Journal of Pharmaceutical and Biomedical Analysis 221 (2022) 115037

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.journals.elsevier.com/journal-of-pharmaceutical-and-biomedical-analysis

Application of experimental design in HPLC method optimisation for the simultaneous determination of multiple bioactive cannabinoids

Caterina Durante [a,*], Lisa Anceschi [b,c], Virginia Brighenti [b], Clarissa Caroli [b,c], Cindy Afezolli [b], Andrea Marchetti [a], Marina Cocchi [a], Stefano Salamone [d,e], Federica Pollastro [d,e], Federica Pellati [b,*]

[a] Department of Chemical and Geological Sciences, University of Modena and Reggio Emilia, Via G. Campi 103, Modena 41125, Italy
[b] Department of Life Sciences, University of Modena and Reggio Emilia, Via G. Campi 103, 41125 Modena, Italy
[c] Clinical and Experimental Medicine PhD Program, University of Modena and Reggio Emilia, Via G. Campi 287, 41125 Modena, Italy
[d] Department of Pharmaceutical Sciences, University of Eastern Piedmont, Largo Donegani 2, 28100 Novara, Italy
[e] PlantaChem srls, via A. Canobio 4/6, 28100 Novara, Italy

ARTICLE INFO

Keywords:
Cannabis sativa L.
Cannabinoids
Design of experiments
Analysis
HPLC

ABSTRACT

The scientific interest in *Cannabis sativa* L. analysis has been rapidly increasing in recent years, especially for what concerns cannabinoids, plant secondary metabolites which are well known for having many biological properties. High-performance liquid chromatography (HPLC) is frequently used for both the qualitative and quantitative analysis of cannabinoids in plant extracts from *C. sativa* and its derived products. Many studies have been focused on the main cannabinoids, such as Δ^9-tetrahydrocannabinolic acid (Δ^9-THCA), cannabidiolic acid (CBDA), cannabigerolic acid (CBGA) and their decarboxylated derivatives, such as Δ^9-tetrahydrocannabinol (Δ^9-THC), cannabidiol (CBD) and cannabigerol (CBG). In addition to the abovementioned compounds, the plant produces other metabolites of the same chemical class, and some of them have shown interesting biological activities.

In the light of this, it is important to have efficient analytical methods for the simultaneous separation of cannabinoids, which is quite complex since they present similar chemical-physical characteristics. The present work is focused on the use of the Design of Experiments technique (DoE) to develop and optimise an HPLC method for the simultaneous separation of 14 cannabinoids. Experimental design optimisation was applied by using a Central Composite Face-Centered design to achieve the best resolution with minimum experimental trials. Five significant variables affecting the chromatographic separation, including ammonium formate concentration, gradient elution, run time and flow rate, were studied. A multivariate strategy, based on Principal Component Analysis (PCA) and Partial Least Squared (PLS) regression, was used to define the best operative conditions. The developed method allowed for the separation of 12 out of 14 cannabinoids. Due to co-elution phenomena, HPLC coupled with a triple quadrupole mass analyser (HPLC-ESI-MS/MS) was applied, monitoring the specific transitions of each compound in the multiple reaction monitoring (MRM) mode. Finally, the optimised method was applied to *C. sativa* extracts having a different cannabinoid profile to demonstrate its efficiency to real samples.

The methodology applied in this study can be useful for the separation of other cannabinoid mixtures, by means of appropriate optimisation of the experimental conditions.

1. Introduction

Cannabis sativa L. is an annual herbaceous plant belonging to the Cannabaceae family. It has been used for recreational purposes over the centuries [1]. Nowadays, some varieties have been approved for medical use in many countries [1]. In addition, non-psychoactive varieties of the plant of *C. sativa* (also known either as industrial hemp or simply hemp) are a well-known source of both fibre and food [1].

C. sativa contains different classes of bioactive compounds, such as cannabinoids, terpenes, flavonoids and so on [2]. The most

* Corresponding authors.
E-mail addresses: caterina.durante@unimore.it (C. Durante), federica.pellati@unimore.it (F. Pellati).

https://doi.org/10.1016/j.jpba.2022.115037
Received 28 July 2022; Received in revised form 6 September 2022; Accepted 7 September 2022
Available online 9 September 2022

representative ones are cannabinoids, a class of terpenophenolic compounds accumulated mainly in female inflorescences [3]. Currently, more than 500 natural compounds have been identified from *C. sativa* and, among them, more than 100 are cannabinoids [4]. In *C. sativa*, cannabinoids are biosynthesized and accumulated as cannabinoic acids, which are subsequently decarboxylated into their neutral forms [4].

Recreational-type varieties of *C. sativa* have a high content of Δ^9-tetrahydrocannabinol (Δ^9-THC) and its biosynthetic precursor Δ^9-tetrahydrocannabinolic acid (Δ^9-THCA). Plants belong to "drug-type" varieties have a standard content of both Δ^9-THC and cannabidiol (CBD), the second being the main non-psychoactive compound present in *C. sativa*. Fibre-type plants of *C. sativa* are characterized by a low content of Δ^9-THC, typically below 0.2% *w/w*, and a high content of non-psychoactive compounds, mainly CBD and cannabigerol (CBG).

In recent years, the use of *C. sativa* for medicinal purposes has rapidly grown. Cannabinoids have been investigated for many potential therapeutic applications in the treatment of inflammation, pain, neurodegenerative disorders, epilepsy, cancer and more [5–10]. In parallel, there has been an increasing demand for accurate analysis of cannabinoids in various products from C. *sativa* and, for this reason, it is highly important to develop reliable and broad-spectrum methods for qualitative and quantitative analysis of these compounds. The two most common instrumental techniques for the analysis of cannabinoids include gas chromatography (GC) and high-performance liquid chromatography (HPLC), the second being the most suitable one for cannabinoid analysis, since it does not involve thermal stress of the injected compounds [11]. Indeed, heat can cause the decarboxylation of the cannabinoic acids into their neutral form [11].

Several HPLC methods have been described in the literature for this purpose [12–14] and most of them have been focused on the separation of the main cannabinoids, including Δ^9-THCA, cannabidiolic acid (CBDA), cannabigerolic acid (CBGA), cannabichromenic acid (CBCA) and their decarboxylated derivatives, such as Δ^9-THC, CBD, cannabigerol (CBG) and cannabichromene (CBC). The crucial aspect of the methods is the separation of some critical pairs of cannabinoids, including CBD/CBG and Δ^9-THC/Δ^8-THC [12,14–18]. Moreover, few papers have included varinic cannabinoids, such as cannabigerovarin (CBGV) and cannabichromevarin (CBCV), while cannabinerol (CBNR), a geometric isomer of CBG, has never been taken into consideration in HPLC methods described up to now in the literature.

In the light of all the above, the aim of this work was to develop an innovative HPLC method, providing a short run time and adequately resolving the analytes of interest, under reversed-phase conditions. For this aim, a Central Composite Face-Centered (CCF) design was applied, and the best experimental conditions were obtained considering the interaction between the significant factors and, therefore, their effects on the response (resolution of the different chromatographic peaks). Principal component analysis (PCA) was performed as an explorative analysis to obtain an overview on the whole data set without forcing any model and to extract relevant information. Partial Least Squared (PLS) regression was used to model the obtained results and the relevant effects were chosen according to their variable influence on projection (VIP) value.

The use of a Design of Experiment (DoE)-based analytical strategy to simultaneously balance chromatographic parameters to ensure optimal separation in HPLC has been already described in several studies [19–21]. However, to the best of our knowledge, this is the first study that aimed at the optimisation of a HPLC method by means of a synergistic use of DoE and PLS approaches for the multi-component analysis of different cannabinoids. Indeed, when more variables must be modelled, it could be more useful to fit a PLS model since, considering the covariance of the investigated matrices, it is able to simultaneously represent the variation of all responses.

2. Material and methods

2.1. Chemical and solvents

Standard solutions of CBDA, CBD, CBGA, CBG, Δ^9-THCA and CBN (1 mg/mL either in acetonitrile or methanol) were purchased from Restek (Milan, Italy). Pure Δ^8-THC, CBC, CBCA, CBDV, CBCV, CBGV and CBNR were provided by PlantaChem with a purity of 98%. A mixture of the 14 cannabinoids at 40 μg/mL was prepared in methanol (MeOH) and used in all the analysis for the method optimisation.

HPLC grade acetonitrile (ACN), MeOH, ethanol (EtOH), formic acid (HCOOH) and ammonium formate were from Sigma-Aldrich (Milan, Italy). Water (H_2O) was purified using a MilliQ Plus185 system from Millipore (Milford, MA, USA).

2.2. Plant samples

Three ethanolic extracts from different *C. sativa* varieties were analysed in this work for demonstration purposes, including one from a recreational-type variety, which was kindly provided by the Toxicology Laboratory of the Forensic Institute of the Department of Biomedical, Metabolic and Neurosciences of the University of Modena and Reggio Emilia, one from a medical-type variety, which was kindly provided by a local pharmacy and prepared under prescription, and one from a fibre-type variety (Bernabeo), belonging to those approved by the European Union (EU). All the extracts were prepared following a previously validated method for the extraction of cannabinoids [17].

2.3. HPLC-UV/DAD analysis

The HPLC-UV/DAD analysis of pure cannabinoids was carried out on a Hewlett-Packard 1100 series system (Palo Alto, CA, USA), equipped with a quaternary pump, a manual injector, a thermostated column compartment and a Diode-Array detector (DAD). An Ascentis Express C_{18} column (150 mm × 3.0 mm I.D., 2.7 μm, Supelco, Bellefonte, PA, USA) at 25 °C was used. The injection volume was 6 μL. A mobile phase composed of 0.1% HCOOH in H_2O with ammonium formate at different concentrations (2, 5 and 8 mM) (A) and 0.1% HCOOH in ACN (B) was selected. The DAD detector was set to acquire UV-Vis spectra in the range 200–500 nm and chromatograms were recorded at 210 nm for the analysis of decarboxylated cannabinoids and 220 nm for the analysis of cannabinoid acids.

The concentration of the ammonium formate, the flow rate of the mobile phase, the duration of the run, and the initial and final concentration of solvent B were different for each of the chromatographic runs and were established by DoE. Each analysis was performed in duplicate.

2.4. HPLC-ESI-MS/MS analysis

The HPLC-ESI-MS/MS analysis was carried out on an Agilent Technologies 1200 series LC system (Santa Clara, CA, USA), equipped with a degasser, a quaternary pump, an autosampler, a thermostated column compartment, and an Agilent Technologies 6410 series (Santa Clara, CA, USA) triple quadrupole mass analyser. The column and the chromatographic parameters applied were the same of those of the HPLC-UV/DAD method.

The MS acquisition was carried out with an electrospray ionisation source (ESI) operating both in the positive and in the negative ion mode. The source parameters for both polarities were set as follows: spray gas pressure (N_2) 35 psi, gas temperature 300 °C, gas flow 10 L/min, electrospray voltage 3.5 kV. The triple-quadrupole mass analyzer was operated in the Multiple Reaction Monitoring (MRM) mode. The transitions of the individual compounds were first optimized. To do this, each standard was directly injected into the ESI source at the concentration of 2 ppm. Optimized MRM transitions for each cannabinoid are shown in Table 1.

Table 1
Optimised MRM parameters for each compound.

Peak number	Compound	t_R (min)	Ione mode	Precursor ion (*m/z*)	Fragmentor (V)	Product ion (*m/z*)	CE (eV)	CAV (V)
1	CBGV	6.4	+	289	100	165	12	1
						123	32	2
						108	50	4
2	CBDV	6.5	+	287	100	231	15	2
						165	18	1
						135	15	2
3	CBDA	8.4	–	357	120	339	15	2
						313	15	2
						245	28	4
4	CBGA	8.9	–	359	130	341	15	2
						315	18	2
						297	22	2
5	CBG	9.4	+	317	100	193	10	4
						123	35	2
6	CBNR	9.4	+	317	100	193	13	4
						137	35	1
						123	35	2
7	CBD	9.9	+	315	110	259	15	2
						193	18	4
						123	32	2
8	CBCV	12.8	+	287	90	231	12	2
						205	15	4
						165	20	1
9	CBN	13.7	+	311	100	293	12	1
						241	18	2
						223	18	8
10	CBCA	14.2	–	357	120	339	15	2
						313	20	2
						243	20	2
11	Δ^9-THC	15.9	+	315	90	259	18	2
						233	18	2
						193	25	4
12	Δ^8-THC	16.3	+	315	100	259	18	6
						207	20	4
						193	25	1
13	CBC	17.9	+	315	100	233	12	2
						193	20	4
						123	35	2
14	Δ^9-THCA	18.3	–	357	120	313	22	1
						245	30	4
						191	35	2

2.5. Experimental design and data analysis

Five main factors, including ammonium formate concentration, initial and final percentage of solvent B, gradient run time and flow rate, affecting the chromatographic separation of the investigated cannabinoids, were optimised by means of a Central Composite Face-Centered design (CCF) [22,23]. The CCF design, composed by a Factorial Design and a Star Design, allowed us to estimate the linear terms, the interactions between variables and the quadratic terms. Twenty-seven experiments were planned, which consisted of 16 corner experiments, 10 axials experiments and one central point.

In this study, all the experiments were twice replicated with the aim to evaluate the significance of the effects of each term, and the different chromatographic runs were randomly performed to ensure a genuine replication [24].

Thirteen different responses, corresponding to resolution values computed for each pair of consecutive (in terms of retention time) investigated cannabinoids, were considered and a PLS algorithm was used to model the results obtained, allowing us to simultaneously represent the variation of all responses [25]. The obtained PLS regression coefficients were considered as an estimate of the effect induced by the corresponding term on the response, and they were also used to point out the level (high or low) of the most influential parameters, to set the optimal chromatographic conditions.

Since the model interpretation based on regression coefficients could be cumbersome with thirteen responses, the variable importance in the projection (VIP) scores [24] was used, and model terms with VIP-values higher than one were considered important in given model.

To further evaluate whether the VIP selected model term had a significant effect on the response, a critical value was computed for each of the selected terms as well, considering an estimation of its error variance and standard errors from replicated runs [25]. In the specific case of performing duplicate runs, it is possible to obtain a pooled estimate of the experimental run variance, using the following equation [25]:

$$s^2 = \frac{\sum d^2}{2g} \quad (1)$$

where d is the difference of duplicate and g the degrees of freedom (in this case equal to the number of experimental runs). To determine which model terms were certainly real, a significant value of Student's t at the 5% was used at g degree of freedom; thus the 95% confidence interval for each effect was given by product of Student's t by the pooled estimate variance computed for each response.

2.6. Software

PCA and PLS regression were carried out by using PLS_Toolbox 8.9.2 software (Eigenvector Research Inc., Manson, WA, USA) for MATLAB®. Designs of experiments were planned with MODDE 9.1 (Umetrics AB, Umeå, Sweden).

3. Results and discussion

3.1. Optimisation of the chromatographic conditions by experimental design

The aim of the present work was to define and optimize the experimental conditions to be used for the HPLC simultaneous separation of 14 cannabinoids (Fig. 1). Δ^9-THC and Δ^8-THC are positional isomers. Δ^9-THC, Δ^8-THC, CBD and CBC have the same molecular weight and they are structural isomers, as well as Δ^9-THCA, CBDA and CBCA. The same is for CBCV and CBDV, which have the same molecular weight and they are structural isomers. Finally, CBG and CBNR are geometric isomers.

In the first part of the work, a DoE approach was employed with three main issues:

1) using a minimum number of experiments to find the optimal experimental conditions;
2) exploring the influence of the main investigated effects and their interactions on the chromatographic separation;
3) generating a mathematical model able to describe the relation among the significant factors and investigated response.

Before performing the DoE approach, an investigative analysis was carried out to evaluate the chromatographic profile of a standard mixture containing all the compounds of interest, by applying the conditions reported in the literature [15,26,27]. The identification of the different analytes was achieved by the comparison of retention time with those obtained by the analysis of pure standards under the same experimental conditions. Early eluting analytes (i.e. having t_R around 1.5–3.5 min) were poorly resolved, while CBN-CBNV-CBCA (t_R 4.5 min) and CBC-Δ^9-THCA (t_R 9.0 min) were totally co-eluted.

Therefore, a two levels CCF design was set up to investigate the following conditions:

1) Addition of ammonium formate to the mobile phase A (AF_{conc}) in a concentration ranging from 2 to 8 mM. Indeed, the presence of ammonium formate in mobile phase A at a different concentration could influence the selectivity factor, increasing the ionic strength and slightly changing the pH as well. Using 0.1% HCOOH and ammonium formate concentrations of 2, 5 and 8 mM, the pH values of mobile phase A were 2.8, 3.1 and 3.3, respectively.
2) Influence of different flow rates, ranging from 0.3 to 0.6 mL/min.
3) The gradient of the elution by simultaneously changing time (from 10 to 20 min) and the initial (Bi%) and final (Bf%) percentage of solvent B, ranging from 70% to 80% and from 90% to 100%, respectively.

For each experiment, the resolution values of CBGV (from CBDV,

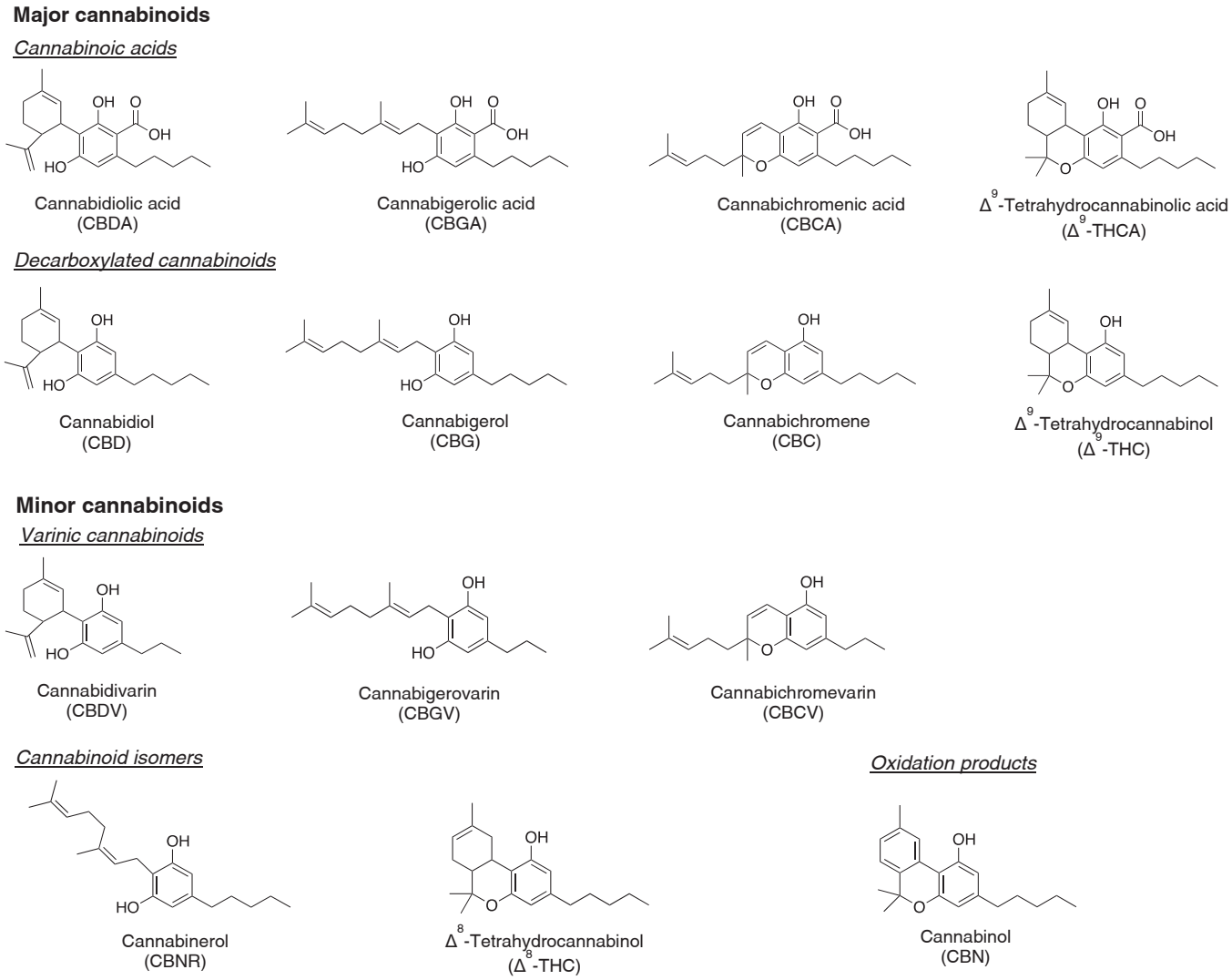

Fig. 1. Chemical structures of investigated cannabinoids.

Table 2
Factors and levels used in the Central Composite Face-Centered (CCF) design.

Factors	Abbreviation	Unit	Lower level (–)	Middle level (0)	Higher level (+)
Ammonium formate concentration	AF_{conc}	mM	2	5	8
Flow rate	Flow	mL/min	0.3	0.45	0.6
Run time	Time	min	10	15	20
Initial percentage of solvent B	Bi%	%	70	75	80
Final percentage of solvent B	Bf%	%	90	95	100

Table 3
Experimental plan used for the optimisation of chromatographic procedure. The symbols (+) and (–) correspond to the high and low level for each factor, respectively.

Run N°	AF_{conc}	flow	time	Bi%	Bf%
N1	–	–	–	–	+
N2	+	–	–	–	–
N3	–	+	–	–	–
N4	+	+	–	–	+
N5	–	–	+	–	–
N6	+	–	+	–	+
N7	–	+	+	–	+
N8	+	+	+	–	–
N9	–	–	–	+	–
N10	+	–	–	+	+
N11	–	+	–	+	+
N12	+	+	–	+	–
N13	–	–	+	+	+
N14	+	–	+	+	–
N15	–	+	+	+	–
N16	+	+	+	+	+
N17	–	0	0	0	0
N18	+	0	0	0	0
N19	0	–	0	0	0
N20	0	+	0	0	0
N21	0	0	–	0	0
N22	0	0	+	0	0
N23	0	0	0	–	0
N24	0	0	0	+	0
N25	0	0	0	0	–
N26	0	0	0	0	+
N27	0	0	0	0	0

R1), CBDV (from CBDA, R2), CBDA (from CBGA, R3), CBGA (from CBG, R4), CBG (from CBNR, R5), CBNR (from CBD, R6), CBD (from CBCV, R7), CBCV (from CBN, R8), CBN (from CBCA, R9), CBCA (from Δ^9-THC, R10), Δ^9-THC (from Δ^8-THC, R11), Δ^8-THC (from CBC, R12) and CBC (from Δ^9-THCA, R13) were measured and the results obtained are shown in Table A (Supplementary Material). The factors investigated, their levels of variation and the experimental plan are shown in Tables 2 and 3, respectively.

To take into account the optimisation of all the investigated resolutions and their different range values, a desirability function [28] was built considering the minimum and the maximum value of each response across all the experiments (Eq. 2), and the geometrical mean of all the computed individual desirabilities D_i (Eq. 3).

$$d_k\left(R_i^{(k)}\right) = 0.8 \quad \times \quad \frac{R_i^{(k)} - \quad L_k}{U_k - \quad L_k} + 0.1 \tag{2}$$

$$D_i = \sqrt[N]{\prod_{k=1}^{N} d_k\left(R_i^{(k)}\right)} \tag{3}$$

where:

N is number of investigated responses;
$d_k\left(R_i^{(k)}\right)$ is the value of desirability function for the kth response in the ith experiment;
$R_i^{(k)}$ is the kth resolution value of the ith experiment;
L_k and U_k are the minimum and maximum values of the kth response across all the experiments, respectively.

Each individual desirability function was normalised to vary between 0.1 and 0.9 [28] and the geometrical means obtained were shown in Fig. 2, which describes them in the same order as that shown in

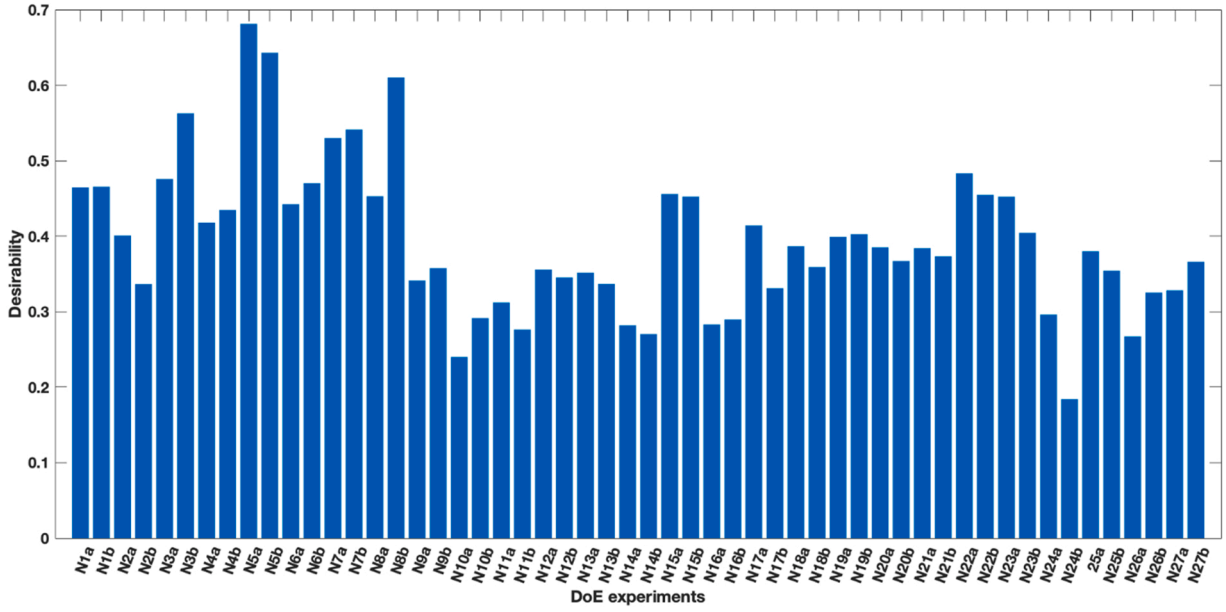

Fig. 2. Optimisation of the experimental conditions based on desirability function. For each run, the two replicates are labelled "a" and "b".

Table A. From a first inspection of the results, it is worth to note that, among the different experiments, the conditions of run 5 (N5a and N5b) could allow to achieve the maximum resolution values, thus suggesting performing the chromatographic separation at a lower value for all the monitored factors, except for the gradient time.

3.2. Principal component analysis (PCA)

To obtain preliminary information on the relationship between the results from the DoE experiments and the goodness of the replicates, a PCA exploratory analysis was carried out on the data set arranged in a bi-dimensional matrix of 54 (DoE experiments) × 12 (resolution values) dimensions. The R5 value (i.e. CBG-CBNR resolution) was not included in data analysis, since it was always zero for all the experimental runs, showing an impossibility to resolve the respective two peaks with the adopted column and experimental conditions. The data matrix was auto-scaled and a 2 PCs model, which explained the 66% of the total variance in fit, was considered.

The bi-plot, i.e. the joint scores (represented by grey triangles) and loading plot (represented by black circles) of the first principal component (PC1) vs. the second one (PC2), is shown in Fig. 3. To facilitate the reading and interpretation of the results, the outcomes of replicated are labelled as "a" and "b", preceded by the same number. From a closer inspection of the figure, it is possible to point out a good reproducibility within the different replicates (same number in the label) along PC1. Notwithstanding a heterogeneous distribution of the samples in the whole PC subspace, some similarities emerge among different chromatographic runs. In particular, N5 and N8 runs (characterised by low values for both initial and final percentage of solvent B) have the highest PC1 scores values. The PC2 components discriminated N9 and N15 runs (both performed at low level of AF_{conc} and Bf% and at high level of Bi%), placed at highest PC2 scores, from N10, N12, N16, characterized by a setting of AF_{conc}, Bi% at their high level.

As far as the resolution values are concerned, differences in their distribution could be observed in the biplot as well. There are many resolution values which seems to be directly correlated and lied at positive values for PC1. Furthermore, these values are opposite with respect to R1 value. On the other hand, PC2 mainly discriminated R9 and R13 (at positive PC2 values) from R10 (at negative PC2 values). Based on these results, it is evident that the investigated chromatographic conditions differently affect cannabinoid resolutions, and it could be further unlikely to find a unique optimal solution that is valid for all the investigated responses. Indeed, some resolutions (R2-R3-R7-R8) seem to be particularly influenced by low values for both initial and final percentage of solvent B, while R13 and R9 resolutions improve at low values of ammonium formate concentration and R10 has an opposite behaviour with respect the previous two responses.

3.3. Partial least squared regression (PLS)

Before starting the PLS analysis, the X matrix was arranged to consider the influence of five linear, ten two-factors interaction and five quadratic interaction terms. This was followed by twenty additional columns in the matrix obtaining a data matrix of 54 × 25 dimensions. On the other side, the Y matrix was composed of 12 resolution values obtained for the 54 experimental runs.

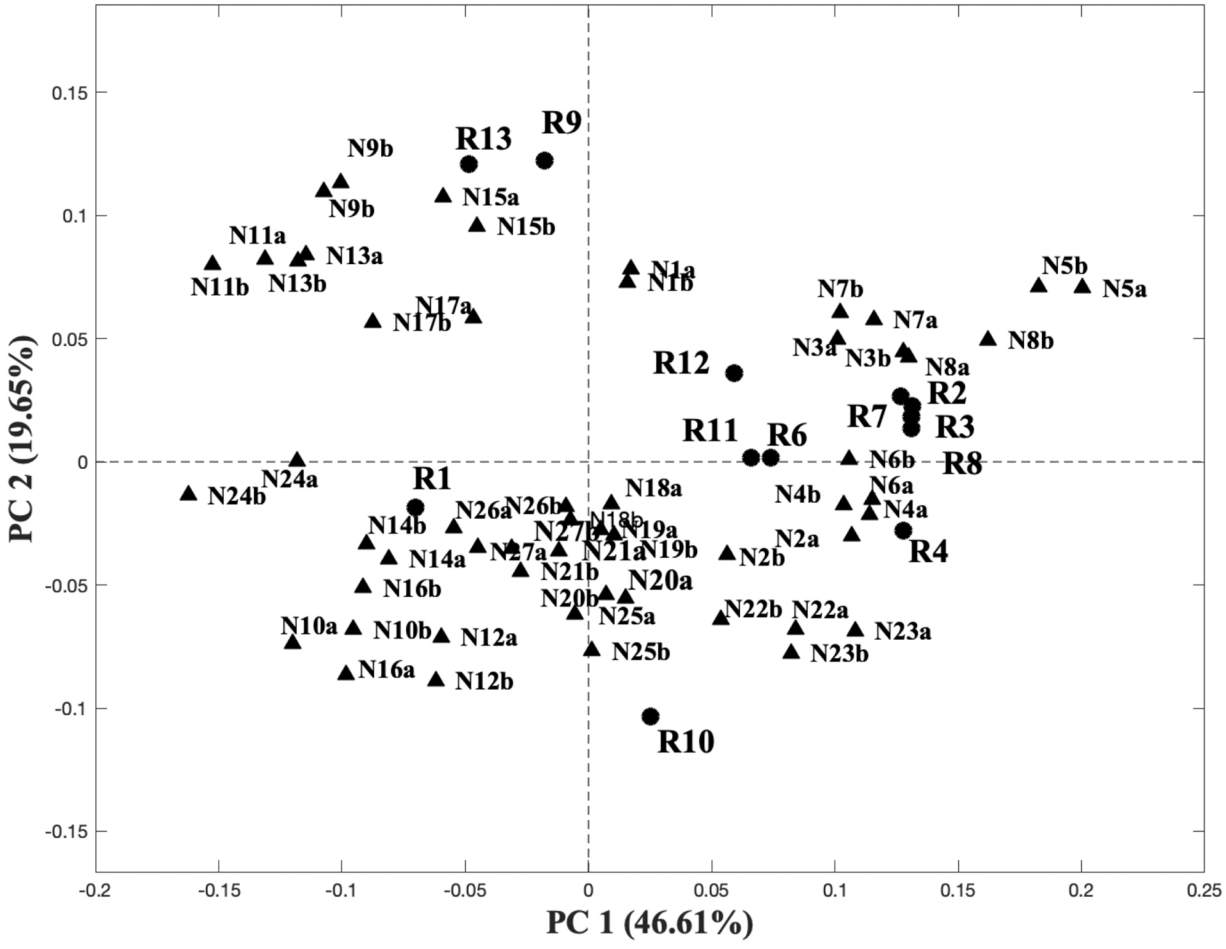

Fig. 3. Biplot of PC1 vs. PC2 of the results obtained from the DoE experimental plan. Each experiment is shown as a grey triangle and it is labelled according to Table 3. Replicates are indicated as 'a' and 'b'. The black circles indicate the variables, i.e. the computed resolutions.

Table 4
PLS Y explained variance in fit (R^2_{fit}) and level of the significative regression coefficients.

		R1 R^2_{fit}: 60%	R2 R^2_{fit}: 88%	R3 R^2_{fit}: 95%	R4 R^2_{fit}: 95%	R6 R^2_{fit}: 42%	R7 R^2_{fit}: 86%	R8 R^2_{fit}: 93%	R9 R^2_{fit}: 85%	R10 R^2_{fit}: 56%	R11 R^2_{fit}: 71%	R12 R^2_{fit}: 82%	R13 R^2_{fit}: 92%
b1	AF_{conc}	n.s.	n.s.	n.s.	(+)	n.s.	n.s.	n.s.	(–)	n.s.	n.s.	(–)	(–)
b2	flow	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	(–)	n.s.	n.s.
b3	time	(+)	n.s.	n.s.	n.s.	(+)	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
b4	%Bi	(+)	(–)	(–)	(–)	(–)	(–)	(–)	(+)	n.s.	(–)	n.s.	(+)
b5	%Bf	(–)	n.s.	n.s.	n.s.	(–)	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
b8	AF×Bi	(+)	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	(–)	(–)	n.s.	n.s.
b9	AF×Bf	n.s.	n.s.	n.s.	n.s.	(–)	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

n.s.: not-significant term; (–): lower level; (+): higher level.

A five latent variables PLS model was chosen (total Y explained variance, R^2_Y: 78%), according to the lower values of root mean square errors in cross-validation (CV venetian bling procedure with six number of splits). The explained variance related to each individual factor in fit and the most influential parameters, selected considering as their VIP values (Fig. A1, Supplementary Material) as the standard errors of the selected effects from replicated runs, are summarized in Table 4. The symbols (+) and (−) correspond to the high and low level for each factor, respectively, and they are individuated considering the sign of the respective regression coefficients (Fig. A2, Supplementary Material). A positive coefficient sign indicates that the best condition to obtain the maximum response is achieved when a high value of this factor is set. Besides, a term with a negative coefficient provides the maximum of the analytical response when its low value is considered. As regards the explained variance related to each individual response, R1, R6 and R10 are poorly modelled with a value of explained variance equal or lower than 60%, while in the other cases satisfactory results are obtained with R^2_Y higher than 70%. By looking at Table 4, the predominant effect could

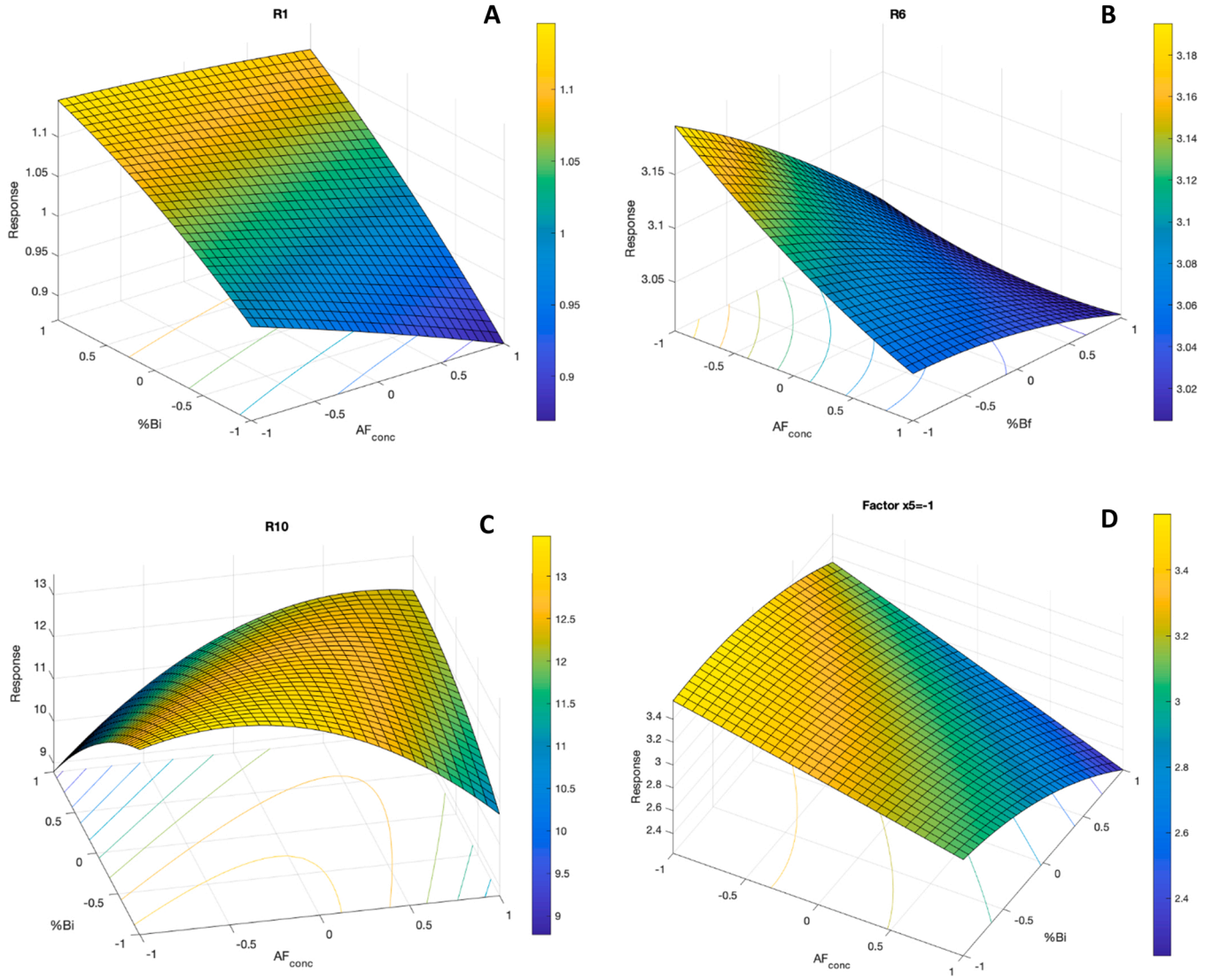

Fig. 4. Surface plots showing the effects of initial percentage of solvent B (%Bi) and ammonium formate concentration (AF_{conc}) on R1 (A), R10 (B) and R11 (C) values and the effects of the final percentage of solvent B (%Bf) and AF_{conc} on R6 (D) values.

be attributed to the initial percentage of solvent B (Bi%), since almost all the resolution values seem to be inversely affected by its value. To some extent, working at a low level of ammonium formate concentration (AF_{conc}), flow and final percentage of solvent B (Bf%) appears to increase some resolution values. However, the effects of Bi%, Bf% and AF_{conc} could not be separately interpreted, because of the importance of their interactions in some cases. Indeed, important issues can be deduced from the interactive terms (AF×Bi and AF×Bf) for R11 and R6 resolutions, where the negative sign suggests that the three factors behave in an opposite manner, i.e. to increase this response the initial and final percentages of solvent B must be used at low level, while keeping ammonium formate concentration at a high level.

The positive sign of the interactive terms AF×Bi in the case of R1 resolution reveals that the two factors behave in a positive manner, i.e. the ammonium formate concentration should be increased to increase the response, while keeping the final percentage of solvent B at a high level.

However, in the case of multivariate models, it is of utmost importance to consider the surface plots when interactions are significant, since the behaviour of two factors could be influenced by the optimal conditions of the same factors involved in other interactions and/or in prediction of other responses. Therefore, the surface plots showing the effects of ammonium formate concentration and initial and final percentage of solvent B on R1, R6, R10 and R11 resolutions are shown in Fig. 4A-D. In all cases, the maximum of resolution is achieved by applying all these factors at their low values.

The gradient run time plays an important role in the elution of cannabinoids represented by R1 and R6 responses: in particular, a high time of gradient is directly correlated to both resolution values. Although other terms were selected according to their VIP values higher than one, they are not included in the discussion of the results because not statistically significant (the values of their respective effects resulted to be lower than the critical values computed considering the standard errors on replicates).

Finally, the critical resolution values (R1, R6 and R10), poorly explained by the previous model, were singularly modelled by means of a PLS model and the results, in terms of number of selected latent variables (LVs), explained variance and the most influent parameters are shown in Table 5. By looking at these results, it is possible to highlight an improvement in the explained variance for all the three models, confirming the influence of the same parameters observed in the previous 'global' model for R1 and R6 responses. As far as the R10 value is concerned, the effects of ammonium formate concentration and of the interaction AF×Bi are predominant and positive. Therefore, the surface plot, which shows the effects of ammonium formate concentration and initial percentage of solvent B, was considered (Fig. A3, Supplementary Material) and, as obtained in the previous model, the maximum of the resolution is achieved by assuming both the factors at their low values.

3.4. Optimised HPLC conditions

After the evaluation of PCA results and PLS coefficients, it was decided to consider the best compromise for the selection of the optimised experimental parameters, according to both the significance of the effects and the goodness of the fitted model. Therefore, to obtain the higher separation of the investigated cannabinoids, the following conditions were applied: low level for ammonium formate concentration (2 mM), low level for the flow rate (0.3 mL/min), high level for time (20 min) and low level for both initial and final percentage of solvent B (70% and 90%, respectively).

The optimised parameters were then applied to the HPLC-UV/DAD analysis of a mixture of the 14 cannabinoids considered in the study. Fig. 5A shows the HPLC-UV/DAD chromatogram of the standard mixture obtained with the optimized experimental conditions. As it can be seen from the chromatogram, the analytical method developed allowed us to separate 12 out of the 14 compounds of interest. Indeed, as previously observed, the resolution R5 was always zero for all the experimental runs. CBG and CBNR are indeed geometric isomers; therefore, their separation on a C_{18} chromatographic column represents a critical issue.

To discriminate CBG/CBNR, it was thus necessary to change the detector type and to use a triple quadrupole mass analyser, which was operated in the MRM mode, both in the positive and negative ion mode, depending on the chemical structure of the compound. The selected MRM transitions, together with the main HPLC-ESI-MS/MS parameters, are shown in Table 1. Two or three transitions were monitored for each precursor ion and, for each transition, the collision energy (CE) and the cell accelerator voltage (CAV) were optimised.

Fig. 5B shows a representative Total Ion Current (TIC) chromatogram of the cannabinoids analysed in this work under the optimised conditions. The chromatographic profile appears the same of that of the HPLC-UV/DAD analysis. Thanks to the optimisation of the MRM transitions for each analyte, it was possible to discriminate between CBG and CBNR, and CBGV and CBDV, which partially co-eluted. As shown in Table 1, the product ions at 193 and 123 *m/z* were monitored for CBG. These ions are also characteristic for CBNR, which, however, differs from CBG for another peculiar product ion at 137 *m/z*. By monitoring these ions with their optimised parameters, it was possible to identify CBNR, thanks its product ion at 137 *m/z* and to distinguish it from CBG. Fig. 6 shows the representative product ions monitored for CBGV (*m/z* 123), CBDV (*m/z* 135), CBG (*m/z* 193) and CBNR (*m/z* 137).

Moreover, the HPLC-ESI-MS/MS method developed was applied to real samples of *C. sativa* extracts having a different phytochemical profile to demonstrate its suitability for the analysis of complex matrices. Fig. 7 shows the TIC chromatograms of the extracts analysed under the optimised conditions. The extract from the recreational-type variety (Fig. 7A) was mainly composed of Δ^9-THCA and Δ^9-THC. The extract

Table 5
Y-PLS explained variance in fit (R^2_{fit}) and level of the significative regression coefficients.

		R1 R^2_{fit}: 75% 2 LVs	R6 R^2_{fit}: 58% 2 LVs	R10 R^2_{fit}: 67% 2 LVs
b1	AF_{conc}	n.s.	n.s.	(+)
b2	flow	n.s.	n.s.	n.s.
b3	time	n.s	(+)	n.s.
b4	%Bi	(+)	(–)	n.s.
b5	%Bf	n.s	(–)	n.s.
b8	AF×Bi	n.s	n.s	(+)

n.s.: not-significant term; (–): lower level; (+): higher level.

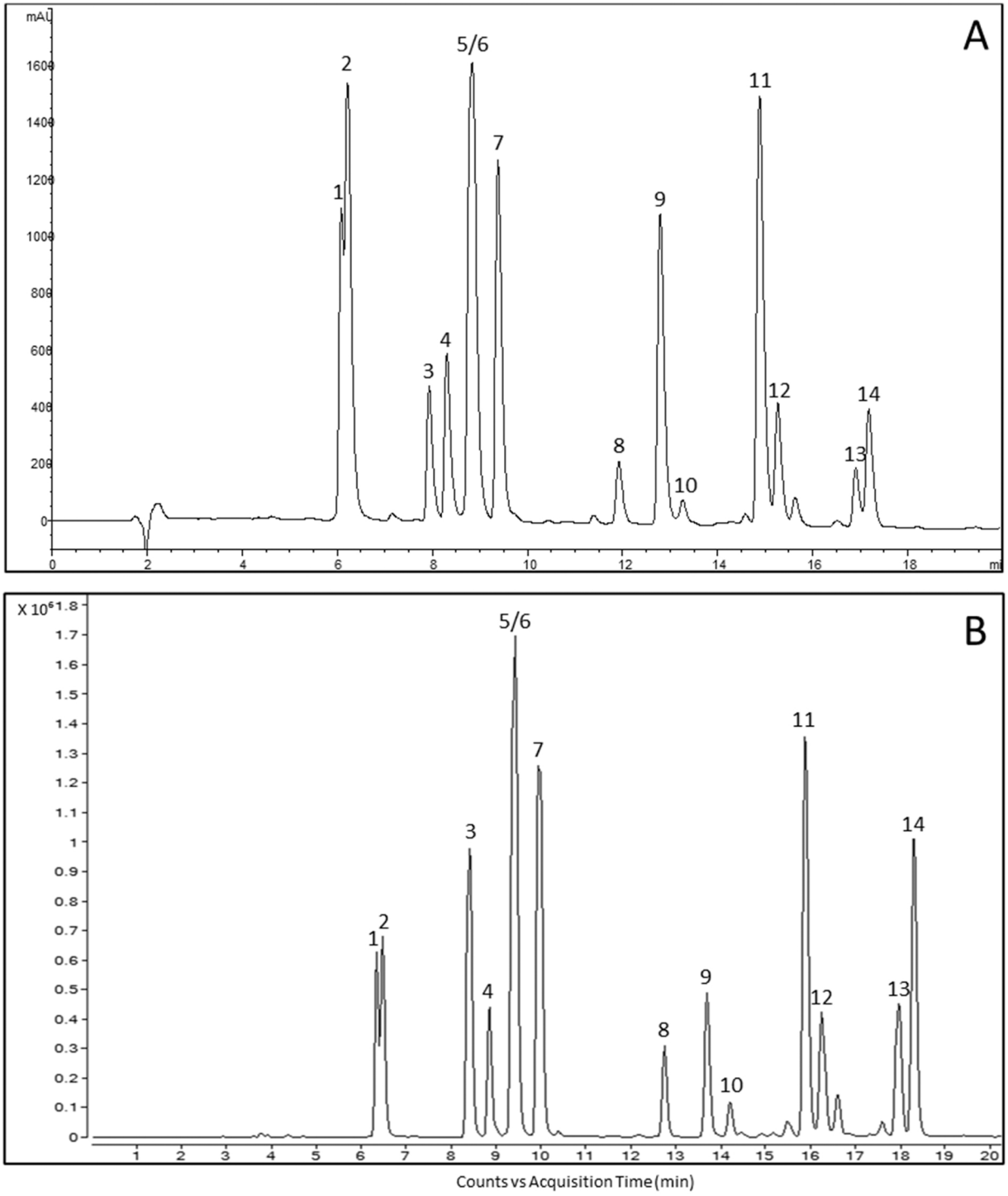

Fig. 5. Representative HPLC-UV/DAD chromatogram of the cannabinoid mixture recorded at 210 nm, obtained under the optimised conditions (A). Representative Total Ion Current (TIC) chromatogram from the HPLC-ESI-MS/MS analysis of the cannabinoid mixture (B). For peak identification, see Table 1.

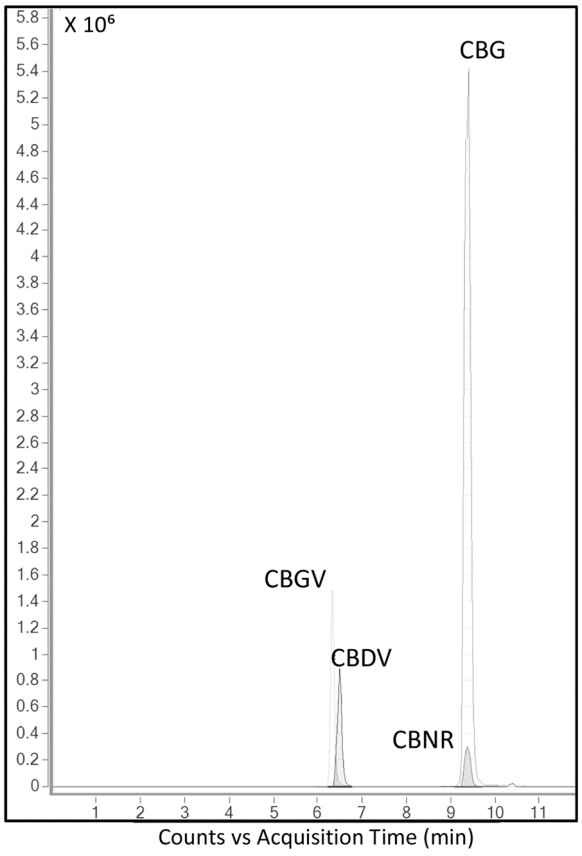

Fig. 6. Multiple Reaction Monitoring (MRM) chromatogram of the cannabinoid mixture, focused on the retention window of CBGV, CBDV, CBG and CBNR, with highlighted the ions at *m/z* 193 for CBG, *m/z* 137 for CBNR, *m/z* 135 for CBDV and *m/z* 123 for CBGV.

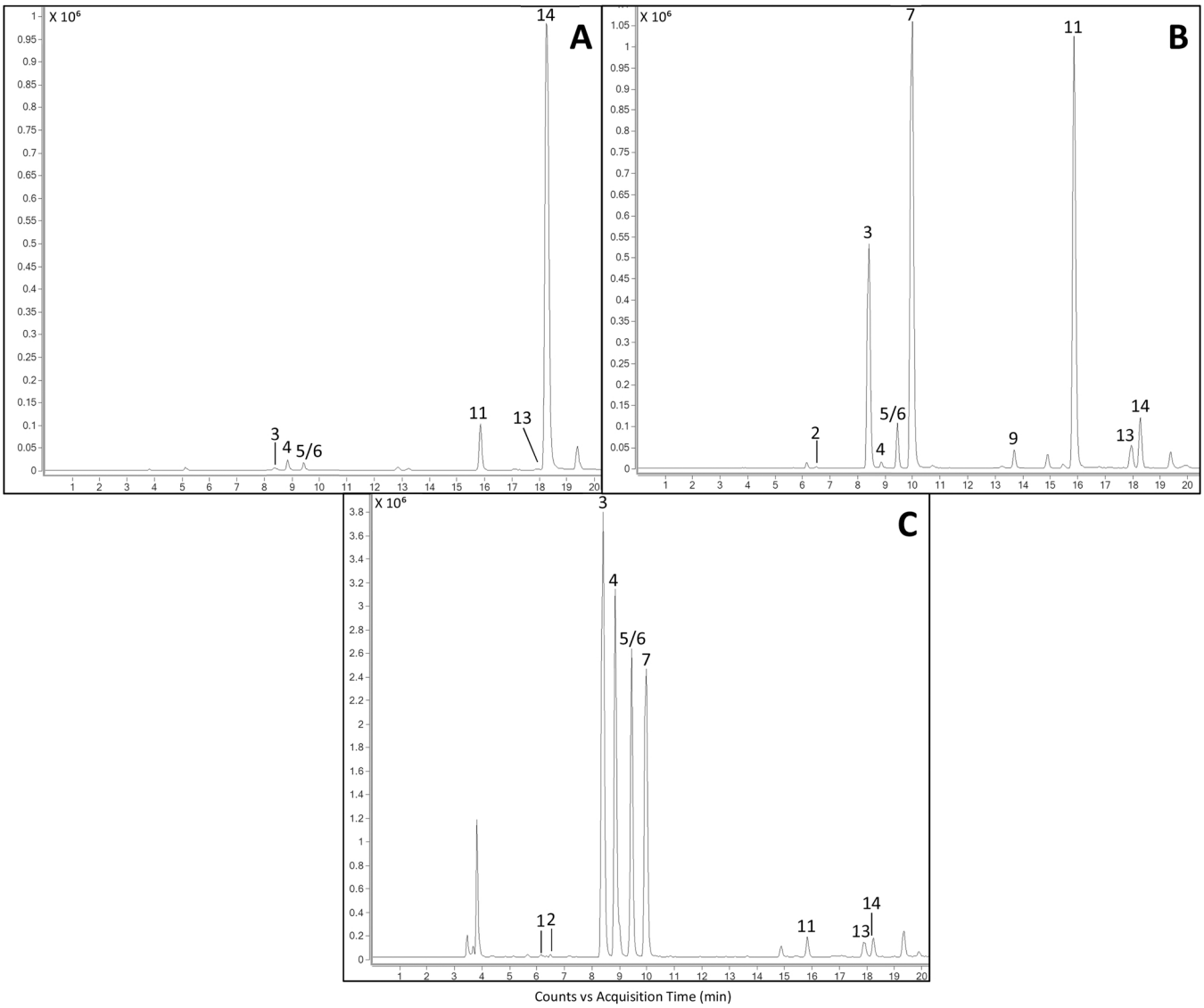

Fig. 7. Representative TIC chromatograms from the HPLC-ESI-MS/MS analysis of recreational-type (A), medical-type (B) and fibre-type (C) *C. sativa* extracts. For peak identification, see Table 1.

from the medical-type *C. sativa* variety (Fig. 7B) was rich in Δ^9-THCA, Δ^9-THC, CBDA and CBD. CBN, which is a metabolite derived from Δ^9-THC, was also detected and it is usually found in cannabis samples that have been exposed to conditions of light and temperature [17]. Concerning the extract from the fibre-type *C. sativa* variety (Fig. 7C), the most representative cannabinoids were CBDA, CBD, CBGA and CBG. The peaks corresponding to CBGV and CBDV had very low intensity as these compounds are produced by the plant in small amounts, being minor cannabinoids.

In general, the cannabinoid profiles of the different extracts from *C. sativa* varieties are in good agreement with the literature [18], underlying the reliability of the developed method in the analysis of these metabolites in plant extracts.

4. Conclusions

The aim of the present work was the optimisation of a new HPLC method for the simultaneous separation of fourteen cannabinoids, using the DoE technique. The CCF experimental design was applied to design the experiments to carry out and to evaluate the interactive effects of the five studied parameters.

Data analysis was initially performed by PCA as an explorative analysis and, subsequentially, by PLS to model the obtained results. VIP scores were used for the selection of the most important terms. However, for the sake of clarity, it is important to highlight that, by considering the complexity of the nature of the analytes investigated, the highest response was not the optimal one, but the best compromise among the different experimental conditions. By applying the analytical strategy described in this work, it was possible to optimise the separation of all investigated cannabinoids except for CBG and CBNR, due to their geometric isomerism. The co-elution issue of these analytes was solved by using a triple quadrupole mass analyser and monitoring the specific MRM transitions for each compound.

The developed method was used for the qualitative profiling of cannabinoids in extracts obtained from different *C. sativa* varieties, thus providing to be a valid tool for a subsequent application in multiple fields of cannabinoid analysis.

Finally, the proposed analytical methodology can be surely useful for the separation of a larger number of cannabinoids, since it allows for the knowledge of the influence of experimental conditions and their interaction on the chromatographic retention of the studied analytes, improving their resolution as well as supporting their identification and

quantification.

CRediT authorship contribution statement

Caterina Durante: Software, Data curation, Formal analysis, Conceptualization, Methodology, Visualization, Writing – review & editing. **Lisa Anceschi:** Investigation, Methodology, Formal analysis, Writing – original draft. **Virginia Brighenti:** Formal analysis, Investigation, Methodology, Writing – original draft. **Clarissa Caroli:** Investigation. **Cindy Afezolli:** Investigation. **Andrea Marchetti:** Methodology. **Marina Cocchi:** Methodology, Software. **Stefano Salamone:** Investigation, Methodology. **Federica Pollastro:** Methodology, Resources, Writing – review & editing. **Federica Pellati:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Visualization, Writing – review & editing.

Declaration of Competing Interest

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

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jpba.2022.115037.

References

[1] S.A. Bonini, M. Premoli, S. Tambaro, A. Kumar, G. Maccarinelli, M. Memo, A. Mastinu, *Cannabis sativa*: a comprehensive ethnopharmacological review of a medicinal plant with a long history, J. Ethnopharmacol. 227 (2018) 300–315, https://doi.org/10.1016/j.jep.2018.09.004.

[2] C.M. Andre, J.F. Hausman, G. Guerriero, *Cannabis sativa*: the plant of the thousand and one molecules, Article 19, Front. Plant. Sci. 7 (2016), https://doi.org/10.3389/fpls.2016.00019.

[3] C.A.S. Tanney, R. Backer, A. Geitmann, D.L. Smith DL, Cannabis glandular trichomes: a cellular metabolite factory, Article 721986, Front. Plant Sci. 12 (2021), https://doi.org/10.3389/fpls.2021.721986.

[4] L.O. Hanuš, S.M. Meyer, E. Muñoz, O. Taglialatela-Scafati, G. Appendino, Phytocannabinoids: a unified critical inventory, Nat. Prod. Rep. 33 (2016) 1347–1448, https://doi.org/10.1039/c6np00074f.

[5] L. Corsi, F. Pellati, V. Brighenti, N. Plessi, S. Benvenuti, Chemical composition and in vitro neuroprotective activity of fibre-type *Cannabis sativa* L. (Hemp), Curr. Bioact. Compd. 15 (2018) 201–210, https://doi.org/10.2174/1573407214666180809124952.

[6] A.M. Costa, L. Senn, L. Anceschi, V. Brighenti, F. Pellati, G. Biagini, Antiseizure effects of fully characterized non-psychoactive *Cannabis sativa* L. extracts in the repeated 6-Hz corneal stimulation test, Article 1259, Pharmaceuticals 14 (2021), https://doi.org/10.3390/ph14121259.

[7] L. Anceschi, A. Codeluppi, V. Brighenti, R. Tassinari, V. Taglioli, L. Marchetti, L. Roncati, A. Alessandrini, L. Corsi, F. Pellati, Chemical characterization of non-psychoactive *Cannabis sativa* L. extracts, in vitro antiproliferative activity and induction of apoptosis in chronic myelogenous leukaemia cancer cells, Phyther. Res. 36 (2022) 914–927, https://doi.org/10.1002/ptr.7357.

[8] W. Häuser, D.P. Finn, E. Kalso, N. Krcevski-Skvarc, H.-G. Kress, B. Morlion, S. Perrot, M. Schäfer, C. Wells, S. Brill, European Pain Federation (EFIC) position paper on appropriate use of cannabis-based medicines and medical cannabis for chronic pain management, Er. J. Pain 22 (2018) 1547–1564, https://doi.org/10.1002/ejp.1297.

[9] G. Comi, A. Solari, L. Leocani, D. Centonze, S. Ostero-Romero, Italian consensus on treatment of spasticity in multiple sclerosis, Eur. J. Neurol. 27 (2020) 445–453, https://doi.org/10.1111/ene.14110.

[10] G. Watt, T. Karl, *In vivo* evidence for therapeutic properties of cannabidiol (CBD) for Alzheimer's disease, Article 20, Front. Pharmacol. 8 (2017), https://doi.org/10.3389/fphar.2017.00020.

[11] C. Citti, D. Braghiroli, M.A. Vandelli, G. Cannazza, Pharmaceutical and biomedical analysis of cannabinoids: a critical review, J. Pharm. Biomed. Anal. 147 (2018) 565–579, https://doi.org/10.1016/j.jpba.2017.06.003.

[12] D. Calzolari, G. Magagnini, L. Lucini, G. Grassi, G. Appendino, S. Amaducci, High added-value compounds from Cannabis threshing residues, Ind. Crops Prod. 108 (2017) 558–563, https://doi.org/10.1016/j.indcrop.2017.06.063.

[13] R. Pacifici, E. Marchei, F. Salvatore, L. Guandalini, F.P. Busardò, S. Pichini, Evaluation of cannabinoids concentration and stability in standardized preparations of cannabis tea and cannabis oil by ultra-high performance liquid chromatography tandem mass spectrometry, Clin. Chem. Lab Med 55 (2017) 1555–1563, https://doi.org/10.1515/cclm-2016-1060.

[14] M. Ternelli, V. Brighenti, L. Anceschi, M. Poto, D. Bertelli, M. Licata, F. Pellati, Innovative methods for the preparation of medical Cannabis oils with a high content of both cannabinoids and terpenes, Article 113296, J. Pharm. Biomed. Anal. 186 (2020), https://doi.org/10.1016/j.jpba.2020.113296.

[15] M. Mandrioli, M. Tura, S. Scotti, T.G. Toschi, Fast detection of 10 cannabinoids by RP-HPLC-UV method in *Cannabis sativa* L, Article 2113, Molecules 24 (2019), https://doi.org/10.3390/molecules24112113.

[16] L. Song, S. Carlson, G. Valenzuela, M. Chao, S.B. Pathipaka, Development of a validated method for rapid quantification of up to sixteen cannabinoids using ultra-high-performance liquid chromatography diode-array detector with optional electrospray ionization time-of-flight mass spectrometry detection, Article 462953, J. Chromatogr. A. 1670 (2022), https://doi.org/10.1016/j.chroma.2022.462953.

[17] V. Brighenti, F. Pellati, M. Steinbach, D. Maran, S. Benvenuti, Development of a new extraction technique and HPLC method for the analysis of non-psychoactive cannabinoids in fibre-type *Cannabis sativa* L. (hemp), J. Pharm. Biomed. Anal. 143 (2017) 228–236, https://doi.org/10.1016/j.jpba.2017.05.049.

[18] V. Brighenti, L. Marchetti, L. Anceschi, M. Protti, P. Verri, F. Pollastro, L. Mercolini, D. Bertelli, C. Zanardi, F. Pellati, Separation and non-separation methods for the analysis of cannabinoids in *Cannabis sativa* L, Article 114346, J. Pharm. Biomed. Anal. 206 (2021), https://doi.org/10.1016/j.jpba.2021.114346.

[19] B.A. Moussa, M.A. Mahrouse, M.G. Fawzy, Application of experimental design in HPLC method optimization and robustness for the simultaneous determination of canagliflozin, empagliflozin, linagliptin, and metformin in tablet, Article e5155, Biomed. Chromatogr. 35 (2021), https://doi.org/10.1002/bmc.5155.

[20] P.K. Sahu, N.R. Ramisetti, T. Cecchi, S. Swain, C.S. Patro, J. Panda, An overview of experimental designs in HPLC method development and validation, J. Pharm. Biomed. Anal. 147 (2018) 590–611, https://doi.org/10.1016/j.jpba.2017.05.006.

[21] M.R. de Souza, M. Koetz, P.R. Limberger, A.T. Henriques, DoE-assisted development and validation of a stabilityindicating HPLC-DAD method for simultaneous determination of five cannabinoids in *Cannabis sativa* L. based on analytical quality by design (AQbD) concept (in press), Phytochem. Anal. (2022), https://doi.org/10.1002/pca.3154.

[22] M. Casale, C. Malegori, P. Oliveri, E. Liberto, P. Rubiolo, C. Bicchi, C. Cordero, Chapter 10: chemometrics: basic principles and applications, Food Chem. Funct. Anal. (2020) 403–451, https://doi.org/10.1039/9781788015752-00403.

[23] R. Leardi, Experimental design in chemistry: a tutorial, Anal. Chim. Acta 652 (2009) 161–172, https://doi.org/10.1016/j.aca.2009.06.015.

[24] G.E.P. Box, J.S. Hunter, W.G. Hunter (Wiley series in probability and statistics). Statistics for Experimenters. Design, innovation and discovery, 2nd edition.,, Wiley, 2005.

[25] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2001) 109–130, https://doi.org/10.1016/S0169-7439(01)00155-1.

[26] E.M. Mudge, P.N. Brown, Determination of cannabinoids in *Cannabis sativa* dried flowers and oils by LC-UV: Single-laboratory validation, first action 2018.10, 4890-493, J. AOAC Int. 103 (2021), https://doi.org/10.5740/jaoacint.19-0197.

[27] P. Galettis, M. Williams, R. Gordon, J.H. Martin, A simple isocratic HPLC method for the quantitation of 17 cannabinoids, Aust. J. Chem. 74 (2021) 453–462, https://doi.org/10.1071/CH20380.

[28] V. Giannetti, M. Boccacci Mariani, F. Marini, A. Biancolillo, Effects of thermal treatments on durum wheat pasta flavour during production process: a modelling approach to provide added-value to pasta dried at low temperatures, Article 121955, Talanta 225 (2021), https://doi.org/10.1016/j.talanta.2020.121955.