Food Research International 43 (2010) 965–971

Contents lists available at ScienceDirect

Food Research International

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

Amino acid profiles and quantitative structure–property relationships for malts and beers

Alicia B. Pomilio [a,*,1], Pablo R. Duchowicz [b,1], Miguel A. Giraudo [c,1], Eduardo A. Castro [b,1]

[a] PRALIB (CONICET, UBA), División Química Estructural y Espectroscopía, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956, C1113AAD Buenos Aires, Argentina
[b] Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas INIFTA, Universidad Nacional de La Plata, CCT La Plata-CONICET, Diag. 113 y 64, C.C. 16, Suc. 4, B1900 La Plata, Argentina
[c] Carrera de Ciencia y Tecnología de los Alimentos, Universidad Nacional de Lanús, 29 de Setiembre 3901, B1826 Lanús, Provincia de Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 11 September 2009
Accepted 19 January 2010

Keywords:
Quantitative structure–property relationships
Multi-variable Linear Regression
Replacement method
Aminograms
Malts
Blonde and black beers

ABSTRACT

Experimental amino acid concentrations of blonde and black commercial beers, brewed in Argentina, as well as national malts were subjected for the first time to Quantitative Structure–Property Relationships (QSPRs). Thus, Dragon theoretical descriptors were derived for a set of optimised amino acid structures with the purpose of assessing QSPR models. We used the statistical Replacement Method for designing the best multi-parametric linear regression models, which included structural features selected from a pool containing 1497 constitutional-, topological-, geometrical-, and electronic-type molecular descriptors. In this work QSPR results were in good agreement with experimental amino acid profiles, thus demonstrating the predictive power of the designed QSPRs. QSPR-modelling was used to predict aminograms, and was also used to estimate non-available amino acid concentrations for these malts, and beers. The developed QSPR approach showed to be an useful tool for discriminating among blonde and dark beers, and malts. This is a new application of the QSPR theory to food, in particular to chemical biomarkers of malts and beers.

1. Introduction

Argentina is one of the countries in the American continent with an old beer tradition, although is also a wine country. The origin of the beer in Argentina can be traced back to the seventeenth century, when the Jesuits were settled in this land. However, some historians state 1744 as start of beer production outside religious centres and monasteries. Baron Emile Bieckert is considered the first brewer in Argentina, founding the first brewery in Argentina under his name. In 1852 Otto Bemberg, arrived to Argentina, and in 1887 built his brewery in Quilmes, 25 km away from Buenos Aires. Bemberg was very important in the development of barley and hops crops as well as crown taps, and the widespread of breweries in Argentina, Paraguay and Brazil.

The manufacture of home-made beer began in the last 12 or 15 years in Argentina. Nowadays there are hundreds of artisanal breweries all over the country.

World producers of beer and other alcoholic beverages tend to be large multinational corporations (Jernigan, 2009). Argentina is included in this world trend, although there are also so far smaller producers ranging from brewpubs to regional breweries.

As it is known, brewing involves two major biological systems, e.g., malted barley (malt) and yeast. Usually, the brewing industry uses well-established tests to assess malt quality, but these often fail to predict malt-associated problem fermentations. Once the malt is converted into the wort, yeast converts carbohydrates of the wort into alcohol and other products that influence appearance, flavour, and taste. The fermentable carbohydrate composition of wort and the manner in which it is used by yeast during brewery fermentation have a direct influence on fermentation efficiency and quality of the final product.

The use of starter cultures became a common practice after methods for the isolation of pure yeast strains were developed. Moreover, efforts have been undertaken to improve *Saccharomyces* yeast strains, first by classical genetic methods and later by genetic engineering (Donalies, Nguyen, Stahl, & Nevoigt, 2008).

Despite the regional variations, beer is categorised into two main types based on the temperature of the brewing which influences the behaviour of yeast used during the brewing process, e.g., *lagers*, and the more regionally distinct *ales*, brewed at a higher temperature. Different roasting times and temperatures are used to produce a variety of colours of malt from the same grain.

* Corresponding author. Tel./fax: +54 011 4814 3952.
E-mail addresses: pomilio@ffyb.uba.ar, abpomilio@sinectis.com.ar (A.B. Pomilio), duchow@inifta.unlp.edu.ar, pabloducho@gmail.com (P.R. Duchowicz), mgiraudo@unla.edu.ar (M.A. Giraudo), castro@quimica.unlp.edu.ar, eacast@gmail.com (E.A. Castro).

[1] All authors contributed equally to this work. This work was part of the Ph.D. Thesis of Miguel A. Giraudo, approved with awards on May 21, 2009.

0963-9969/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodres.2010.01.006

Most of the yeasts collected from actual lager fermentations have hybrid genomes consisting of *Saccharomyces cerevisiae*, *Saccharomyces bayanus*, and the "lager-type" sequences, and seem to lack *Saccharomyces bayanus* var. *uvarum* sequences (Rainieri et al., 2006).

Determination of beer authenticity is a crucial issue in beer quality control and safety. For example, infrared (MIR) and near infrared (NIR) spectroscopy (Iñón, Garrigues, & de la Guardia, 2006), together with attenuated total reflectance–Fourier transform infrared spectroscopy (ATR–FTIR) (Llario, Iñón, Garrigues, & de la Guardia, 2006; Iñón et al., 2006), has been used for assessing beer quality parameters. It is also necessary to use chemical biomarkers to characterise malts and beers as well as other foods, e.g., flavonoids (Gorinstein et al., 2007), and amino acids (Molnár-Perl, 2005). The fingerprint of food allows its univocal identification and characterisation.

The profile of free amino acids and ammonia (aminogram) of beers and malts has been previously reported from our laboratories (Giraudo et al., 2004). Amino acids were analysed by high performance liquid chromatography (HPLC) with fluorescence and UV detection (Elfakir, 2005; Pomilio, 1994, 2004).

The experimental ranges of amino acid concentrations have been now further subjected to Quantitative Structure–Property Relationships (QSPRs) (Hansch & Leo, 1995) in order to distinguish these malts and beers. Therefore, QSPR models have been proposed, which are able to estimate the properties of the amino acids, taking into account that the resulting effects are a consequence of the molecular structures.

In past years, the Multi-variable Linear Regression (MLR) technique has proved to be of valuable applicability in establishing predictive QSPR by performing an exhaustive analysis of a pool containing a great number of structural molecular descriptors. Linear models are more general, and can transparently reveal the effects upon the property being modelled of the inclusion/exclusion of structural variables present in the model. The main advantage of developing linear regression models, when compared to non-linear ones, is the fact that they do not involve too many optimisation parameters to be found for model building, just the regression coefficient for each model's numerical variable. In addition, some recent studies have applied the Gram-Schmidt procedure in the recent proposed Spectral–Structure Activity Relationship (S-SAR) method (Putz & Lacrămă, 2007). The advantage of this technique is that it enables to replace the classic Multi-variable Linear Regression analysis by purely algebraic models with some conceptual and computational advantages, having both ecotoxicological, environmental, and anticancer bioactivity applications (Chicu & Putz, 2009; Lacrămă, Putz, & Ostafe, 2007; Putz & Lacrămă, 2007; Putz, Putz, Lazea, Ienciu, & Chiriac, 2009). Then, the chemical structures of the amino acids were translated into molecular descriptors, thus describing different relevant features of these compounds by a mathematical formula for each food matrix, which has been proposed from the Chemical Graph Theory, Information Theory, Quantum Mechanics, etc. (Katritzky, Lobanov, & Karelson, 1995; Todeschini & Consonni, 2000).

Suitable descriptors were selected from those available to characterise the properties of the compounds under consideration, e.g., amino acids. In the present analysis, we explored a pool containing 1497 theory-based descriptors computed by means of the Dragon software, and established QSPRs for each beer and malt samples under analysis by means of the Replacement Method technique (Duchowicz, Castro, Fernández, & González, 2005; Duchowicz, Mercader, Fernández, & Castro, 2008; Duchowicz, Talevi, Bruno-Blanch, & Castro, 2008).

As far as we know this is the first report on the development of QSPR models based on the amino acids of some types of beers and malts.

2. Materials and methods

2.1. Samples

Commercial samples of national and international blonde beers (trademarks: Andes, Bieckert, Brahma, Budweiser, Córdoba, Heineken, Iguana, Isenbeck, Liberty, Palermo, Quilmes Cristal, Quilmes light, Río Segundo Rubia, Santa Fe, Schneider, Warsteiner) and two dark beers (trademarks: Brahma Morena, Quilmes Bock), brewed in Argentina, as well as two national malts (trademarks: Malt of Salta Brewery; MaltiFe of Grupo Estrella SA), all marketed in Argentina and in the region, were used. Sample vials were kept closed until analysis. Samples were degassed conveniently, filtered through Millex LCR 13 filters, and directly derivatized with 6-AQC, as indicated in the commercial kit. Norleucine (Sigma) was used as internal standard.

2.2. Instrumentation and chromatographic conditions

An Agilent liquid chromatograph Model 1050 was used, composed of a ternary pump, on-line vacuum degasser, automatic injector, column oven, microbore connectors; UV–Vis variable Model C detector with 1 μl-cell; Model 1046 fluorescence detector and HP 3396 II Integrator. A C_{18} Hypersil column (200 mm length and 2.1 mm i.d.; 5 μm of particle diameter; pore diameter of 12 nm), and a guard column of 3.9 × 20 mm were used. A ternary system was used as mobile phase: *Solvent A:* 140 mM sodium acetate, 17 mM triethanolamine, pH 5.05, containing 1 mM disodium edetate. *Solvent B:* Acetonitrile, HPLC grade. *Solvent C:* Water, HPLC grade, as previously reported (Giraudo et al., 2004). *Flow rate:* 0.33 ml/min. *Detection:* UV: 248 nm. Fluorescence: excitation wavelength = 250 nm and emission wavelength of 395 nm. *Oven temperature:* 40 °C. *Injection volume:* 5 μl.

2.3. Analysis of free amino acids

Free amino acids present in the malts and beers under study were analysed by the AccQ·Tag technique of Waters Corporation, as previously reported (Giraudo et al., 2004). We used the Waters commercial kit provided by D'Amico Sistemas SA (Buenos Aires, Argentina) for precolumn derivatization.

AccQ·Tag kit contained an ampoule with 17 standards of amino acids each with a concentration of 2.5 mM, except for cystine, which had a concentration of 1.25 mM. Once the profile of amino acids was obtained, the concentrations of the main amino acids were calculated by comparison with the AccQ·Tag kit of amino acids' standards and the internal standard concentration.

Water was obtained by filtration through a Milli-Q system and samples were filtered through Millex-SLCR13 Millipore filters (Biopore, Buenos Aires, Argentina). All solvents were filtered daily through GVWP04700 Millipore filters (Buenos Aires, Argentina).

Experimental concentration ranges were expressed in ppm, and accounted for the average of three replications of each sample.

Once the chromatograph was optimised with the standards of amino acids (variation coefficient less than 5%), alternately three samples and a standard were injected.

2.4. Statistical analysis

The statistical program NWA Quality Analyst release 6.1 (2007) was used to study the accuracy and precision of the method applied.

2.5. QSPR studies

2.5.1. Data set of amino acids

The experimental amino acid concentration of malts and beers were measured in our laboratories, as indicated above. The concentration of each amino acid in the aminograms was expressed as the mean concentration range (ppm), which was then converted into logarithm units for modelling purposes [$\log_{10}$ range (ppm)].

2.5.2. Geometry optimisation and theoretical descriptors calculation

We pre-optimised the structure of each amino acid with the Molecular Mechanics Force Field (MM+) procedure included in the Hyperchem 6.03 package (Hypercube, Inc.). We chose the levogyre enantiomer because of its natural occurrence. The Semi-Empirical Method PM3 (Parametric Method-3) was used to refine the resulting geometries by the Polak–Ribière algorithm, and a gradient norm limit of 0.01 kcal/Å.

We computed 1497 theoretical molecular descriptors using the Dragon software (Milano Chemometrics and QSAR Research Group) including different types of descriptors, such as Constitutional, Topological, Geometrical, Charge, GEometry, Topology, and Atom-Weights AssemblY (GETAWAY), Weighted Holistic Invariant Molecular (WHIM) descriptors, 3D-Molecule Representation of Structures based on Electron diffraction (3D-MoRSE), Molecular Walk Counts, BCUT descriptors, 2D-Autocorrelations, Aromaticity Indices, Randic Molecular Profiles, Radial Distribution Functions, Functional Groups, Atom-Centred Fragments, Empirical and Property-based descriptors (Todeschini & Consonni, 2000). Finally, we added quantum-chemical descriptors to the pool, which had not been provided by the Dragon program, such as molecular dipole moments, total energies, homo–lumo energies, and homo–lumo gap ($\Delta_{\text{homo–lumo}}$).

2.5.3. Model search

We applied the Replacement Method (RM) as a selection approach for molecular descriptors. This algorithm, that had been proposed by our research group some years ago ((Duchowicz et al., 2005; Duchowicz, Mercader et al., 2008; Duchowicz, Talevi et al., 2008), is an efficient optimisation tool, which gives rise to multi-parametric linear regression QSPR models by searching the set D of D descriptors for an optimal subset d of $d<<D$ ones with minimum model's standard deviation S. The quality of the results achieved with this technique is quite close to that obtained by performing an exact (combinatorial) full search (FS) of molecular descriptors, although, of course, requires much less computational work. The RM provides models with better statistical parameters than the Forward Stepwise Regression procedure, and similar ones to the more elaborated Genetic Algorithms (Mitchell, 1998). We used the computer Matlab 5.0 system for all calculations (The MathWorks, Inc., 2004).

2.5.4. Model validation

The most important step for a QSPR analysis is to obtain properly validated models having predictive value of global applicability, and not limited to function only correlatively. The Leave-One-Out Cross Validation procedure (*loo*) was practiced over each linear regression. The parameters R_{loo} and S_{loo}, correlation coefficients, and standard deviation of Cross Validation each measured the stability of the developed QSPRs upon inclusion/exclusion of compounds. R_{loo} should be higher than 0.50 for obtaining a validated model (Golbraikh & Tropsha, 2002).

We also applied a rigorous and more realistic validation process that consisted of omitting from the complete molecular set some amino acids which constituted the 'test set', denoted here as *val*. This subdivision was performed in order to assess whether the found QSPRs had any predictive ability for estimating the concentrations on the independent test set of 'fresh' compounds, that were not involved during the model fitting using the 'training set' compounds, denoted as *train*. We first selected the molecules of both the training and test series, previously to model search, so that both sets shared similar qualitative structure–property features.

We checked the robustness of the structure–property equations established in this study by the so-called *y*-randomization (Wold, Eriksson, & Clementi, 1995) in order to demonstrate that these equations did not result from happenstance. The *y*-randomization technique consisted of scrambling the concentration values in such a way that they did not account for the respective amino acids. Upon analysing 5,000,000 cases of *y*-randomization for each developed QSPR, the smallest standard deviation obtained for the model by using this procedure (S_{rand}) was compared to that found when considering the true calibration (S). If $S_{rand} > S$, then it would be expected that the correlations found were not fortuitous, and resulted in actual structure–property relationships.

Furthermore, the Kubinyi function (*FIT*) (Kubinyi, 1994) has been proposed as a statistical parameter that closely related to the Fisher ratio (F), but avoided the main disadvantage of the latter that was too sensitive to changes in small d values and poorly sensitive to changes in large d values. The *FIT* criterion had a low sensitivity to changes in small d values and a substantially increasing sensitivity for large d values. The higher the *FIT* was, the better was the quality of the linear regression equation.

3. Results and discussion

3.1. Amino acid analysis in beer and malt samples

The amino acid content is a fingerprint for the authenticity of beverages such as malts and beers, and can be used as a specific biomarker to identify the end product.

The occurrence of some amino acids can be explained taking into account both the original presence in the wort, and/or the yeast consumption/release during beer fermentation. In other words, the nature and relative amount of the amino acids can be related to the wort composition as well as to beer fermentation conditions.

Experimental relative percentages of the free amino acids of commercial blonde and black beers, and malts are shown in Table 1. Brief experimental details have been indicated in Section 2.

According to these results, blonde beers contained approximately one third of total free amino acids of dark beers, and malts, as expected, contained 25% more amino acids than dark beers. Proline, as expected, is the main amino acid in these malts and beers. Proline concentration increases during barley germination, reaching the maximum level after 8 days. Furthermore, proline is only slightly uptaken from wort under anaerobic conditions, and slowly absorbed under aerobic conditions (Perpète, Santos, Bodart, & Collin, 2005). This is the reason of finding the highest proline concentration in the samples of blonde and black beers (Table 1), being higher in beers than in malts, followed by cystine, alanine, arginine, serine, valine, lysine and isoleucine in order of importance. These chemical parameters influence flavour, since yeast metabolises the amino acids by several biochemical mechanisms. As it is known, of all the biochemical processes taking place in beer production, proteolysis is primarily responsible, since it produces free amino acids and peptides that either contribute necessarily to flavour or are precursors behaving as substrates for various enzymatic reactions. It should be noted that amino acids also influence the colour of the final product, since during cooking of the amino acids, especially proline, they react with maltose (Maillard) yielding maltol and coloured products known as melanoi-

Table 1
Relative percentages of amino acids and ammonia in beer and malt samples.

Amino acids	Blonde beer (pale lager) samples (n = 16)	Black beer (dark lager) samples (n = 2)	Malt samples (n = 2)
Range of relative percentages			
Aspartic acid	2.9–0.7	3.5–1.6	3.2–1.9
Serine	4.8–1.5	8.2–4.5	7.5–4.6
Alanine	8.6–5.8	8.1–3.2	11.5–6.1
Cystine	18.9–7.9	7.3–3.1	11.2–6.1
Proline	43.2–25.9	45.7–30.8	26.5–20.8
Lysine	2.5–0.3	1.5–0.6	6.3–4.1
Arginine	6.7–2.8	4.8–2.9	7.5–4.1
Ammonia	3.7–1.5	3.4–1.9	4.8–2.7
Glutamic acid	1.8–0.6	3.1–1.9	2.1–1.0
Threonine	2.1–0.6	2.2–0.9	4.1–2.1
Isoleucine	5.0–1.5	2.1–0.8	5.8–3.8
Leucine	3.5–1.4	1.4–1.0	8.5–6.2
Phenylalanine	4.6–1.4	3.2–1.9	8.3–4.1
Tyrosine	6.1–2.2	4.5–1.9	6.3–4.1
Methionine	1.1–0.1	0.5–0.2	2.2–1.0
Valine	5.8–2.2	4.5–2.2	7.3–4.9
Histidine	3.3–1.3	3.3–2.1	2.6–1.2
Glycine	5.3–1.4	5.1–2.1	2.4–1.1

dines. Finally, the amino acids behave on the biological stability of the final product.

As it is known, there is also a direct relationship between free amino acids and the alcohols obtained after fermentation, this being one of the main topics of the final product quality.

Non-experts cannot easily make a direct comparison of the amino acid profiles of beer and malt samples due to the high similarity of some concentration ranges. Therefore, we tried to find out a way to transform mathematically these experimental values into models in order to establish whether these concentrations are of value for distinguishing the samples, and further to discriminate among the samples clearly. Hence, we applied the QSPR theory to the experimental amino acid concentrations in order to obtain the respective models. As expected we obtained three different QSPR equations, which accounted for malts, blonde and black beers, respectively, despite the commercial brand used.

3.2. QSPR analysis of amino acid profile in Argentine malts and beers

General QSPR equations were established on the tested amino acid concentration profiles in order to achieve predictive values on new fresh data. This predictive ability of each mapped QSPR equation was tested carrying out an internal validation (Leave-One-Out Cross Validation Technique), and also by an external validation (using some molecules as a test set, but not for data adjustment). Some molecules of the training set were removed, thus improving the model's quality performance. Furthermore, all models designed in this work followed the "Rule of Thumb" (Tute, 1990), which stated that at least five or six data points should be present for each fitting parameter.

Each QSPR described herein contained *N*, the number of amino acids; *R*, the correlation coefficient; *res*, the residual for a given amino acid (difference between the experimental and predicted concentration); *p*, the significance of the model. *FIT* accounted for the Kubinyi function (Kubinyi, 1994), and S_{rand} was the model's standard deviation of *y*-randomization. Decision criteria that were simultaneously analysed for determining the model's size of the training sets investigated, such as the determination of the optimal *d* to be included in each QSPR, were the following: (a) the lowest

Table 2
Brief description for the theoretical descriptors of the established QSPRs.

Descriptor	Category	Type	Brief description
BELp4	2D	BCUT	Lowest eigenvalue No. 4 of Burden matrix/weighted by atomic polarizabilities
RDF040e	3D	Radial Distribution Function	Radial Distribution Function-4.0/weighted by atomic Sanderson electronegativities
Mor16 m	3D	3D-MoRSE	3D-MoRSE-signal 16/weighted by atomic masses
Mor06v	3D	3D-MoRSE	3D-MoRSE-signal 06/weighted by atomic van der Waals volumes
LDip	3D	Charge	Local dipole index
HATS3v	3D	GETAWAY	Leverage-weighted autocorrelation of lag 3/weighted by atomic van der Waals volumes.

Table 3
Numerical values of the theoretical descriptors involved in various QSPRs designed in this study.

No.	Levo-amino acids	BELp4	RDF040e	Mor16 m	Mor06v	LDip	HATS3v
1	Aspartic acid	0.756	3.000	0.070	0.042	0.368	0.179
2	Serine	0.449	3.133	0.014	0.003	0.309	0.207
3	Alanine	0.384	4.648	−0.186	−0.077	0.272	0.229
4	Cystine	0.385	5.830	−0.036	0.097	0.286	0.233
5	Proline	0.473	9.877	−0.184	−0.453	0.222	0.225
6	Lysine	1.075	6.081	−0.181	0.289	0.179	0.116
7	Arginine	1.101	8.312	−0.234	0.216	0.179	0.105
8	Ammonia	0.000	0.000	0.012	−0.087	0.009	0.000
9	Glutamic acid	0.980	5.582	−0.116	0.039	0.323	0.148
10	Threonine	0.831	5.390	0.089	0.154	0.304	0.193
11[a]	Isoleucine	1.071	10.032	−0.044	0.167	0.205	0.142
12	Leucine	1.087	5.338	−0.119	0.201	0.201	0.132
13	Phenylalanine	1.027	8.737	0.083	−0.173	0.173	0.164
14	Tyrosine	1.053	8.482	−0.044	0.000	0.235	0.179
15	Methionine	1.051	4.137	0.073	0.403	0.248	0.131
16	Valine	1.070	7.084	−0.064	0.119	0.216	0.168
17[b]	GABA	0.596	4.363	−0.156	0.103	0.232	0.170
18	Histidine	1.084	6.897	0.298	−0.402	0.281	0.162
19[b]	Asparagine	0.871	11.199	−0.122	0.121	0.274	0.181
20	Glycine	0.241	0.932	0.026	−0.138	0.331	0.249

[a] Test set amino acid.
[b] Estimated amino acid.

value for the *S* parameter; (b) the lowest S_{loo} value; (c) the highest *FIT* parameter; (d) the lowest number of outlier amino acids exceeding 2.*S*, 2.5.*S* or 3.*S*; and (e) the lowest value for the maximal inter-correlation between descriptors in the model. A brief description of the theoretical descriptors involved in QSPR Eqs. (1)–(3), and the corresponding numerical values of these descriptors are shown in Tables 2 and 3, respectively.

3.2.1. Blonde (lager) beer samples

The structure–property relationship indicated in Eq. (1) was obtained.

$$\log_{10}(\text{mean rel.}\%) = -1.145(\pm 0.2)\cdot \text{BELp4} + 0.153(\pm 0.02)\cdot \text{RDF040e} + 0.526(\pm 0.1) \quad (1)$$

$d = 2$, $N_{train} = 16$, $N_{train}/d = 8$, $R = 0.911$, $S = 0.173$, $FIT = 3.158$, $p < 10^{-5}$, outliers (>2*S*) = 0, $R_{loo} = 0.872$, $S_{loo} = 0.206$, $S_{rand} = 0.260$

Eq. (1) was used to predict the mean relative percentages of the *test set* molecules **4** (cystine) and **11** (isoleucine) that were not taken into account for adjustment. Eq. (1) was also applied to estimate amino acids **17** (GABA) and **19** (asparagine), whose data were not available. The plot of predicted mean relative percentages *vs.* experimental mean relative percentages (Fig. 1a) for pale lager samples shows that the proposed model (Eq. (1)) fits a straight line for both the training and test set data. All predictions were included in Table 4.

3.2.2. Dark lager (bock) samples

The structure–property relationship, indicated in Eq. (2), was obtained.

$$\log_{10}(\text{mean rel.}\%) = -1.659(\pm 0.4)\cdot \text{Mor16m} - 1.642(\pm 0.2)\cdot \text{Mor06v} + 0.417(\pm 0.05) \quad (2)$$

$d = 2$, $N_{train} = 16$, $N_{train}/d = 8$, $R = 0.905$, $S = 0.196$, $FIT = 2.951$, $p < 10^{-5}$, outliers (>3*S*) = 0, $R_{loo} = 0.851$, $S_{loo} = 0.242$, $S_{rand} = 0.287$

Eq. (2) was used to predict the mean relative percentages of the test set molecules **4** (cystine) and **11** (isoleucine) that were not taken into account for adjustment. Eq. (2) was also applied to estimate the amino acids **17** (GABA) and **19** (asparagine), whose data were not available. The plot of predicted mean relative percentages *vs.* experimental mean relative percentages (Fig. 1b) for dark lager samples shows that the proposed model (Eq. (2)) fits a straight line for both the training and test set data. Predicted values are shown in Table 5.

3.2.3. Malt samples

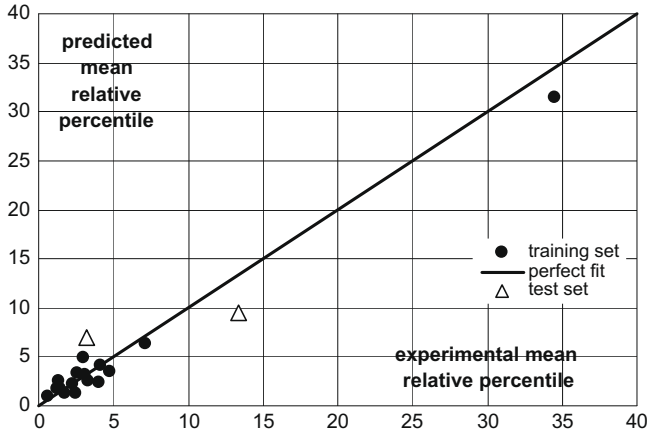
It was possible to take up the structure–property relationship with a simple model of two descriptors provided the molecule **20** (glycine) was removed from the *training set* (Eq. (3)).

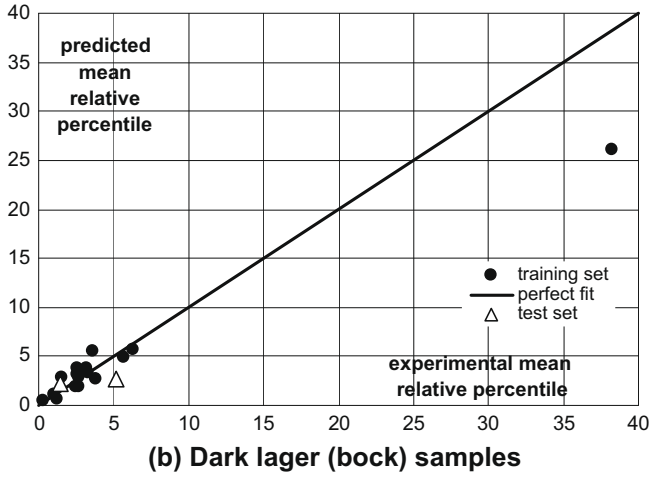
$$\log_{10}(\text{mean rel.}\%) = -4.438(\pm 0.8)\cdot \text{LDip} + 6.811(\pm 1)\cdot \text{HATS3v} + 0.638(\pm 0.1) \quad (3)$$

$d = 2$, $N_{train} = 15$, $N_{train}/d = 7.5$, $R = 0.867$, $S = 0.168$, $FIT = 1.919$, $p < 10^{-5}$, outliers (>2*S*) = 0, $R_{loo} = 0.794$, $S_{loo} = 0.207$, $S_{rand} = 0.177$

Eq. (3) was used to predict the mean relative percentages of the molecules of the *test set* **4** (cystine) and **11** (isoleucine) that were not taken into account for adjustment. Eq. (3) was also applied to estimate non-available data of amino acids **17** (GABA) and **19** (asparagine).

The plot of predicted mean relative percentages *vs.* experimental mean relative percentages (Fig. 1c) for malt samples shows that the proposed model (Eq. (3)) fits a straight line for both the training and test set data. Predicted values are shown in Table 6.





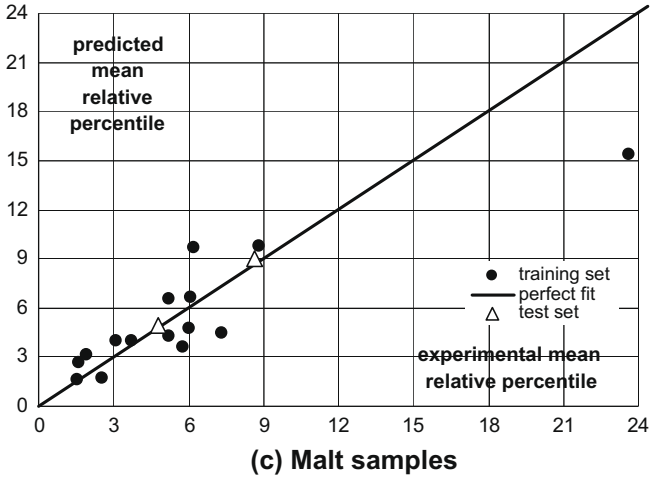


Fig. 1. Experimental and QSPR predicted mean relative percentages for pale lager, dark lager and malt samples analysed in this work. (a) Pale lager samples. (b) Dark lager samples. (c) Malt samples.

All these predictive linear QSPR models were able to capture the essential structural features of the amino acids related to their concentration in malts and beers (pale and dark lagers). It is worth to mention that these QSPRs involved a combination of 2D- and 3D-type molecular descriptors in order to achieve the best aminogram predictions.

Table 4
Experimental (Exp.) and Eq. (1) predicted (Pred.) mean amino acid relative percentages for pale lager samples.

No.	Levo-amino acids	Lowest limits	Higher limits	Exp. mean values	Pred. mean values	Differences
1	Aspartic acid	0.70	2.90	1.80	1.32	0.48
2	Serine	1.50	4.80	3.15	3.10	0.05
3	Alanine	5.80	8.60	7.20	6.29	0.91
4[a]	Cystine	7.90	18.90	13.40	9.52	3.88
5	Proline	25.90	43.20	34.55	31.49	3.06
6	Lysine	0.30	2.50	1.40	1.69	−0.29
7	Arginine	2.80	6.70	4.75	3.46	1.29
8	Ammonia	1.50	3.70	2.60	3.35	−0.75
9	Glutamic acid	0.60	1.80	1.20	1.82	−0.62
10	Threonine	0.60	2.10	1.35	2.52	−1.17
11[a]	Isoleucine	1.50	5.00	3.25	6.88	−3.63
12	Leucine	1.40	3.50	2.45	1.26	1.19
13	Phenylalanine	1.40	4.60	3.00	4.89	-1.89
14	Tyrosine	2.20	6.10	4.15	4.17	-0.02
15	Methionine	0.10	1.10	0.60	0.91	-0.31
16	Valine	2.20	5.80	4.00	2.44	1.56
17[b]	GABA	–	–	–	3.25	–
18	Histidine	1.30	3.30	2.30	2.20	0.10
19[b]	Asparagine	–	–	–	17.59	–
20	Glycine	1.40	5.30	3.35	2.47	0.88

[a] Test set amino acid.
[b] Estimated amino acid.

Table 5
Experimental (Exp.) and Eq. (2) predicted (Pred.) mean amino acid relative percentages for dark lager samples.

No.	Levo-amino acids	Lowest limits	Higher limits	Exp. mean values	Pred. mean values	Differences
1	Aspartic acid	1.60	3.50	2.55	1.70	0.85
2	Serine	4.50	8.20	6.35	2.45	3.90
3	Alanine	3.20	8.10	5.65	7.11	−1.46
4[a]	Cystine	3.10	7.30	5.20	2.07	3.13
5	Proline	30.80	45.70	38.25	29.23	9.02
6	Lysine	0.60	1.50	1.05	1.75	−0.70
7	Arginine	2.90	4.80	3.85	2.82	1.03
8	Ammonia	1.90	3.40	2.65	3.46	−0.81
9	Glutamic acid	1.90	3.10	2.50	3.51	−1.01
10	Threonine	0.90	2.20	1.55	1.04	0.51
11[a]	Isoleucine	0.80	2.10	1.45	1.64	−0.19
12	Leucine	1.00	1.40	1.20	1.92	−0.72
13	Phenylalanine	1.90	3.20	2.55	3.66	−1.11
14	Tyrosine	1.90	4.50	3.20	3.09	0.11
15	Methionine	0.20	0.50	0.35	0.43	−0.08
16	Valine	2.20	4.50	3.35	2.12	1.23
17[b]	GABA	–	–	–	3.21	–
18	Histidine	2.10	3.30	2.70	3.82	−1.12
19[b]	Asparagine	–	–	–	2.63	–
20	Glycine	2.10	5.10	3.60	3.98	−0.38

[a] Test set amino acid.
[b] Estimated amino acid.

4. Conclusions

A large world consumption of malts as well as blonde and black beers justifies an exhaustive chemical study addressed to enhance quality and authenticity control.

Suitable experimental conditions have been previously obtained in our laboratories (Giraudo et al., 2004) for the amino acid analysis of malts and beers. According to our findings, it is important to get the respective aminograms in order to establish a specific pattern or fingerprint for the characterisation of each malt and beer.

Amino acid concentration ranges were further submitted to QSPR-modelling. Then, the goal of the present work was to develop for the first time predictive QSPR theoretical models for the amino acid concentrations, and further to contribute to the discrimination among malts, blonde and black beers.

Therefore, QSPR analysis was applied in order to calculate discrimination functions for the samples and to find the most useful descriptors in the differentiation among sample equations. The results obtained from QSPR can be considered very satisfactory, since QSPR equations were able to evaluate quantitative differences in the amino acid profiles of the commercial beers and malts investigated, so allowing not only to discriminate between malts and beers, but also the type of beer (pale or dark lagers).

QSPR models had a predicted value, and were also used to estimate non-available amino acid concentrations for the samples described herein.

QSPR showed to be a suitable tool for data discrimination and visualisation of similarities among malt and beer samples.

This is a new application of the QSPR theory to food, in particular to chemical biomarkers of malts and beers.

Table 6
Experimental (Exp.) and Eq. (3) predicted (Pred.) mean amino acid relative percentages for malt samples.

No.	Levo-amino acids	Lowest limits	Higher limits	Exp. mean values	Pred. mean values	Differences
1	Aspartic acid	1.9	3.2	2.55	1.68	0.87
2	Serine	4.6	7.5	6.05	4.75	1.30
3	Alanine	6.1	11.5	8.80	9.80	−1.00
4[a]	Cystine	6.1	11.2	8.65	9.04	−0.39
5	Proline	20.8	26.5	23.65	15.34	8.31
6	Lysine	4.1	6.3	5.20	4.31	0.89
7	Arginine	4.1	7.5	5.80	3.62	2.18
8	Ammonia	2.7	4.8	3.75	3.97	−0.22
9	Glutamic acid	1.0	2.1	1.55	1.63	−0.08
10	Threonine	2.1	4.1	3.10	4.02	−0.92
11[a]	Isoleucine	3.8	5.8	4.80	4.96	−0.16
12	Leucine	6.2	8.5	7.35	4.42	2.93
13	Phenylalanine	4.1	8.3	6.20	9.72	−3.52
14	Tyrosine	4.1	6.3	5.20	6.53	−1.33
15	Methionine	1.0	2.2	1.60	2.69	−1.09
16	Valine	4.9	7.3	6.10	6.67	−0.57
17[b]	GABA	–	–	–	5.85	–
18	Histidine	1.2	2.6	1.90	3.12	−1.22
19[b]	Asparagine	–	–	–	4.52	–
20	Glycine	1.1	2.4	1.75	–	–

[a] Test set amino acid.
[b] Estimated amino acid.

Acknowledgements

We thank the National Research Council of Argentina [Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina], Universidad Nacional de La Plata, Universidad Nacional de Lanús, and Universidad de Buenos Aires (Argentina) for financial support; and Ministerio de Ciencia y Técnica (Mincyt, Argentina) for electronic bibliography facilities. MAG is Chairman of Food Science and Food Technology of Universidad Nacional de Lanús (Argentina). ABP, EAC, and PRD are Research Members of CONICET. None of the authors had any personal or financial conflict of interest.

References

Chicu, S. A., & Putz, M. V. (2009). Köln-Timişoara molecular activity combined models toward interspecies toxicity assessment. *International Journal of Molecular Sciences, 10*(10), 4474–4497. doi:10.3390/ijms10104474.

Donalies, U. E., Nguyen, H. T., Stahl, U., & Nevoigt, E. (2008). Improvement of *Saccharomyces* yeast strains used in brewing, wine making and baking. *Advances in Biochemical Engineering/Biotechnology, 111*, 67–98. doi:10.1007/978-3-540-70536-9.

Duchowicz, P. R., Castro, E. A., Fernández, F. M., & González, M. P. (2005). A new search algorithm of QSPR/QSAR theories: Normal boiling points of some organic molecules. *Chemical Physics Letters, 412*(4–6), 376–380. doi:10.1016/j.cplett.2005.07.016.

Duchowicz, P. R., Mercader, A. G., Fernández, F. M., & Castro, E. A. (2008). Prediction of aqueous toxicity for heterogeneous phenol derivatives by QSAR. *Chemometrics and Intelligent Laboratory Systems, 90*(2), 97–107. doi:10.1016/j.chemolab.2007.08.006.

Duchowicz, P. R., Talevi, A., Bruno-Blanch, L. E., & Castro, E. A. (2008). New QSPR study for the prediction of aqueous solubility of drug-like compounds. *Bioorganic and Medicinal Chemistry, 16*(17), 7944–7955. doi:10.1016/j.bmc.2008.07.067.

Elfakir, C. (2005). 1.2. HPLC. In I. Molnár-Perl (Ed.), *Quantitation of amino acids and amines by chromatography: Methods and protocols*. *Journal of Chromatography Library* (vol. 70, pp. 119–267). Amsterdam, Boston, Heidelberg: Elsevier. ISBN 0-444 52050-3.

Giraudo, M., Sánchez Tuero, H., Pomilio, A., Pavesi, R., Markowski, I., Guirin, G., et al. (2004). Determination of free amino acids and ammonium in blonde and black beers and malts of the argentinean market by RP-HPLC with fluorescent detection. *Alimentaria, 359*, 85–89. ISSN: 0300-5755. INIST-CNRS, Cote INIST: 21624, 35400012136686.0140..

Golbraikh, A., & Tropsha, A. (2002). Beware of q^2! *Journal of Molecular Graphics and Modelling, 20*(4), 269–276. doi:10.1016/S1093-3263(01)00123-1.

Gorinstein, S., Caspi, A., Libman, I., Leontowicz, H., Leontowicz, M., Tashma, Z., et al. (2007). Bioactivity of beer and its influence on human metabolism. *International Journal of Food Sciences and Nutrition, 58*(2), 94–107. doi:10.1080/09637480601108661.

Hansch, C., & Leo, A. (1995). *Exploring QSAR: Fundamentals and applications in chemistry and biology*. Washington, DC: American Chemical Society.

Iñón, F. A., Garrigues, S., & de la Guardia, M. (2006). Combination of mid- and near-infrared spectroscopy for the determination of the quality properties of beers. *Analytica Chimica Acta, 571*(2), 167–174. doi:10.1016/j.aca.2006.04.070.

Jernigan, D. H. (2009). The global alcohol industry: An overview. *Addiction, 104*(Suppl. 1), 6–12. doi:10.1111/j.1360-0443.2008.02430.x.

Katritzky, A. R., Lobanov, V. S., & Karelson, M. (1995). QSPR: The correlation and quantitative prediction of chemical and physical properties from structure. *Chemical Society Reviews, 24*(4), 279–287. doi:10.1039/cs9952400279.

Kubinyi, H. (1994). Variable selection in QSAR studies II. A highly efficient combination of systematic search and evolution. *Quantitative Structure–Activity Relationships, 13*(4), 393–401. doi:10.1002/qsar.19940130403.

Lacrămă, A.-M., Putz, M. V., & Ostafe, V. (2007). A Spectral-SAR model for the anionic–cationic interaction in ionic liquids: Application to *Vibrio fischeri* ecotoxicity. *International Journal of Molecular Sciences, 8*(8), 842–863. doi:10.3390/i8080842.

Llario, R., Iñón, F. A., Garrigues, S., & de la Guardia, M. (2006). Determination of quality parameters of beers by the use of attenuated total reflectance–Fourier transform infrared spectroscopy. *Talanta, 69*(2), 469–480. doi:10.1016/j.talanta.2005.10.016.

Mitchell, M. (1998). *An introduction to genetic algorithms*. 0-262-63185-7. Cambridge, MA: The MIT Press.

Molnár-Perl, I. (2005). Quantitation of amino acids and amines by chromatography: Methods and protocols. *Journal of Chromatography Library, 70*, 7–9. ISBN 0-444-52050-3.

Perpète, P., Santos, G., Bodart, E., & Collin, S. (2005). Uptake of amino acids during beer production: The concept of a critical time value. *Journal of American Society of Brewing Chemists, 63*(1), 23–27. doi:10.1094/ASBCJ-63-0023.

Pomilio, A. B. (1994). Separación de aminoácidos mediante cromatografía líquida de alta resolución. *Acta Bioquímica Clínica Latinoamericana, 28*(4), 527–554. ISSN: 0325-2957.

Pomilio, A. B. (2004). *Actualización en cromatografía líquida de alta resolución. Aplicaciones clínicas y medioambientales*. Serie Química. Buenos Aires: IDECEFYN. ISBN: 987-98643-7-9.

Putz, M. V., & Lacrămă, A.-M. (2007). Introducing spectral structure activity relationship (S-SAR) analysis. Application to ecotoxicology. *International Journal of Molecular Sciences, 8*(5), 363–391. doi:10.3390/i8050363.

Putz, M. V., Putz, A.-M., Lazea, M., Ienciu, L., & Chiriac, A. (2009). Quantum-SAR extension of the spectral-SAR algorithm. Application to polyphenolic anticancer bioactivity. *International Journal of Molecular Sciences, 10*(3), 1193–1214. doi:10.3390/ijms10031193.

Rainieri, S., Kodama, Y., Kaneko, Y., Mikata, K., Nakao, Y., & Ashikari, T. (2006). Pure and mixed genetic lines of *Saccharomyces bayanus* and *Saccharomyces pastorianus* and their contribution to the lager brewing strain genome. *Applied and Environmental Microbiology, 72*(6), 3968–3974. doi:10.1128/AEM.02769-05.

Todeschini, R., & Consonni, V. (2000). *Handbook of molecular descriptors. Book series: Methods and principles in medicinal chemistry*. Weinheim: Wiley-VCH.

Tute, M. S. (1990). History and objectives of quantitative drug design. In C. Hansch, P. G. Sammes, & J. B. Taylor (Eds.), *Comprehensive medicinal chemistry. Quantitative drug design* (Vol. 4, pp. 1–31). Oxford: Pergamon Press.

Wold, S., Eriksson, L., & Clementi, S. (1995). Statistical validation of QSAR results. In H. van de Waterbeemd (Ed.), *Chemometric methods in molecular design. Book series: Methods and principles in medicinal chemistry* (pp. 309–318). Weinheim: VCH (Chapter 5, ISBN 3-527-30044-9).