic and whey samples are disregarded, a two-dimensional representation of the specific methyl deuterium content against the overall carbon-13 parameter (Fig 1) illustrates the existence of a rough linear relationship within the whole population of the plant species. However, this property is mainly imposed by the metabolic distinction between C_3 molecules, relatively depleted in ^{13}C, and richer C_4 and CAM species. Simple linear relationships no longer hold satisfactorily within populations restricted to a given photosynthetic metabolism.

Better correlations are observed between the deuterium parameters of water and of site II of ethanol. Thus for the whole set of natural samples the following relationship is satisfied: ($n = 10$ groups, $r = 0.7$)

$$\delta D_{II} = 0.48 \delta D_W^S + 180 \qquad (5)$$

In the same respect hydrogeological investigations have shown that good correlations exist between the hydrogen and oxygen isotope ratios of natural waters (Craig 1961). The slope of the line relating the δD and $\delta^{18}O$ values of meteoric waters from world-wide stations is close to 8. However, smaller coefficients have been determined for regional precipitations from semi-arid and arid regions (Yurtsever and Gat 1981) and non-equilibrium evapotranspiration effects contribute to a strong lowering of the slope for leaf water or orange juice water, for example (Lesaint *et al* 1974; Allison *et al* 1985; Bricout and Koziet 1985). If

we consider the whole set of data obtained for juices extracted from different botanical species such as citrus, pineapple, cane sugar, apple or whey, a satisfactory correlation is found between the deuterium and oxygen parameters ($n = 5$ groups, $r = 0.8$)

$$\delta D_W^S = 5.6\delta^{18}O_W^S - 21 \quad (6)$$

However, noticeably different behaviour of the three parameters δD_{II}, δD_W^S and $\delta^{18}O_W^S$ is exhibited within well defined, more reduced categories of plant products. The introduction of all hydrogen, carbon and oxygen isotopic vectors into subsequent data analyses therefore remains statistically meaningful. Unfortunately, due to unavailability of appropriate IRMS facilities at the beginning of this study, the oxygen parameters could not be obtained for all samples. Consequently most multivariate analyses of the results have been performed in a space with only four dimensions.

Principal component analysis of the isotopic data

Multivariate analysis is a very efficient tool for appraising the discriminating potential of a set of variables associated with a large number of individuals. A principal component analysis has first been performed on the whole set of data. Figure 2A represents the different groups of ethanols and the four isotopic parameters, $\delta^{13}C$, δD_I, δD_{II}, δD_W^S in the plane defined by the two main factors which represent 87% of the total variance (Table 2). In Fig 2B the same individuals and variables are described after a quartimax rotation of the three initial factors which allowed for 95% of the variance (Harman 1976) (Table 2). Quartimax rotation produces a factor matrix for which the variance of the squared factor loadings is maximum. This procedure therefore maximises the contribution of each variable to the first two components which still explain 87% of the total variance. The isotopic parameters δD_I and $\delta^{13}C$ are clearly identified to one quartimax axis whereas δD_{II} and δD_W^S correspond to the other axis. This factorial analysis shows that the deuterium content in the methyl group of ethanol and the overall carbon-13 content, which are mainly governed by the filiation from glucose, exhibit a high discrimination power. The type of photosynthetic metabolism plays an important role in this discrimination. Thus C_4 sugars extracted from cane or maize and CAM sugars from pineapple, which are all characterised by high deuterium and carbon-13 contents, lead to ethanol samples situated on the upper side of the diagram. In contrast, the most depleted C_3 sugars (eg beet, manioc) are fermented into ethanol samples which appear on the lower side. The horizontal axis corresponds to isotopic variables mainly connected to the isotopic behaviour of the aqueous fermentation medium (Martin *et al* 1986). This factorial axis enables further distinction to be achieved between plants characterised by the same metabolism and undergoing comparable fractionation phenomena during the photosynthesis of glucose. Citrus and apple, and pineapple and cane, for example, are well separated by this factor. However, it should be emphasised that the discrimination along the horizontal axis accounts for properties of a fermentation medium which may be composed of natural juices, of added waters or of mixtures of both.

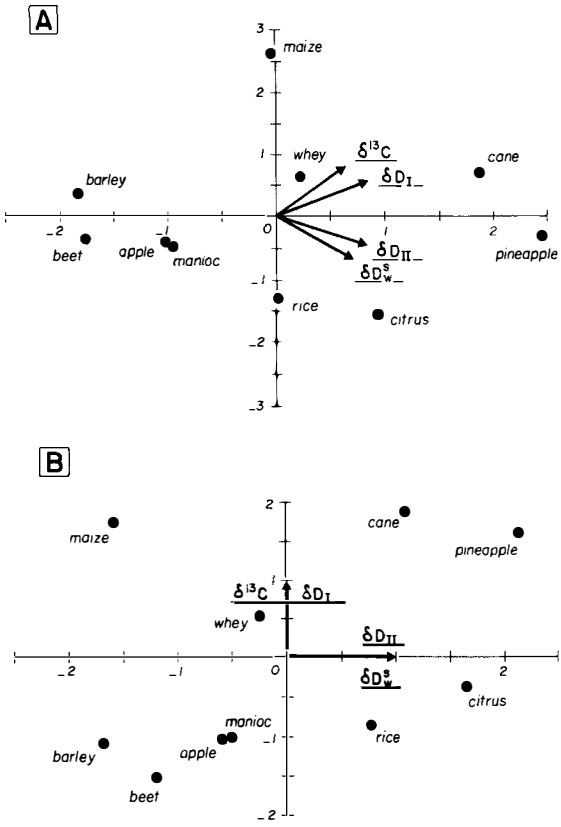

Fig. 2. Principal component analysis of the results in a four-dimensional isotopic space δD_I, δD_{II}, $\delta^{13}C$, δD^s_w. In diagram A the individuals are represented in the plane of the two main original axes which express 87.4% of the total variance. Diagram B presents the samples in the plane of the two main axes resulting from a quartimax rotation of the three initial factors (Harman 1976).

TABLE 2
Principal component analysis

Loadings		*Variables*	δD_I	δD_{II}	$\delta^{13}C$	δD^s_w
A	Original Axes	1st comp (52.4%)	0.530	0.555	0.422	0.482
		2nd comp (35.0%)	0.444	−0.423	0.591	−0.524
B	Quartimax Rotation	1st comp (44.6%)	0.150	0.687	0.005	0.711
		2nd comp (42.5%)	0.674	0.111	0.730	0.019

The first two components represent more than 87% of the variance. A corresponds to the original axes (Fig 2A) and **B** corresponds to a quartimax rotation of the original axes (Fig 2**B**). The loadings of the original or quartimax axes can be used to compute the coordinates of an unknown sample for which the isotope ratios δD_I, δD_{II}, $\delta^{13}C$ and δD^s_w have been determined (Table 1) and autoscaled according to eqn (**2**). The means (standard deviations) of the four isotopic variables for the whole population considered (10 origins) are respectively (‰): $\delta D_I = -337.1\ (48.8)$; $\delta D_{II} = -187.1\ (19.5)$; $\delta^{13}C = -22.2\ (6.9)$; $\delta D^s_w = -25.4\ (18.1)$.

Group representation and group classification

On the basis of the four isotopic variables δD_I, δD_{II}, $\delta^{13}C$ and δD^S_W, we have carried out a hierarchical clustering analysis searching for the degree of similarity of the ethanols in terms of the botanical origin of their precursor. Eleven groups representing the plant species have been considered. In the case of cereals two groups of C_4 plants, maize and sorghum, and three groups of C_3 plants, barley, rice and wheat, have been retained. Due to the relatively small number of samples investigated, only one group has been made up from tubers (manioc, potato) and Leguminosae (lentil, bean). In the case of fruits three groups are distinguished: pineapple, citrus and apple. The two sugar plants, beet and sugar cane, constitute the two last plant groups. For the sake of comparison we have also considered two very different kinds of alcohols: whey ethanol, which comes from animals, and synthetic ethanol which derives from fossil products. The rectangular table containing the mean isotopic values of each group was converted into a proximity table on the basis of the Euclidian distances between the groups. The similarity tree is represented in Fig 3, and the distribution of the different categories of samples is also given. A relative degree of dissimilarity, d%, is defined by referring all the Euclidian distances to the largest distance which separates the synthetic ethanols from the whole set of 'natural' products. If the tree is cut just after this first branching, two main groups are well recognised. One of them contains all C_3 plants and the other collects together both C_4 and CAM plant species and whey. The next level of discrimination distinguishes whey from C_4-CAM plants. Citrus is then separated from the other C_3 plants, and a branching occurs between the CAM fruit, pineapple, and the two C_4 plants, sugar cane and maize. Among the C_3 plants, the similarity tree then separates at successive steps, a sugar plant (beet), a cereal (barley) and a fruit (apple). The manioc–rice branching occurs at the top of the tree, and the euclidian distances are still smaller between wheat and barley and between sorghum and maize. It should be recalled, however, that the different groups are built from very unequal numbers of individuals. Certain groups such as wheat and sorghum, which contain a restricted number of samples, are defined with a lower degree of confidence and this may be the source of some discrepancies in the ordering of the very last branchings.

It is obvious from this similarity tree that the discrimination between sugars does not occur primarily in terms of the main plant categories such as fruits and cereals. Thus, even among C_3 plants, an apple fruit, for example, may be closer to a cereal such as rice than the rice is itself to another cereal such as barley. From a related point of view, the monomeric or polymeric chemical state of the carbohydrate pool is not a primary criterion of distinction. Within the category of plant sugars the level of similarity probably integrates complex influences of physiological properties which exhibit variable degrees of sensitivity to environmental factors. The botanical species of the plant therefore seems to constitute a suitable criterion for classifying the individuals on the basis of the isotopic variables. However, since the hierarchical analysis is only grounded on proximity parameters between groups and does not take into account the expansion of these groups, the discriminating power has been rationalised on a more quantitative basis in further multivariate treatments of the data.

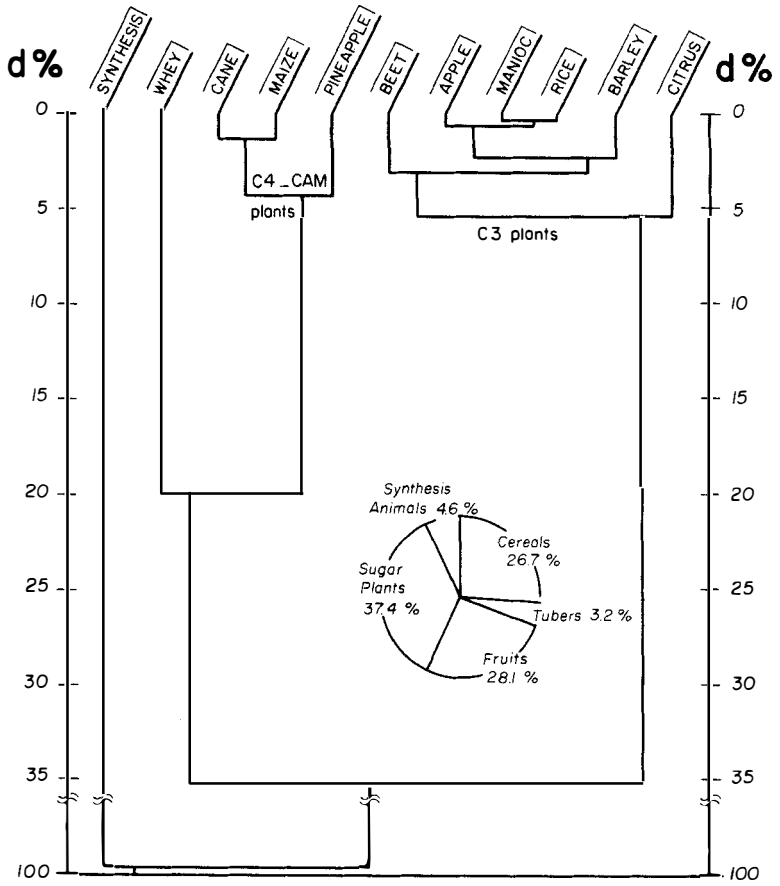

Fig. 3. Sample distribution and hierarchical clustering of groups of ethanols built according to the botanical origin of the sugar precursor. The classification is made on the basis of a euclidian similarity between groups characterised by the four isotopic parameters δD_I, δD_{II}, $\delta^{13}C$, δD^s_w. The rectangular table containing the isotope ratios of the groups of ethanols from different origins is transformed into a triangular proximity table using the mean distance between groups (Euclidian distance) as a similarity criterion. The synthetic ethanols, for which no δD^s_w values intervene, were arranged according to only three isotope ratios but this procedure does not perturb the arrangement order of the other groups.

Discriminating ability of the method

The multi-site approach founded on site-specific natural isotope fractionation (SNIF–NMR) greatly enhances the analytical potential of the isotopic methods. The principal component analysis suggests that, in a number of cases, ethanols from a given plant can be assigned to a well defined group. The distinction between synthetic and plant origins, frequently possible on the sole basis of the overall isotope contents, is further substantiated and new possibilities are offered for the inference of the botanical origin. The assignment power of the method can be checked by factor discriminant (FDA) analyses. We have considered in particular the case of the three groups of fruits, pineapple, citrus and apple. The FDA matrix was first built from the values of the four isotopic variables δD_I, δD_{II}, $\delta^{13}C$ and

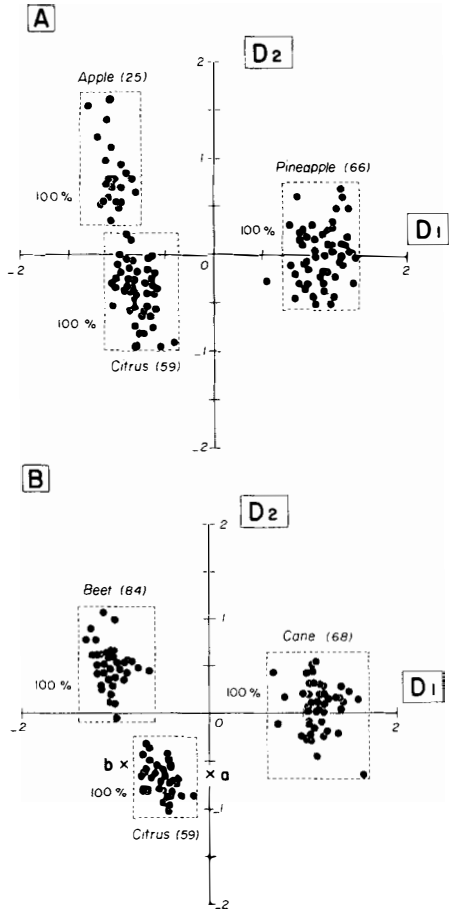

Fig. 4. Examples of discriminant analyses carried out on individual samples representing different kinds of fruits and sugar plants. The rectangles are the two-dimensional projections of the parallelepipeds corresponding to the natural dispersion of the samples. The numbers of samples are given within parentheses. The discriminant analysis illustrated in A involves 150 samples of fruits for which four isotopic variables have been measured. Diagram B shows that all samples of citrus, beet and cane are perfectly assigned and that even a 10% addition of cane (point a) or beet (point b) is detected.

δD_W^S associated with 150 fruit samples: 66 from pineapple, 59 from citrus and 25 from apple. Figure 4A shows that the classification is excellent since all the samples are perfectly assigned. When the $\delta^{18}O$ value of the juice is considered as a fifth isotopic variable, the discriminant power of the second axis is slightly improved. Species that correspond to an earlier branching in the similarity tree are even better characterised. Thus a 100% discrimination is reached between sugars photosynthesised in citrus, sugar beet and cane sugar as demonstrated by an FDA analysis involving 211 samples (Fig 4B). In contrast, at the highest degrees of similarity, group overlapping may be rather severe. As successfully achieved in the case of grape (Martin *et al* 1988), complementary approaches taking into account the influence of environmental variables need to be developed to improve the analytical assignment of C_3 cereals such as manioc, wheat and rice, for example.

The high discriminating ability accessible for species such as citrus, beet and cane gives the method a unique efficiency for detecting and quantifying adulterations by exogenous sugars. For example, much interest has been devoted to the problem of juice adulteration, and mass spectrometry determinations of overall isotope contents have proved very useful (Nissembaum *et al* 1974; Bricout and Koziet 1987; Doner *et al* 1987). These methods are especially helpful for detecting the addition of C_4 sugars (cane) to C_3 juices. However, they are less efficient for quantitating eventual additions of exogenous C_3 sugar. In this respect, Fig 4B shows that two samples of a citrus juice containing 10% of cane sucrose (a) or 10% of beet sucrose (b) are both located well outside the citrus domain. The site-specific dimension introduced into the present multi-site multi-element approach by the SNIF–NMR method greatly improves the analytical distinction within C_3 species and is therefore expected to be of prime interest for the authentication of natural juices or natural ethanols.

ACKNOWLEDGEMENTS

We are very grateful to the following persons and institutions for having provided us with typical samples of fruits, cereals, sugars and ethanols: Citrus Institute Lake Alfred, USA (J Attaway); UNGDA, Paris (S Tourlière); INRA, Dijon and Guadeloupe (P Dupuy); SNEA, Braselfa, Brazil (J Bouvier); Tipiak, Nantes (A Bouleau); Bric-Fruit, Nantes (M Mineau); CANA, Ancenis; Arômes de Bretagne, Antrain; Katell-Roc, Serent; Verniers, Cossé le Vivien; SIAO, Beaumont.

REFERENCES

Allison G B, Gat G R, Leaney F W J 1985 The relationship between deuterium and oxygen-18 delta values in leaf water. *Chem Geol* **58** 145–156.

Bricout J, Koziet J 1985 Détection de l'addition de sucre dans les jus d'orange par analyse isotopique. *Science des Aliments* **5** 197–204.

Bricout J, Koziet J 1987 Control of the authenticity of orange juice by isotopic analysis. *J Agric Food Chem* **35** 758–760.

Bricout J, Menoret Y 1975 Teneur en isotopes stables du rhum et des principaux alcools de bouche. *Ann Technol Agric* **24** 247–254.

Bricout J, Fontes J C, Merlivat L, Pusset M 1975 Sur la composition en isotopes stables de l'éthanol. *Ind Alim Agric* **92** 375–378.

Craig H 1957 Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide. *Geochim Cosmochim Acta* **12** 133–149.

Craig H 1961 Isotopic variations in meteoric waters. *Science* **133** 1702–1704.

Doner L W, Ajie H O, Sternberg L S L, Milburn J M, De Niro M J, Hicks K B 1987 Detecting sugar beet syrups in orange juice by D/H and $^{18}O/^{16}O$ analysis of sucrose. *J Agric Food Chem* **35** 610–612.

Dunbar J, Schmidt H L, Woller R 1983 Möglichkeiten des Nachweises der Zuckerung von Wein über die Bestimmung von Wasserstoff-Isotopenverhältnissen. *Vitis* **22** 375–386.

Gonfiantini R 1978 Standards for stable isotope measurements in natural compounds. *Nature* **271** 534–536.

Harman H H 1976 *Modern Factor Analysis*. The University of Chicago Press, Chicago, IL.

Koziet J, Bricout J 1978 Determination du pourcentage de whisky de grain dans les whiskies commerciaux par spectrométrie de masse isotopique. *Ann Nutr Alim* **32** 941–946.

Krueger H W, Reesman R H 1982 Carbon isotope analyses in food chemistry. *Mass Spectrometry Rev* **1** 205–236.

Lesaint C, Merlivat L, Bricout J, Fontes J C, Gautheret R 1974 Sur la composition en isotopes stables de l'eau de la tomate et du maïs. *C R Acad Sci D* **278** 2925–2930.

Malinowski E R, Howery D G 1980 *Factor Analysis in Chemistry*. Wiley, New York.

Martin G J, Martin M L 1981a Marquage isotopique en abondance naturelle. Application de la résonance magnétique nucléaire haute résolution du deutérium à l'étude de composés vinyliques. *C R Acad Sci II* **293** 31–33.

Martin G J, Martin M L 1981b Deuterium labelling at the natural abundance level as studied by high field quantitative ^{2}H NMR. *Tetrahedron Lett* **22** 3525–3528.

Martin G J, Martin M L 1988 The site-specific natural isotope fractionation–NMR method applied to the study of wines. In: *Modern Methods of Plant Analysis, Vol 6*, eds Linkens H F, Jackson J F. Springer-Verlag, Berlin, pp 258–275.

Martin G J, Martin M L, Mabon F, Michon M J 1982 Identification of the origin of natural alcohols by natural abundance hydrogen-2 nuclear magnetic resonance. *Anal Chem* **54** 2380–2382.

Martin G J, Sun X Y, Guillou C, Martin M L 1985 NMR determination of absolute site-specific natural isotope ratios of hydrogen in organic molecules. Analytical and mechanistic applications. *Tetrahedron* **41** 3285–3296.

Martin G J, Zhang B L, Naulet N, Martin M L 1986 Deuterium transfer in the bioconversion of glucose to ethanol studied by specific isotope labelling at the natural abundance level. *J Amer Chem Soc* **108** 5116–5122.

Martin G J, Guillou C, Martin M L, Cabanis M T, Tep Y, Aerny J 1988 Natural factors of isotope fractionation and the characterization of wines. *J Agric Food Chem* **38** 316–322.

Misselhorn K, Brückner H, Müssig-Zufika M, Grafarend W 1983 Nachweis des Rohstoffs bei hochrektifiziertem Alkohol. *Branntweinwirtschaft* 162–170.

Neath P S, Sokal R 1973 *Numerical Taxonomy*. W H Freeman, San Francisco.

Nissembaum A, Lifshitz A, Stepek Y 1974 Detection of citrus adulteration using the distribution of natural stable isotopes. *Lebensm Wiss Technol* **7** 152–152.

O'Leary M H 1981 Carbon isotope fractionation in plants. *Phytochemistry* **20** 553–567.

Park R, Epstein S 1960 Carbon isotope fractionation during photosynthesis. *Geochim Cosmochim Acta* **21** 110–126.

Rauschenbach P, Simon H, Stichler W, Moser H Z 1979 Vergleich der Deuterium- und Kohlenstoff-13-gehalte in Fermentation- und Synthese Ethanol. *Z Naturforsch* **34c** 1–4.

Schmidt H L 1986 Food quality control and studies on human nutrition by mass spectrometric and nuclear magnetic resonance isotope ratio determination. *Fresenius Z Anal Chem* **324** 760–766.

Shaka A J, Keeler J, Frenkiel T, Freeman R 1983 An improved sequence for broadband decoupling: WALTZ-16. *J Magn Res* **32** 335–338.

Simpkins W A, Rigby D 1982 Detection of the illicit extension of potable spirituous liquors using $^{13}C/^{12}C$ ratios. *J Sci Food Agric* **33** 898–903.

Smith B N, Epstein S 1971 Two categories of $^{13}C/^{12}C$ ratios for higher plants. *Plant Physiol* **47** 380–384.

Sternberg L, Deniro M J, Ting L P 1984 Carbon hydrogen and oxygen isotope ratios of cellulose from plants having intermediary photosynthetic modes. *Plant Physiol* **74** 104–107.

Volle M 1985 *Analyse des Données*. Economica, Paris.

Whelan T, Sackett W M, Benedict C R 1970 Carbon isotope discrimination in a plant possessing the C_4 dicarboxylic acid pathway. *Biochem Biophys Res Comm* **41** 1205–1210.

Yurtsever Y, Gat J R 1981 Atmospheric waters in stable isotope hydrology. *Technical Reports Series No 210*. IAEA, Vienna, pp 103–142.

Ziegler H 1970 Diskriminierung von Kohlenstoff- und Wasserstoffisotopen: Zusammenhänge mit dem Photosynthesemechanismus und den Standortbedingungen. *Ber Deutsch Bot Ges* **92** 169–184.