Talanta 130 (2014) 128–136

Contents lists available at ScienceDirect

Talanta

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

Box–Behnken factorial design to obtain a phenolic-rich extract from the aerial parts of *Chelidonium majus* L.

Clara Grosso [a], Federico Ferreres [b,*], Angel Gil-Izquierdo [b], Patrícia Valentão [a], Maria Sampaio [a], Júlio Lima [a], Paula B. Andrade [a,*]

[a] *REQUIMTE/Laboratório de Farmacognosia, Departamento de Química, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, no. 228, 4050-313 Porto, Portugal*

[b] *Research Group on Quality, Safety and Bioactivity of Plant Foods, Department of Food Science and Technology, CEBAS (CSIC), P.O. Box 164, 30100 Campus University Espinardo, Murcia, Spain*

ARTICLE INFO

Article history:
Received 28 April 2014
Received in revised form
13 June 2014
Accepted 19 June 2014
Available online 27 June 2014

Keywords:
Box–Behnken design
Chelidonium majus
HPLC–DAD–ESI/MS^n
Flavonoids

ABSTRACT

A Box–Behnken design (BBD) was developed to study the influence of four parameters (X_1: % methanol; X_2: extraction time; X_3: extraction temperature; X_4: solid/solvent ratio) on two responses, namely extraction yield and phenolics content of the aerial parts of *Chelidonium majus* L. The model presented a good fit to the experimental results for the extraction yield, being significantly influenced by X_1 and X_4. On the other hand a parameter reduction was necessary to run the model for phenolics content, showing that only X_1 and X_2 had great influence on the response. Two best extraction conditions were defined: X_1=76.8% MeOH, X_2=150.0 min, X_3=60.0 °C and X_4=1:100 and X_1=69.2%, X_2=150 min, X_3=42.5 °C and X_4=1:100.

Moreover, the HPLC–DAD–ESI/MS^n analysis conducted with the center point sample revealed the presence of 15 alkaloids and 15 phenolic compounds, from which the 9 flavonoids and 3 hydroxycinnamic acids are described for the first time. Only phenolic compounds were quantified by a validated HPLC–DAD method, the pair quercetin-3-*O*-rutinoside+quercetin-3-*O*-glucoside dominating all the 29 extracts. This study is of great importance for future works that seek to apply the phenolic profile to the quality control of *C. majus* samples.

1. Introduction

The need to optimize process variables is emerging in different fields of research and industry, due to the increasing demand for fast, simple and efficient methodologies with reduced costs and wastes. The use of statistical tools, such as response surface methodology (RSM), allows finding the best set of independent variables or factors (input variables) that produce the optimum response (output variable). The advantage of this kind of approach is that it enables obtaining more information about the variables and their interactions with fewer experiments than the traditional univariate procedures [1,2]. Box–Behnken design (BBD), a type of RSM, is a second-order multivariate technique based on three-level incomplete factorial design. BBD has been widely applied in the past decade to optimize the extraction procedure of bioactive compounds from natural sources, such as phenolic compounds [3–8].

The determination of phenolic compounds involves a general analytical strategy that, besides a recovery step, includes their structural characterization and quantification. Among the analytical methodologies available, the most widely employed are based on reversed-phase high-performance liquid chromatography (RP-HPLC) coupled to diode array detection (DAD) and/or mass spectrometry (MS) with atmospheric pressure ionization techniques, i.e., electrospray ionization (ESI) or atmospheric pressure chemical ionization (APCI). HPLC–MS, particularly HPLC coupled to tandem MS (HPLC–MS^n), has been recognized as the best tool to analyze biological samples due to its selectivity, sensitivity and speed of analysis [9,10]. HPLC–MS^n allows obtaining more information concerning molecular structure, such as the type of aglycone moiety and substituents present, the stereochemical assignment of terminal monosaccharide units, the sequence of the glycan component, the interglycosidic linkages and the points of attachment of the substituents to the aglycone [9].

The aim of this study was to develop a BBD to optimize the extraction of phenolic compounds from *Chelidonium majus* L. and to perform a systematic characterization of its phenolic profile by HPLC–DAD–ESI/MS^n. Only some free hydroxycinnammic (caffeic, *p*-coumaric and ferulic acids) and hydroxybenzoic (gentisic and

* Corresponding authors. Tel.: +351 220428654; fax: +351 226093390.
E-mail addresses: federico@cebas.csic.es (F. Ferreres), pandrade@ff.up.pt (P.B. Andrade).

http://dx.doi.org/10.1016/j.talanta.2014.06.043
0039-9140/

p-hydroxybenzoic acids) acids, as well as flavonoids aglycones (quercetin, kaempferol and apigenin) were previously identified by HPLC–DAD, after enzymatic hydrolysis [11,12]. Additionally, Hahn and Nahrstedt [11] isolated and characterized four hydroxycinnamic acid derivatives, namely (−)-2-(*E*)-caffeoyl-D-glyceric acid, (−)-4-(*E*)-caffeoyl-L-threonic acid, (−)-2-(*E*)-caffeoyl-L-threonic acid lactone and (+)-(*E*)-caffeoyl-L-malic acid, by NMR. On the other hand, the most studied bioactive compounds from this species are the isoquinoline alkaloids, which are found in high amounts and revealed antimicrobial, anticancer, anti-inflammatory, adrenolytic, sympatholytic, anticholinesterase and anti-MAO-A activities [13–17].

Taking into account that the most important factors affecting the extraction performance are the type of solvent, extraction time, extraction temperature and solid/solvent ratio [5], our first purpose was to develop a four-factor BBD to find the optimum extraction conditions for obtaining phenolic compounds from *C. majus*. Moreover, the second goal was to develop and validate an HPLC–DAD method for phenolic compounds quantification.

2. Materials and methods

2.1. Plant material

C. majus aerial parts (Lot. 1044 05 13) were purchased from Morais e Costa & CA. Lda (Porto, Portugal). After being powdered to a mean particle size below 910 μm, the plant material was stored desiccated at room temperature, protected from light. The voucher specimen (Cm-A-042013) was deposited at the Laboratory of Pharmacognosy of the Faculty of Pharmacy of Porto University.

2.2. Chemicals and standards

Methanol (MeOH) Lichrosolv and acetic acid 100% were purchased from Merck (Darmstadt, Germany) and VWR (Fontenay-sous-Bois, France), respectively. Caffeic acid and quercetin-3-*O*-rutinoside were obtained from Sigma-Aldrich (St. Louis, MO, USA) and kaempferol-3-*O*-rutinoside, isorhamnetin-3-*O*-glucoside and isorhamnetin-3-*O*-rutinoside were from Extrasynthèse (Genay, France).

2.3. Extraction procedure

All extractions were performed with ca. 1 g of plant material. Different combinations of four parameters were tested to extract phenolic compounds from *C. majus*, namely % of MeOH (0–100%), extraction time (30 min of sonication, followed by 30–90 min of stirring maceration plus 30 min of sonication), temperature (25–60 °C) and solid/solvent ratio (1:50–1:150). After, the solvent was evaporated under reduced pressure and the extracts were stored at −20 °C until further use.

Table 1
Coded and uncoded levels of the four independent variables.

Independent variables	Low		Center		High	
X_1 (% MeOH)	−1	(0%)	0	(50%)	+1	(100%)
X_2 (extraction time)	−1	(90 min)	0	(120 min)	+1	(150 min)
X_3 (extraction temperature)	−1	(25 °C)	0	(42.5 °C)	+1	(60 °C)
X_4 (solid/solvent ratio)	−1	(1:50)	0	(1:100)	+1	(1:150)

2.4. Factorial design

The software Design Expert (version 6.0.8, Stat-Ease Inc., Minneapolis, MN, USA) was used for experimental design, data analysis and model building.

BBD was employed to find the optimum extraction conditions for obtaining the highest extraction yield and amount of phenolic compounds.

BBD requires an experiment number according to

$$N = 2k(k-1)+cp$$

where k is the number of factors (or independent variables) and cp is the number of the center points.

The independent variables chosen for this study, namely percentage of MeOH (X_1), extraction time (X_2), extraction temperature (X_3) and solid–solvent ratio (X_4), were evaluated at three different levels (−1, 0, 1) and coded according to the following equation:

$$x_i = \frac{X_i - X_0}{\Delta X} \quad i = 1,2,3,4$$

where x_i is the coded value of an independent variable, X_i is the actual value of an independent variable, X_0 is the actual value of an independent variable at the center point and ΔX is the step change value of an independent variable. The coded and uncoded levels of the four independent variables are given in Table 1. In total, 29 experiments were performed in triplicate, with five repetitions of the center point (Table 2).

In order to predict the optimal responses (extraction yield and content of phenolic compounds), the following second-order polynomial equation was used to fit the experimental data:

$$Y = \alpha_0 + \sum_{i=1}^{4} \alpha_i X_i + \sum_{i=1}^{4} \alpha_{ii} X_i^2 + \sum_{i=1}^{3} \sum_{j=i+1}^{4} \alpha_{ij} X_i X_j$$

Table 2
Extraction yields (%) and phenolic compounds content (mg/kg of extract) as a function of the four independent variables.

Run	X_1[a]	X_2	X_3	X_4	Extraction yield	Phenolics content
1	100% MeOH	90	42.5	1:100	10.4	2544.8
2	100% Water	90	42.5	1:100	21.3	372.1
3	100% MeOH	150	42.5	1:100	10.3	3746.8
4	100% Water	150	42.5	1:100	21.7	828.2
5	50% MeOH	120	25.0	1:50	14.6	4537.0
6	50% MeOH	120	60.0	1:50	15.6	2496.5
7	50% MeOH	120	25.0	1:150	21.5	6703.4
8	50% MeOH	120	60.0	1:150	20.9	3731.9
9	50% MeOH	120	42.5	1:100	21.5	3612.5
10	100% MeOH	120	25.0	1:100	9.4	3004.0
11	100% Water	120	25.0	1:100	21.8	324.0
12	100% MeOH	120	60.0	1:100	10.3	4397.3
13	100% Water	120	60.0	1:100	20.5	304.9
14	50% MeOH	90	42.5	1:50	15.3	2034.9
15	50% MeOH	150	42.5	1:50	16.3	3567.3
16	50% MeOH	90	42.5	1:150	22.6	2644.4
17	50% MeOH	150	42.5	1:150	21.4	4639.1
18	50% MeOH	120	42.5	1:100	19.3	4200.3
19	100% MeOH	120	42.5	1:50	9.8	2781.1
20	100% Water	120	42.5	1:50	17.1	432.3
21	100% MeOH	120	42.5	1:150	11.4	1934.7
22	100% Water	120	42.5	1:150	22.4	306.3
23	50% MeOH	90	25.0	1:100	17.6	3468.9
24	50% MeOH	150	25.0	1:100	20.7	3664.4
25	50% MeOH	90	60.0	1:100	17.6	4239.7
26	50% MeOH	150	60.0	1:100	21.7	3985.3
27	50% MeOH	120	42.5	1:100	20.4	3331.0
28	50% MeOH	120	42.5	1:100	18.5	2973.0
29	50% MeOH	120	42.5	1:100	19.6	2954.2

[a] X_1 – solvent (% MeOH); X_2 – extraction time (min); X_3 – extraction temperature (°C); X_4 – solid/solvent ratio.

where Y represents the response variables, α_0 is a constant, α_i, α_{ii} and α_{ij} are the linear, quadratic and interactive coefficients, respectively. X_i and X_j are the independent variables [5,18].

Only for the amount of phenolic compounds a transformation squared root was required, since the ratio of maximum and minimum response was higher than 10.

2.5. HPLC–DAD–ESI/MS^n qualitative analysis

The center point extract was dissolved in MeOH/water mixture (1:1) and submitted to sonication, centrifugation at 12,000 rpm and filtration through 0.2 μm membrane.

Chromatographic analyses were carried out on a Luna C_{18} (2) 100A column (150 × 1.0 mm, 3 μm particle size; Phenomenex, Macclesfield, UK). The mobile phase consisted of two solvents: water (1% acetic acid) (A) and methanol (B), starting with 20% B and using a gradient to obtain 50% B at 30 min and 80% B at 40 min. The flow rate was 20 μL/min and the injection volume 3 μL. Spectral data from all peaks were accumulated in the range 240–400 nm and chromatograms were recorded at 280, 320 and 350 nm. The HPLC–DAD–ESI/MS^n analyses were carried out in an Agilent HPLC 1200 series equipped with a diode array detector and mass detector in series (Agilent Technologies, Waldbronn, Germany). The HPLC consisted of a binary pump (model G1376A), an autosampler (model G1377A) refrigerated at 4 °C (G1330B), a degasser (model G1379B) and a photodiode array detector (model G1315D). The HPLC system was controlled by ChemStation software (Agilent, v. B.01.03-SR2). The mass detector was a Bruker ion trap spectrometer (model HCT Ultra) equipped with an electrospray ionization interface and was controlled by LCMSD software (Agilent, v. 6.1). The ionization conditions were adjusted at 300 °C and 4.0 kV for capillary temperature and voltage, respectively. The nebulizer pressure and flow rate of nitrogen were 5.0 psi and 3 L/min, respectively. The full scan mass covered the range from m/z 100 up to m/z 1200. Collision-induced fragmentation experiments were performed in the ion trap using helium as the collision gas, with voltage ramping cycles from 0.3 up to 2 V. Mass spectrometry data were acquired in the negative ionization mode for the study of phenolic compounds and in the positive mode for alkaloids. MS^2 was carried out in the automatic mode.

2.6. HPLC–DAD quantitative analysis

For phenolic compounds quantification, 20 μL of each extract (29 in total) were analyzed in triplicate on an analytical HPLC unit (Gilson), using a Luna C_{18} column (250 × 4.60 mm, 5 μm particle size; Phenomenex, Torrance, CA, USA). The mobile phase consisted of water (1% acetic acid) (A) and methanol (B), starting with 20% B and using a gradient to obtain 50% B at 30 min and 80% B at 40 min. The flow rate was 0.8 mL/min. Detection was achieved with a Gilson diode array detector. Spectral data from all peaks were collected in the range of 200–400 nm. The data were processed on Unipoint System software (Gilson Medical Electronics, Villiers le Bel, France). Peak purity was checked by the software contrast facilities. Phenolic compounds quantification was achieved by the absorbance recorded in the chromatograms at 320 nm (for hydroxycinnamic acids) and 350 nm (for flavonoids) relative to calibration curves carried out with five concentrations (in triplicate) of each standard.

Compounds **9**, **12**, **13** and **22** were not quantified because they were present in trace amounts. Compounds **19** and **25** were quantified as kaempferol-3-*O*-rutinoside, the pair **20**+**21** as quercetin-3-*O*-rutinoside, compounds **23**, **24** (or **23**+**24**) and **26**+**Ac**+**27** as caffeic acid, compound **28** as isorhamnetin-3-*O*-glucoside and compound **29** as isorhamnetin-3-*O*-rutinoside.

With the purpose of validating the HPLC–DAD method, linearity, limit of detection (LOD), limit of quantification (LOQ), precision and accuracy were determined. Linearity was evaluated from the correlation coefficients (R^2) of the regression curves obtained for each standard.

LOD and LOQ were calculated from the residual standard deviation (σ) of the regression curves and the slopes (S), according to the following equations: LOD = 3.3σ/S and LOQ = 10σ/S.

Precision (reproducibility) was determined by calculating the relative standard deviation (R.S.D.) from repeated injections of the sample corresponding to the center point (extract 18). Intraday precision was calculated from five replicate injections performed in the same day, while interday precision was determined with five injections done in 5 consecutive days.

For accuracy (recovery) evaluation, extracts at the center point were spiked with isorhamnetin-3-*O*-glucoside at three different levels: low (1.6×10^{-3} mg/mL), medium (1.6×10^{-2} mg/mL) and high (8.2×10^{-2} mg/mL). This compound was chosen because it was the only one with available standard that did not co-eluted with others.

3. Results and discussion

3.1. Phenolic compounds and alkaloids characterization

The screening of the hydromethanolic extract of the aerial parts of *C. majus* by RP-HPLC–DAD–ESI/MS^n revealed a chromatographic profile (350 and 280 nm) in (Fig. 1) which first part corresponded to peaks with UV spectra of alkaloids and that ionized in the positive mode. The second part of the chromatogram showed

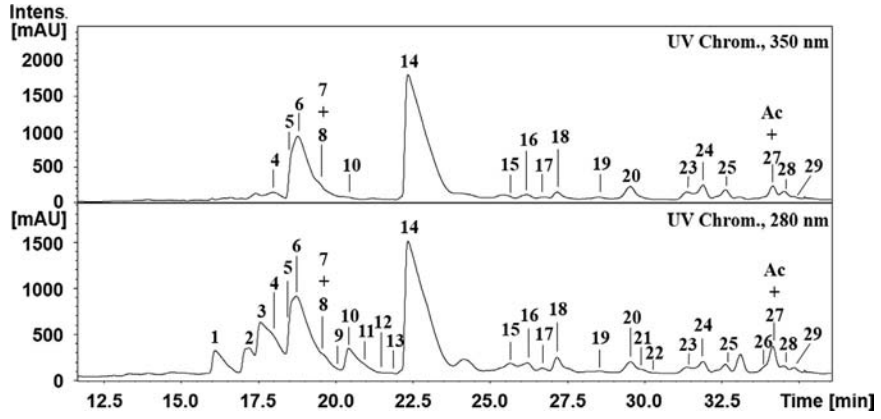

Fig. 1. HPLC-UV (350 and 280 nm) profile of *C. majus* extract obtained with extraction conditions at the center point. Identity of compounds as in Tables 1 and 2. Ac: uncharacterized cinnamoyl acid derivative.

other kind of compounds, less abundant or in trace amounts, possessing UV spectra of flavonoids and that suffered ionization in the negative mode.

3.1.1. Alkaloids

The study of alkaloids was carried out by comparison of their UV and MS spectra with literature data [19–26]. As so, dihydroberberine (**1**), protopine (**2**), allocryptopine (**3**), chelidonine (**4**), coptisine (**6**), tetrahydrocoptisine (**10**), tetrahydroberberine (**11**), berberine (**14**), norchelidonine (**15**), chelerythrine (**18**), which have been already described in this species [27–32, among others], were tentatively characterized. Besides these, other alkaloids were also detected, though their complete characterization was not possible: alkaloid A (**5**), alkaloid B (**7**), alkaloid C (**8**), alkaloid D (**16**) and alkaloid E (**17**) (Fig. 1, Table 3).

3.1.2. Phenolic compounds

As indicated above, flavonoids were found in low or even trace amounts. Therefore the majority of their UV spectra could not be properly observed, although band I was at ca. 350 nm (Table 4), indicating that the hydroxyl at position 3 of the flavonol was not free. The structural characterization of the flavonoids was mainly based on their MS spectra. Four diglycosides (**19**, **20**, **25** and **29**) and five monoglycosides (**21**, **22** and **26**–**28**) (Fig. 1, Table 4) were detected, corresponding to kaempferol (**19**, **25** and **27**) ([Aglyc-H]- at *m/z* 285), quercetin (**20**–**22** and **26**) ([Aglyc-H]$^-$ at *m/z* 301) and isorhamnetin (**28** and **29**) ([Aglyc-2H/H]$^-$ at *m/z* 314/315) derivatives (Table 4). In the MS^2 of the diglycoside **19**, besides the ion of the deprotonated aglycone (*m/z* 285), ions at *m/z* 447 (base peak) and 431 were also observed, which correspond to the loss of rhamnosyl (−146) and hexosyl (−162) radicals, respectively. The loss of those radicals was not accompanied by the loss of water (−164/−180), indicating the absence of interglycosidic linkage [33,34]. This fragmentation is typical of di-*O*-glycosides with sugar residues linked to different phenolic hydroxyls. Moreover, the preferential fragmentation of flavonol-3,7-di-*O*-glycosydes occurs at the C7-OH, giving rise to the base peak [35] and, therefore, compound **19** can be tentatively characterized as kaempferol-3-*O*-hexoside-7-*O*-rhmanoside. Kite and Veitch [36] differentiated kaempferol-3-*O*-glucoside-7-*O*-rhamnoside from kaempferol-3-*O*-galactoside-7-*O*-rhamnoside by the relative abundance of the ion at *m/z* 431, which corresponds to ca. 50% for the glucosyl derivative, while for the galactosyl one it is $<10\%$. Thus, compound **19** can be kaempferol-3-*O*-glucoside-7-*O*-rhamnoside. The MS fragmentation pattern of compound **25** was similar to that of compound **19**, but the two sugars were equal (rhamnose) and the ion at *m/z* 431 ([(M-H)-146]$^-$) was observed as base peak (Table 3). So, it can be labeled as kaempferol-3,7-di-*O*-rhmanoside (**25**). The diglycosides **20** and **29** are two rhamnosyl-hexosides, and in their MS fragmentations practically just the ion of their deprotonated aglycones (base peak) was observed, at *m/z* 301 (quercetin) and 315 (isorhamnetin), respectively. This type of diglycosides fragmentation, in which ions resulting from the break

Table 3
*R*t, UV and MS: $[M]^+/[M+H]^+$ and $MS^2[M]^+/[M+H]^+$ data of alkaloids from *C. majus*. [a]

Compound		*R*t (min)	UV (nm)	$[M]^+/[M+H]^+$ *m/z*	$MS^2[M]^+/[M+H]^+$ *m/z* (%)
1	Dihydroberberine	16.1	290	338	190(100)
2	Protopine	17.1	288	354	190(100), 149(80)
3	Allocryptopine	17.5	–	370	352(50), 188(100)
4	Chelidonine	17.9	–	354	323(100), 305(40), 295(10), 275(50)
5	Alkaloid A	18.4	–	338	190(100)
6	Coptisine	18.7	268, 348, 360, 460	320	293(100)
7	Alkaloid B	19.4	–	370	339(100), 321(25), 290(50)
8	Alkaloid C	19.6	–	354	190(100)
10	Tetrahydrocoptisine	20.4	288, 348	324	176(100), 149(60)
11	Tetrahydroberberine	20.9	–	340	176(100), 149(20)
14	Berberine	25.5	264, 276sh, 338, 346, 426	336	321(100), 292(10)
15	Norchelidonine	25.7	290	340	322(100)
16	Alkaloid D	26.3	290, 320	356	338(100), 308(40)
17	Alkaloid E	26.7	252, 300, 380	368	350(100), 306(20), 276(10)
18	Chelerythrine	27.2	272, 285sh, 320, 338, 426	348	333(100), 318(10), 304(20)

[a] Main observed fragments. Other ions were found but they have not been included.

Table 4
*R*t, UV and MS: $[M-H]^-$ and $MS^2[M-H]^-$ data of flavonoids from *C. majus*.[a]

Compound[b]		*R*t (min)	UV (nm)	$[M-H]^-$ *m/z*	$MS^2[M-H]^-$, *m/z* (%)		
					-146	-162	$[Aglc-H/2H]^-$
19	kaempf-3-glc-7-rhmn	28.4	–	593	447(100)	431(47)	285(33)
20	querct-3-rhmn(1→6)glc	29.4	256, 264sh, 298sh, 354	609			301(100)
21	querct-3-glc	29.8	–	463			301(100)
22	querct-3-gluc	30.5	–	477			301(100)
25	kaempf-3,7-di-rhmn	32.5	264, 316sh, 348	577	431 (100)		285(50)
26	querct-3-rhmn	33.7	–	447			301(100)
27	kaempf-3-glc	34.0	–[c]	447			285(100)
28	isorhamnt-3-glc	34.4	–	477			314(100)
29	isorhamnt-3-rhmn(1→6)glc	34.7	–	623			315(100)

[a] Main observed fragments. Other ions were found but they have not been included.
[b] kaempf: kaempferol; querct: quercetin; isorhamnt: isorhamnetin; glc: glucoside; gluc: glucuronide; rhmn: rhamnoside.
[c] Compound **27** co-eluted with a cinnamoyl acid derivative not fully characterized, which did not allowed to observe its UV spectrum.

of the interglycosidic linkage are not observed, is typical of the 1→6 linkage, which is hard to be broken [33,34]. Compound **20** chromatographically matched with the rutin (quercetin-3-*O*-rhamnosyl(1→6)glucoside) standard. Thus, and taking into account its chromatographic mobility, compound **29** can be labeled as isorhamnetin-3-*O*-rhamnosyl(1→6)glucoside.

As above indicated, the remaining flavonoids are monoglycosides. Compounds **21**, **27** and **28** presented similar MS spectrum, characterized by the loss of a 162 amu fragment (hexosyl radical) to give rise to the ions of their deprotonated aglycones as base peak (**21**, *m/z* 301 [quercetin-H]$^-$; **27**, *m/z* 285 [kaempferol-H]$^-$; **28**, *m/z* 314 [isorhamnetin-2H]$^-$). The chromatographic mobility in reverse phase of compound **21** relative to that of rutin (**20**) indicates that it should be quercetin-3-*O*-glucoside. Moreover, the respective chromatographic mobility of compounds **21**, **27** and **28** showed that they are substituted by the same hexose. Thus, compounds **27** and **28** can be labeled as kaempferol-3-*O*-glucoside and isorhamnetin-3-*O*-glucoside, respectively. The MS fragmentation of compound **22** exhibited the loss of 176 amu (glucuronoyl radical), indicating that it is quercetin-3-*O*-glucuronide. The loss of

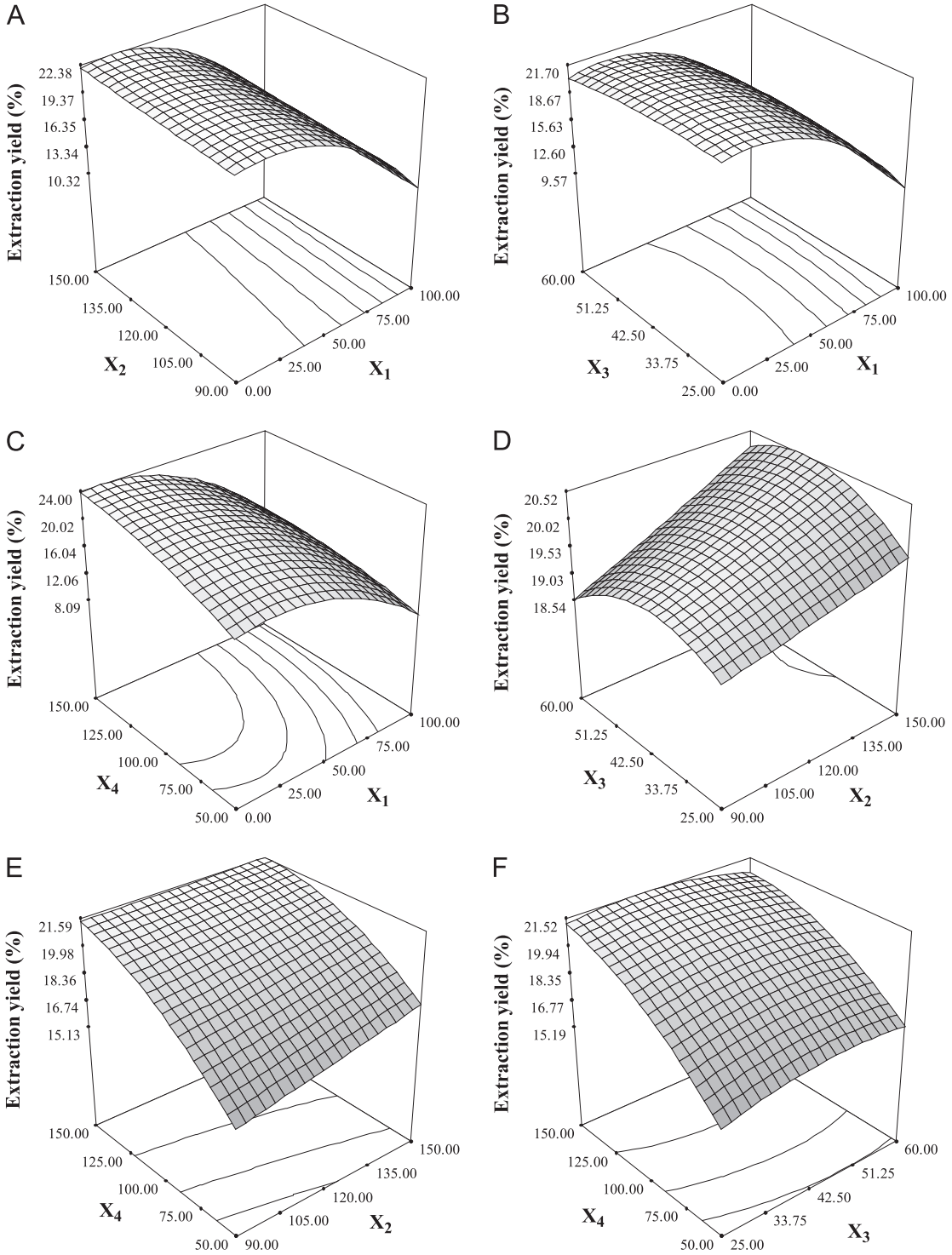

Fig. 2. Three-dimensional (3D) response surface and contour plots for extraction yield. (A) $X_1 \times X_2$; (B) $X_1 \times X_3$; (C) $X_1 \times X_4$; (D) $X_2 \times X_3$; (E) $X_2 \times X_4$; and (F) $X_3 \times X_4$.

Table 5
Regression equation, R^2, LOD and LOQ of the standards.

Compounds[a]	Regression equation	R^2	Range of concentrations (mg/mL)	LOD[b] (mg/mL)	LOQ[c] (mg/mL)
caffeic acid	$y=3.8\times10^{9}x-3.8\times10^{6}$	0.9925	5.3×10^{-5}–5.3×10^{-2}	1.1×10^{-5}	3.3×10^{-5}
querct-3-*O*-rut	$y=7.1\times10^{8}x-2.8\times10^{6}$	0.9959	5.2×10^{-4}–5.2×10^{-1}	7.9×10^{-5}	2.4×10^{-4}
isorhamnt-3-*O*-glc	$y=7.5\times10^{8}x-2.3\times10^{6}$	0.9975	1.6×10^{-4}–1.6×10^{-1}	4.9×10^{-6}	1.5×10^{-5}
isorhamnt-3-*O*-rut	$y=5.7\times10^{8}x-2.1\times10^{5}$	0.9955	4.0×10^{-4}–4.0×10^{-1}	1.2×10^{-4}	3.6×10^{-4}
kaempf-3-*O*-rut	$y=8.3\times10^{8}x-1.9\times10^{6}$	0.9959	3.4×10^{-4}–3.4×10^{-1}	7.7×10^{-5}	2.3×10^{-4}

[a] querct: quercetin; isorhamnt: isorhamnetin; kaempf: kaempferol; rut: rutinoside; glc: glucoside.
[b] Limit of detection.
[c] Limit of quantification.

Table 6
Extraction yields (%) and phenolic compounds content (mg/kg of extract) as a function of the four independent variables.[a]

Run	Compound 19	Compounds 20+21	Compound 23	Compound 24	Compound 25	Compounds 26+Ac+27	Compound 28	Compound 29	Total (mg/kg extract)
1	n.q.	928.2 ± 24.5	49.8 ± 0.9	119.5 ± 2.7	446.7 ± 37.2	366.2 ± 20.4	277.7 ± 7.0	356.7 ± 37.2	2544.8
2	67.4 ± 3.6	109.8 ± 12.9	17.9 ± 0.3	With **23**	n.q.	22.5 ± 1.0	60.3 ± 3.9	94.2 ± 49.1	372.1
3	157.5 ± 0.3	1872.2 ± 46.5	28.3 ± 7.1	With **23**	132.0 ± 0.2	502.8 ± 77.2	454.0 ± 11.9	600.0 ± 101.8	3746.8
4	82.4 ± 15.8	327.4 ± 51.3	n.q.	n.q.	43.4 ± 3.0	56.4 ± 4.9	118.1 ± 9.1	200.5 ± 18.0	828.2
5	256.4 ± 73.0	2718.8 ± 181.1	18.2 ± 0.4	38.4 ± 3.1	114.8 ± 6.7	250.9 ± 17.7	420.9 ± 32.3	718.6 ± 56.1	4537.0
6	232.0 ± 15.6	1139.5 ± 71.0	104.1 ± 3.8	With **23**	265.9 ± 14.3	198.2 ± 10.5	176.3 ± 13.1	380.5 ± 49.0	2496.5
7	298.7 ± 11.4	4018.5 ± 274.6	34.7 ± 1.8	With **23**	122.5 ± 23.1	321.4 ± 64.0	645.5 ± 13.3	1262.1 ± 109.9	6703.4
8	155.0 ± 1.4	2149.0 ± 171.8	28.1 ± 1.9	With **23**	120.7 ± 8.4	192.2 ± 18.8	371.3 ± 10.7	715.6 ± 73.7	3731.9
9	362.2 ± 49.0	1624.8 ± 49.9	275.8 ± 20.5	With **23**	445.7 ± 31.7	159.9 ± 3.6	295.6 ± 73.4	448.5 ± 55.2	3612.5
10	275.1 ± 34.1	941.9 ± 42.2	57.9 ± 10.2	172.3 ± 4.8	361.4 ± 17.2	375.7 ± 8.7	284.0 ± 25.6	535.7 ± 63.6	3004.0
11	41.4 ± 5.3	94.9 ± 13.1	18.2 ± 0.4	With **23**	57.3 ± 3.7	19.5 ± 0.5	77.2 ± 3.4	15.5 ± 3.5	324.0
12	114.7 ± 27.7	2348.4 ± 77.4	25.6 ± 1.0	With **23**	141.9 ± 6.3	486.2 ± 118.5	592.4 ± 16.4	688.1 ± 58.2	4397.3
13	37.6 ± 0.8	130.4 ± 12.1	n.q.	n.q.	n.q.	16.1 ± 0.3	58.6 ± 11.7	62.2 ± 4.0	304.9
14	137.4 ± 17.9	888.4 ± 52.7	18.6 ± 0.2	39.9 ± 1.6	85.4 ± 6.9	274.5 ± 14.5	224.0 ± 9.9	366.7 ± 18.6	2034.9
15	210.1 ± 15.4	1984.7 ± 211.8	19.9 ± 2.9	With **23**	84.9 ± 4.4	213.7 ± 15.3	361.5 ± 27.4	692.5 ± 70.7	3567.3
16	217.0 ± 1.9	1209.8 ± 20.2	128.0 ± 3.1	With **23**	241.6 ± 5.0	196.7 ± 6.0	224.1 ± 39.4	427.2 ± 96.7	2644.4
17	258.6 ± 34.4	2775.4 ± 198.0	30.4 ± 1.1	With **23**	98.4 ± 1.4	209.6 ± 19.8	438.4 ± 10.4	828.3 ± 33.2	4639.1
18	444.4 ± 82.4	2043.4 ± 137.5	210.8 ± 10.4	With **23**	484.7 ± 44.0	174.9 ± 8.0	283.9 ± 59.3	558.2 ± 20.6	4200.3
19	165.1 ± 14.2	1249.1 ± 150.7	n.q.	n.q.	132.2 ± 5.7	461.9 ± 37.2	363.4 ± 39.4	409.4 ± 73.0	2781.1
20	85.9 ± 18.3	170.4 ± 38.1	n.q.	n.q.	44.8 ± 5.3	17.1 ± 0.3	71.4 ± 5.7	42.7 ± 10.4	432.3
21	172.5 ± 30.4	590.8 ± 110.8	87.4 ± 8.2	With **23**	257.8 ± 9.1	378.8 ± 30.2	188.5 ± 36.9	258.9 ± 11.9	1934.7
22	60.2 ± 16.0	84.0 ± 17.8	17.2 ± 1.9	With **23**	48.4 ± 4.9	18.4 ± 2.0	60.6 ± 3.6	17.5 ± 2.1	306.3
23	382.4 ± 104.0	1549.8 ± 243.7	183.4 ± 17.6	With **23**	495.9 ± 71.4	158.6 ± 18.9	267.2 ± 23.0	431.6 ± 57.0	3468.9
24	201.1 ± 50.1	2059.3 ± 128.8	25.0 ± 0.3	With **23**	79.6 ± 0.8	229.8 ± 3.1	373.8 ± 40.6	695.8 ± 25.9	3664.4
25	171.4 ± 55.2	2541.1 ± 209.9	24.9 ± 0.3	With **23**	86.6 ± 0.5	217.0 ± 4.7	432.4 ± 15.2	766.3 ± 30.6	4239.7
26	202.0 ± 11.2	2359.6 ± 138.4	25.3 ± 0.7	With **23**	82.1 ± 2.7	200.8 ± 4.4	386.0 ± 47.1	729.5 ± 35.2	3985.3
27	338.9 ± 5.9	1697.7 ± 39.2	191.0 ± 4.4	With **23**	353.1 ± 4.1	127.3 ± ± 13.0	193.3 ± 9.0	429.7 ± 13.3	3331.0
28	297.5 ± 8.8	1508.2 ± 188.9	158.5 ± 7.7	With **23**	320.6 ± 11.8	147.1 ± 5.6	189.9 ± 1.5	351.2 ± 13.8	2973.0
29	259.1 ± 1.4	1592.5 ± 128.0	172.5 ± 15.1	With **23**	304.1 ± 18.8	135.1 ± 13.2	160.7 ± 1.6	330.2 ± 39.4	2954.2
Intradays (CV,[b] %)	1.8	1.6	3.3	With **23**	4.1	5.9	10.9	1.8	
Interdays (CV,[b] %)	5.7	5.0	2.9	With **23**	0.2	9.1	18.3	12.3	
Recovery (%)									
8.2×10^{-2} mg/mL							84.7		
1.6×10^{-2} mg/mL							84.0		
1.6×10^{-3} mg/mL							91.7		

[a] n.q. – not quantified.
[b] CV – coefficient of variation.

146 amu (rhamnosyl radical) observed in the MS fragmentation of compound **26** points to quercetin-3-*O*-rhamnoside.

Additionally, three caffeic acid derivatives already described in this species [11] were also detected in our extract, in trace amounts: (**9**) caffeoyl threonic acid (*R*t: 19.9 min; MS: 297 $[M-H]^-$; MS^2(297): 179 (30%, [caffeic acid-H]$^-$), 135(100%, [threonic acid-H]$^-$)), (**12**) caffeoyl glyceric acid (*R*t: 21.3 min; MS: 267 $[M-H]^-$; MS^2(267): 179 (35%, [caffeic acid-H]$^-$), 161 (100%, [caffeic acid-H-18]$^-$), 105 (95%, [glyceric acid-H]$^-$)) and (**13**) caffeoyl malic acid (*R*t: 21.7 min; MS: 295 $[M-H]^-$; MS^2(295): 179 (100%, [caffeic acid-H]$^-$), 133 (85%, [malic acid-H]$^-$)).

Also other two caffeoyl acid derivatives isomers were observed (**23** and **24**) at 31.2 and 31.8 min, displaying the same UV spectrum (294sh, 330 nm) and MS spectrum (MS: 359 [M–H]$^-$; MS^2(359): 223 (20%), 197 (30%), 179 (30%), 161 (100%), which are coincident with those of rosmarinic acid; as so, they can be tentatively considered as rosmarinic acid isomers.

3.2. Phenolic compounds quantification

Phenolic acids and flavonoids quantification were carried out by using external calibration curves. The range of concentrations used, regression equations, R^2, LOD and LOQ values are shown in Table 5. Taking into account these parameters, the HPLC–DAD method showed good linearity and sensitivity.

Intraday (repeatability) and interday (intermediate precision) coefficients of variation were calculated to ascertain the precision of the method. For intraday precision they were below 10%,

showing that the repeatability of the method is good, but for interday precision the coefficients were higher (< 18%) (Table 6). The accuracy of the HPLC–DAD method was in general good, with recoveries higher than 80% for isorhamnetin-3-*O*-glucoside (**28**) (Table 6).

Although all extracts are similar from the qualitative point of view, quantitative differences were observed among them. The methanolic and hydromethanolic extracts were richer in phenolic compounds (in the range of 1934.6 and 6703.3 mg/kg of extract) than the water extracts (between 304.9 and 828.0 mg/kg of extract) (Table 6). Anyway, in all cases the dominant compounds were the pair **20**+**21**.

3.3. Optimization of extraction parameters for phenolic compounds

A 4-factor, 3-level BBD was developed in order to optimize the extraction of phenolic compounds from *C. majus*. Two different responses were considered, namely the extraction yield and the phenolic content.

Concerning the extraction yield (Table S1, Fig. 2), it was clear that two linear (X_1 and X_4) and one quadratic (X_1^2) parameters were significant at the level of $p < 0.0001$, while all the other parameters were not significant ($p > 0.05$). The second-order polynomial equation determined from the experimental results in terms

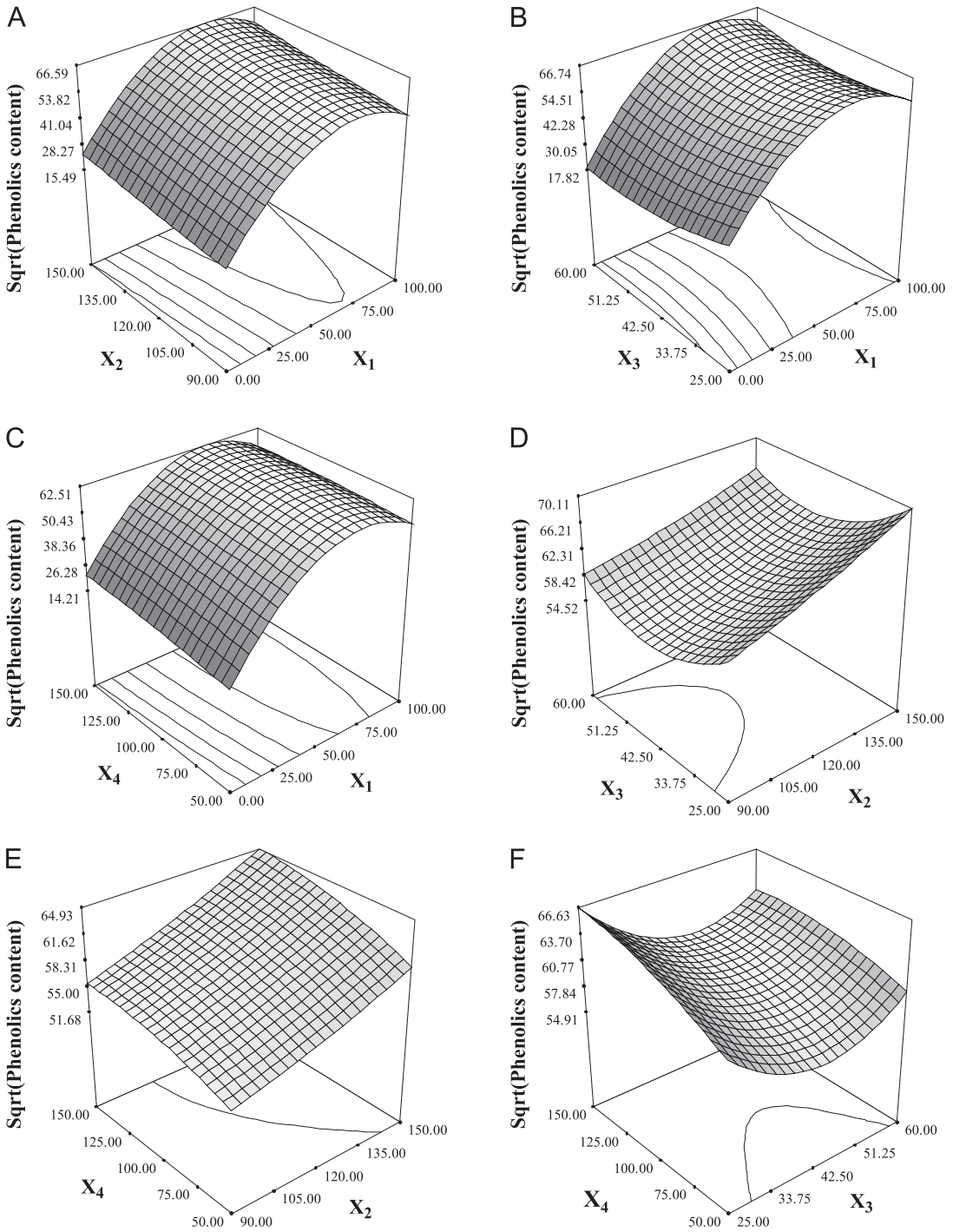

Fig. 3. Three-dimensional (3D) response surface and contour plots for phenolics content. (A) $X_1 \times X_2$; (B) $X_1 \times X_3$; (C) $X_1 \times X_4$; (D) $X_2 \times X_3$; (E) $X_2 \times X_4$; (F) $X_3 \times X_4$.

of coded factors was as follows:

$$\begin{aligned}\text{Extractionyield} = {} & 19.84 - 5.26X_1 + 0.60X_2 + 0.08X_3 + 2.63X_4 \\ & - 3.81{X_1}^2 + 0.03{X_2}^2 - 0.58{X_3}^2 - 0.99{X_4}^2 \\ & - 0.13X_1X_2 + 0.55X_1X_3 - 0.93X_1X_4 + 0.24X_2X_3 \\ & - 0.52X_2X_4 - 0.38X_3X_4\end{aligned}$$

The analysis of variance (ANOVA) result of the model is shown in Table S1 and demonstrates that the model is highly significant. The model presented a high value of correlation coefficient (R^2), adjusted R^2 and predicted R^2, indicating a good correlation between the experimental and predicted values of the response. Moreover, the statistical analysis gave high significant level ($p < 0.0001$) for the model and a non-significant value for lack-of-fit ($p = 0.4411$), revealing that the model can adequately fit the experiment data Fig. 2.

Concerning to phenolics content (Fig. 3), the model showed a worse fit to the experimental results and required a squared root (Sqrt) transformation. Although the lack of fit was not significant ($p = 0.0696$), only two parameters, X_1 and X_1^2, were statistically significant at $p < 0.0001$, indicating that the other three factors are insignificant model terms. In addition, values of R^2, adjusted R^2 and predicted R^2 were lower than those obtained for the response "extraction yield" (Table S2, condition A). Therefore, we decided to reduce the number of parameters to improve the model and we had run again the BBD with just three extraction parameters, resulting in three different designs with 17 experiments each

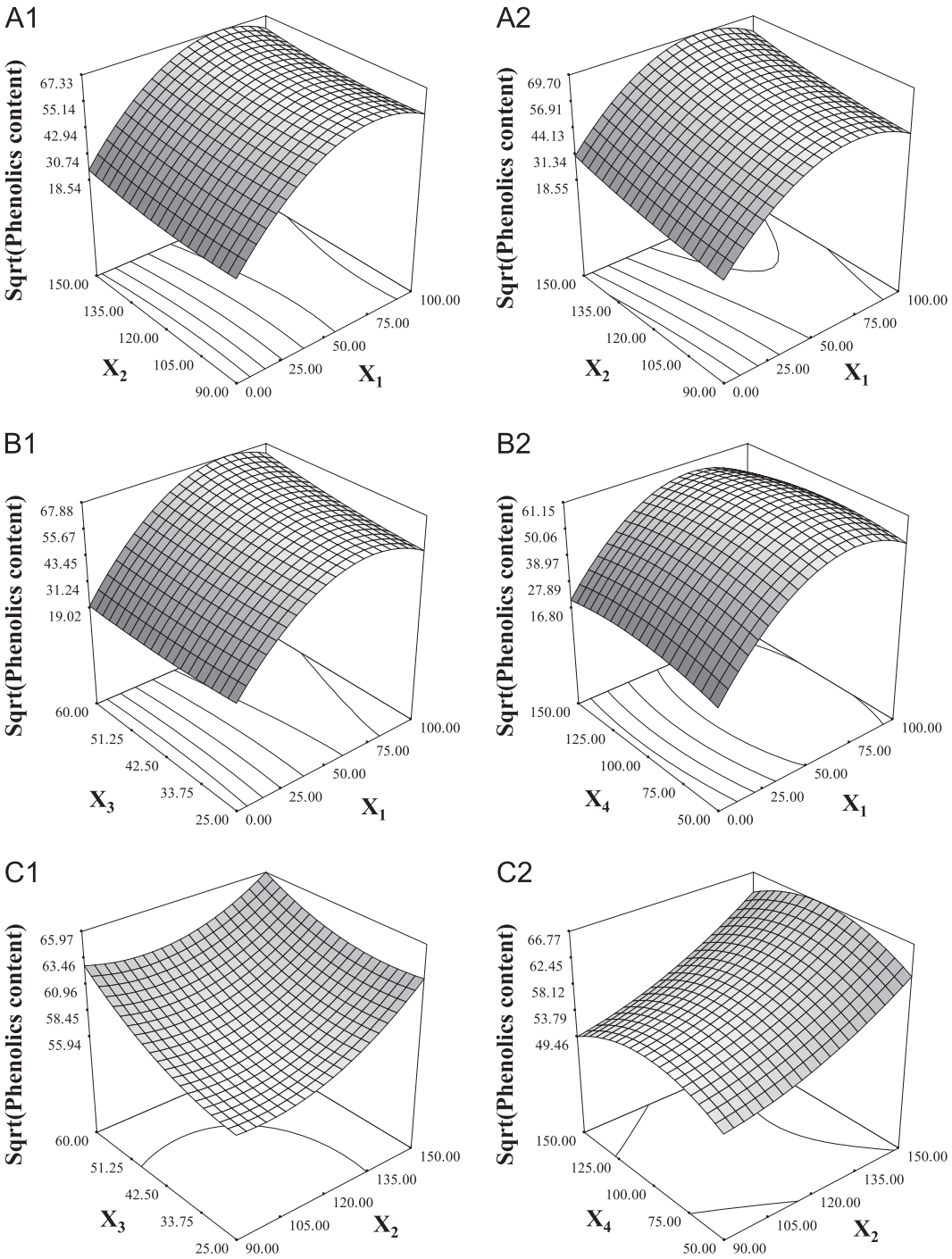

Fig. 4. Three-dimensional (3D) response surface and contour plots for phenolics content with only three parameters. A1–C1: X_1, X_2 and X_3; A2–C2: X_1, X_2 and X_4.

(Fig. 4). Excepting that performed with the independent factors X_1, X_3 and X_4 (X_2 fixed at 120 min), the lack of fit in the other two models was not significant ($p > 0.05$).

As for the model built with the four parameters, the one with just X_1, X_2 and X_3 (X_4 fixed at 1:100) revealed to be significant ($p=0.0004$) and highlighted the influence of the percentage of methanol (X_1 and X_1^2, $p < 0.0001$). The corresponding values of R^2, adjusted R^2 and predicted R^2 were higher than those obtained for the four-parameter model, being 0.9612, 0.9113 and 0.6022, respectively (Table S2, condition B); Fig. 4.

The second-order polynomial equation determined from the experimental results in terms of coded factors was as follows:

$$\begin{aligned}\text{Sqrt(phenolics content)} = {} & 58.30 + 18.66X_1 + 2.49X_2 + 2.47X_3 \\ & - 20.57{X_1}^2 + 2.20{X_2}^2 + 1.42{X_3}^2 \\ & + 0.32X_1X_2 + 3.01X_1X_3 - 0.90X_2X_3\end{aligned}$$

Better results were observed with the three-parameter model constructed with X_1, X_2 and X_4 (X_3 fixed at 42.5 °C). In this model ($p=0.0006$) more terms were significant, namely, X_1 ($p < 0.0001$), X_2 ($p=0.0093$) and X_1^2 ($p < 0.0001$). However, lower correlation coefficients were obtained: $R^2=0.9568$; ajusted $R^2=0.9012$; predicted $R^2=0.5658$ (Table S2, condition C); Fig. 4.

The second-order polynomial equation determined from the model incorporating X_1, X_2 and X_4 parameters in terms of coded factors is the following one:

$$\begin{aligned}\text{Sqrt(phenolicscontent)} = {} & 58.30 + 15.25X_1 + 6.44X_2 + 0.33X_4 \\ & - 20.35{X_1}^2 + 1.98{X_2}^2 - 4.19{X_4}^2 \\ & + 0.32X_1X_2 - 1.37X_1X_4 + 0.52X_2X_4\end{aligned}$$

Taken together these four models designed to explain the phenolics content, it can be concluded that the temperature (X_3) and the solid/solvent ratio (X_4) parameters do not influence the extraction of flavonoids and phenolic acids from *C. majus*.

3.4. Selection of optimum levels to maximize the responses

The selection of the optimum conditions for the four parameters in order to obtain the maximum responses was performed using Derringer's desirability function.

In order to obtain the maximum yield ($Y=24.3\%$, desirability$=1$), the best combination was found to be $X_1=7.0\%$ MeOH, $X_2=90.0$ min, $X_3=27.0$ °C and $X_4=1{:}150$.

Regarding the response "phenolics content" only the best conditions for the models incorporating just three parameters are shown. For situation $X_1-X_2-X_3$ (sqrt(phenolics content)$=71.9$ and fixing X_4 at 1:100, a combination of $X_1=76.8\%$ MeOH, $X_2=150.0$ min and $X_3=60.0$ °C is required. Finally, conditions of $X_1=69.2\%$, $X_2=150$ min and $X_4=1{:}100$ (with X_3 fixed at 42.5 °C) maximize sqrt(phenolics content) in the model using X_1, X_2 and X_4.

4. Conclusions

This is the first report on the systematic study of the phenolic composition of *C. majus*. Apart from fifteen alkaloids, from which five not fully characterized are described for the first time, six hydroxycinnamic acids, three of them not reported before in this species, and nine new flavonoids were characterized by HPLC–DAD–ESI/MS^n. The HPLC–DAD method for phenolics quantification was validated, showing good linearity and repeatability and satisfactory interday precision and recovery.

The Box–Behnken design developed to obtain an extract rich in phenolic compounds highlighted the importance of the % of methanol and the solid/solvent ratio to maximize the extraction yield, while to increase the phenolics content high % of methanol and time of extraction are required.

Acknowledgments

The authors are grateful to the financial support from the European Union (FEDER funds through COMPETE) and National Funds (FCT, Fundação para a Ciência e Tecnologia) through project Pest-C/EQB/LA0006/2013 and from the European Union (FEDER funds) under the framework of QREN through Project NORTE-07-0124-FEDER-000069, to CYTED Program (Ref.112RT0460) CORNUCOPIA Thematic Network and project AGL2011-23690(CICYT). Clara Grosso thanks FCT for the Post-Doc fellowship (SFRH/BPD/63922/2009).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2014.06.043.

References

[1] M.A. Bezerra, R.E. Santelli, E.P. Oliveira, L.S. Villar, L.A. Escaleira, Talanta 76 (2008) 965–977.
[2] L.V. Candioti, M.M. De Zan, M.S. Cámara, H.C. Goicoechea, Talanta 124 (2014) 123–138.
[3] J. Yu, D.V. Dandekar, R.T. Toledo, R.K. Singh, B.S. Patil, Food Chem. 105 (2007) 1026–1031.
[4] W. Liu, Y. Yu, R. Yang, C. Wan, B. Xu, S. Cao, Int. J. Mol. Sci. 11 (2010) 4750–4763.
[5] G. Zhang, L. He, M. Hu, Innov. Food Sci. Emerg. Technol. 12 (2011) 18–25.
[6] L.-C. Zhao, J. Liang, W. Li, K.-M. Cheng, X. Xia, X. Deng, G.-L. Yang, Molecules 16 (2011) 5928–5937.
[7] G. Pan, G. Yu, C. Zhu, J. Qiao, Ultrason. Sonochem. 19 (2012) 486–490.
[8] P.M. Jeganathan, S. Venkatachalam, T. Karichappan, S. Ramasamy, Prep. Biochem. Biotechnol. 44 (2014) 56–67.
[9] A. Ćirić, H. Prosen, M. Jelikić-Stankov, P. Đurđević, Talanta 99 (2012) 780–790.
[10] B. Abad-García, S. Garmón-Lobato, L.A. Berrueta, B. Gallo, F. Vicente, Talanta 99 (2012) 213–224.
[11] R. Hahn, A. Nahrstedt, Planta Med. 59 (1993) 71–75.
[12] A. Wojdyło, J. Oszmiański, R. Czemerys, Food Chem. 105 (2007) 940–949.
[13] S.S. Lee, M. Kai, M.K. Lee, Phytother. Res. 1 (2001) 167–169.
[14] F. Meng, G. Zuo, X. Hao, G. Wang, H. Xiao, J. Zhang, G. Xu, J. Ethnopharmacol. 125 (2009) 494–496.
[15] Q. Zhou, Y. Liu, X. Wang, X. Di, J. Mass Spectrom. 48 (2013) 111–118.
[16] Á. Sárközi, G. Janicsák, L. Kursinszki, Á. Kéry, Chromatographia 63 (2006) S81–S86.
[17] L. Cahlíková, L. Opletal, M. Kurfürst, K. Macáková, A. Kulhánková, A. Hostálková, Nat. Prod. Commun. 5 (2010) 1751–1754.
[18] C.-H. Dong, X.-Q. Xie, X.-L. Wang, Y. Zhan, Y.-J. Yao, Food Bioprod. Process 87 (2009) 139–144.
[19] Y.R. Chen, K.C. Wen, G.R. Her, J. Chromatogr. A 866 (2000) 273–280.
[20] N. Fabre, C. Claparols, S. Richelme, M.-L. Angelin, I. Fourasté, C. Moulis, J. Chromatogr. A 904 (2000) 35–46.
[21] Z.-H. Cheng, Y.-L. Guo, H.-Y. Wang, G.-Q. Chen, Anal. Chim. Acta 555 (2006) 269–277.
[22] M. Liang, W. Zhang, J. Hua, R. Liu, C. Zhang, J. Pharm. Biomed. Anal. 42 (2006) 178–183.
[23] X.-B. Luo, B. Chen, S.-Z. Yao, Phytochem. Anal. 17 (2006) 431–438.
[24] L. Ren, X. Xue, F. Zhang, Q. Xu, X. Liang, J. Sep. Sci. 30 (2007) 833–842.
[25] P. Deevanhxay, M. Suzuki, N. Maeshibu, H. Li, K. Tanaka, S. Hirose, J. Pharm. Biomed. Anal. 50 (2009) 413–425.
[26] C. Wang, S. Wang, G. Fan, H. Zou, Anal. Bioanal. Chem. 396 (2010) 1731–1740.
[27] C.Q. Niu, L.Y. He, J. Chromatogr. A 542 (1991) 193–199.
[28] F. Tomè, M.L. Colombo, Phytochemistry 40 (1995) 37–39.
[29] M.L. Colombo, E. Bosisio, Pharmacol. Res. 33 (1996) 127–134.
[30] S. Sturm, H. Stuppner, Electrophoresis 19 (1998) 3026–3032.
[31] Y.F. Kopyt'ko, T.D. Dargaeva, T.A. Sokol'skaya, E.I. Grodnitskaya, A.A. Kopnin, Pharm. Chem. J. 39 (2005) 40–45.
[32] N.S. Vachnadze, A.D. Bozhadze, D.T. Berashvili, A.J. Bakuridze, V.Y. Vachnadze, Chem. Nat. Compd. 48 (2012) 921.
[33] F. Cuyckens, Y.L. Ma, G. Pocsfalvi, M. Claeys, Analusis 28 (2000) 888–895.
[34] F. Cuyckens, R. Rozenberg, E. Hoffmann, M. Claeys, J. Mass Spectrom. 36 (2001) 1203–1210.
[35] F. Ferreres, R. Llorach, A. Gil-Izquierdo, J. Mass Spectrom. 39 (2004) 312–321.
[36] G.C. Kite, N.C. Veitch, Rapid Commun. Mass Spectrom. 23 (2009) 3125–3132.