Abstract

Phytoestrogens constitute an attractive research topic due to their estrogenic profile and their biological involvement in woman’s health. Therefore, numerous studies are currently performed in natural products chemistry area aiming at the discovery of novel phytoestrogens. The main classes of phytoestrogens are flavonoids (flavonols, flavones), isoflavonoids (isoflavones, coumestans), lignans, stilbenoids as well as miscellaneous chemical groups abundant in several edible and/or medicinal plants, belonging mostly to the Leguminosae family. As for other bioactives, the detection of new structures and more potent plant-derived phytoestrogens typically follows the general approaches currently available in the natural product discovery process. Plant-based approaches selected from traditional medicine knowledge and bioguided concepts are routinely employed. However, these approaches are associated with serious disadvantages such as time-consuming, repeated, and labor intensive processes as well as lack of specificity and reproducibility. In recent years, the natural products chemistry became more technology-driven, and several different strategies have been developed. Structure-oriented procedures and miniaturized approaches employing advanced hyphenated analytical platforms have recently emerged. They facilitate significantly not only the discovery of novel phytoestrogens but also the dereplication procedure leading to the anticipation of major drawbacks in natural products discovery. In this review, apart from the traditional concepts followed in phytochemistry for the discovery of novel biologically active compounds, recent applications in the field of extraction, analysis, fractionation, and identification of phytoestrogens will be discussed. Moreover, specific methodologies combining identification of actives and biological evaluation in parallel, such as liquid chromatography-biochemical detection, frontal affinity chromatography-mass spectrometry and pulsed ultrafiltration-MS will also be presented. Finally, miniaturized methods (microchip and biosensor) will be also discussed.

With the current review, we attempt to give a wide and holistic overview of the different approaches which could be employed in the discovery of new phytoestrogens. On the other hand, we anticipate to attract more scientists to the area of phytoestrogens and to indicate the need of multidisciplinary concepts.

Abbreviations

ACN: acetonitrile
ASE: accelerated solvent extraction
BCD: biochemical detection
BuOH: butanol
CCC: countercurrent chromatography
Co(II)-PAB: cobalt(II)-coated paramagnetic affinity beads
CPC: centrifugal partition chromatography
DAD: diode array detector
DPHP: 2,2-diphenyl-1-picrylhydrazyl
EA: ethyl acetate
ER: estrogen receptor
ERE: estrogen response elements
EtOH: ethanol
FAC: frontal affinity chromatography
FRET: fluorescence resonance energy transfer
FTICR: Fourier transform ion cyclotron resonance
GC-MS: gas chromatography-MS
Hept: heptane
Hex: hexane
HPLC: high-performance liquid chromatography
HRMS: high-resolution mass spectrometry
HSCCC: high-speed CCC

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.
Phytoestrogens are nonsteroidal natural compounds, usually plant constituents or their metabolites, which induce biological responses in mammals mimicking or modulating the action of endogenous estrogens (i.e., 17β-estradiol, estriol, estrone). Semi-synthetic products and synthetic analogues of natural phytoestrogens are also sporadically included [1, 2]. The role of phytoestrogens and their possible protective effect in hormone-related disorders and diseases such as postmenopausal syndrome, osteoporosis, breast and prostate cancer, and cardiovascular complications has been extensively investigated [3–6].

The biological action of phytoestrogens is mainly implemented through their ability to bind to the ER which belongs to the nuclear receptors superfamily. To date, two main isoforms, ERα and ERβ, encoded by two different genes, have been described [7]. The primary structure of ER is characterized by three main functional domains: the area responsible for the binding with the ligand (hormone, phytoestrogens), the region responsible for the interaction with DNA, and the region regulating transcription [8, 9]. The two isoforms are differently distributed to organs and tissues and present a different affinity to ligands.

Phytoestrogens can act either as estrogen agonists or antagonists since they may trigger (i) an estrogenic response, thereby binding and activating the ER, which ultimately induces an estrogenic effect or (ii) an antiestrogenic response, blocking or altering the receptor activation by endogenous estrogens and thereby inhibiting an estrogenic effect. Furthermore, phytoestrogens have also been classified as SERM. Thus, phytoestrogens can act as estrogen agonists and antagonists at the same time usually in an organ-dependent manner resulting in (iii) mixed agonistic/antagonistic properties (selectivity) depending on the receptor content of specific tissues (ERα and ERβ populations) and the concentration of the endogenous estrogens. Finally, phytoestrogens may also have no affinity to ER (iv); in this case they act through alternative mechanisms such as inhibition or activation of specific enzymes implied in an estrogenic response [2, 10] or altering the total amount of free estrogens in the organism. For instance, it has been reported that phytoestrogens interfere with the synthesis of steroid hormones by inhibiting specific enzymes such as aromatase (responsible for the conversion of androgens to estrogens) [11, 12] and 17β-hydroxysteroid oxidoreductase (responsible for the conversion of estrone to estradiol) [13]. Furthermore, it has been suggested that phytoestrogens are able to stimulate, in the liver, the synthesis of proteins which are capable to bind estrogens (SHBG), affecting the overall percentage of free estrogens in the blood circulation [14,15] and thereby exhibiting an indirect estrogenic effect.

On the basis of their chemical structure and in respect to biosynthesis patterns, phytoestrogens may be divided in chalcones, flavonoids (flavones, flavonols, flavanones, isoflavonoids), lignans, stilbenoids, and miscellaneous classes. Particular attention should be given to isoflavonoids, the subgroup of flavonoids which includes amongst others the chemical groups of isoflavones, isoflavonanes, pterocarpanes, and coumestanes (Fig. 1) [16, 17].

These phytochemicals belong to the wide class of polyphenols which are ubiquitous in the plant kingdom and characterized by nonsteroidal structures (Fig. 1). Specific structural characteristics resembling those of endogenous estrogens and consequently demonstrating an estrogenic profile are responsible for their affinity to the ER receptor.
In particular, the estrogenicity of a molecule which is implemented through ER is related to specific structural requirements such as the presence of hydroxyl groups able to form hydrogen bonds with key functional amino acids of the binding pocket of the receptor. On the other hand, these hydroxyl groups should satisfy other critical prerequisites such as the O-O distance (10.9 Å to 12.5 Å) simulating the 3-OH and 17β-OH hydroxy groups of 17β-estradiol [18]. In fact, most of the natural phytoestrogens are characterized by the presence of two aromatic rings, from which at least one bears hydroxyl groups, free or methylated [19].

Isoflavones are probably the most comprehensively investigated group of phytoestrogens. Chemically, they are characterized by a 15-carbon (C6–C3–C6) skeleton and differ from flavones in having the β-ring linked to the C-3 rather than the C-2 position of the central heterocycle (ring C) (Fig. 1). To date, daidzin and genistein are the most studied isoflavones showing significant estrogenic potential. In nature, they occur mainly as glycosides (i.e., daidzin and genistin, respectively) or as methoxylated derivatives (i.e., formononetin and biochanin A) which are metabolized after consumption, liberating the active aglycones. Further biotransformation of isoflavones by intestine microflora in most cases leads to equol and O-desmethylangolensin which both exhibit estrogenic activity [1].

Isoflavones are found almost exclusively in Fabaceae, the largest Leguminosae subfamily [20], but they are also occasionally found in some other angiosperm families [21]. Nevertheless, only a few species provide dietary isoflavones to humans. The major source of isoflavones is soybean (Glycine max) but also other legumes such as the common bean, Phaseolus vulgaris [3].

Other important isoflavone-rich sources are red clover (Trifolium pratense) [22], alfalfa (Medicago sativa) [23], chickpeas (Cicer arietinum) [24], licorice (Glycyrrhiza glabra) [25], and Radix pueraria, a Chinese medicinal preparation containing several root extracts from the Pueraria genus [26].

Another subgroup of isoflavonoids are coumestans which are derived biosynthetically from isoflavones. Coumestrol (Fig. 1) is the best known representative of this group exhibiting significant estrogenicity and being considered as one of the most potent phytoestrogens. The main sources of coumestrol are clover, soybeans,spinaches, and sprouts, but the highest content has been detected in alfalfa (Medicago sativa, Fabaceae) [16,27]. Several flavonoids have also been characterized as phytoestrogens, however, with poorer overall potency compared to isoflavonoids. For instance, the flavanone liquiritigenin and its corresponding chalcone, isoliquiritigenin, has been referred to contribute to the significant estrogenic activity of alfalfa [27]. It is worth mentioning a particular group of flavonoids which are actually prenylated derivatives. The flavanone 8-prenylangolensin is considered one of the most active phytoestrogens, while its prenylated derivative, naringenin, exhibits only a modest estrogenic property. Furthermore, the prenylated chalcone isoxanthohumol together with the prenylated flavanone xanthohumol are the major hops’ (Humulus lupulus, Cannabaceae) phytoestrogens also present in beer [28], however exhibiting a weaker estrogenic potential compared to 8-prenylangolensin (Fig. 1). Interestingly, as to what concerns interindividual differences in the intestinal transformation potential [29], isoxanthohumol is transformed from intestinal bacteria to the active 8-prenylangolensin.

Lignans represent a particular group of phytoestrogens since their metabolites produced after consumption exhibit estrogenic activity in contrast to their predecessors which are almost inactive. Chemically, they are compounds consisting of two phenylpropanoid units linked via a carbon–carbon single bond. The most active lignans are enterodiol and enterolactone which originated from their parent compounds arctigenin, secoisolariciresinol, and matairesinol being formed by the enteric microflora of mammals (Fig. 1) [3,30]. Lignans are widely distributed and can be found in various organs of many plant families. Flaxseed (Linum usitatissimum, Linaceae) is well known for its high content in secoisolariciresinol and matairesinol, but such compounds can also be found in many berries [3]. Arctigenin is consumed with food like safflower seeds (Carthamus tinctorius, Asteraceae) [31], sesame seeds (Sesamum indicum, Pedaliaceae) [32], wheat flower (Triticum aestivum, Poaceae) [33], and from the seeds and roots of Arctium lappa (Asteraceae) [34]. Furthermore, stilbenoids which are also plant polyphenols structurally characterized by the presence of a 1,2-diphenylethylene nucleus have been reported as phytoestrogens (Fig. 1) [35]. The main estrogenic stilbene is resveratrol occurring in grapes and their products such as red wine (Vitis vinifera, Vitaceae) and in peanuts (Arachis hypogaea, Fabaceae) [16,36]. Stilbenoids are also distributed in other families such as Cyperaceae, Pinaceae, and Polygonaceae [35,37,38].

Nevertheless, there are also numerous compounds from miscellaneous chemical classes that have been described as phytoestrogens exhibiting important estrogenicity. A characteristic example is the mycotoxin zearalenone (phenolic resorcylic acid lactone) produced by fungi of the Fusarium species and their metabolites (α- and β-zearalenols) which display significant estrogenic activity [39]. Other examples are compounds from the chemical group of 2-arylbenzofurans (ebenfuran I, II, and III) [40], diarylheptanoids [41], coumarines (psoralen and isopsoralene) [42], fatty acids (α-linoleic acid) [43], diterpenoids such as tanshinones (abietane-type diterpenes) [41], and terpenoids (dehydroyabietic acid, Δ9-tetrahydrocannabinol) [44,45]. Also, compounds from the chemical group of lactones such as specific butenolides or from the chemical class of alkaloids (casimiroin, berberine) have also been reported (Fig. 1) [46–48]. Overall, phytoestrogens occupy a central position in natural product research due to their potential health effects on peri/postmenopausal women, but also on men, children, and animals (e.g., sheep), which have been conclusively and extensively reviewed by Cornwell et al. [3].

Currently, many studies in the area of natural products are ongoing and focused on the discovery of novel phytoestrogens. Usually, classical approaches in natural product chemistry are used for the detection, isolation, and further complete identification of phytoestrogens. Nonetheless, new approaches, techniques, and methodologies have been recently introduced enabling the accelerated, specific, more sensitive and cost-effective discovery thereof.

After a brief overview of the classical approaches and their recent improvements, this review will focus on modern advances in the dereplication and discovery of novel phytoestrogens. Our main goal is to briefly present the traditional and current status in the area of phytoestrogen’s discovery from natural sources and the potentials which emerge from the recent technological advances. We aim to practically assist and attract pharmacognosists and natural product chemists to work on secondary metabolites relevant for hormone-related diseases but also challenge other scientists outside the area of phytochemistry to share efforts aiming at more multidisciplinary concepts.
Traditional and Modern Concepts in Natural Products Discovery

Following the traditional work-flow of natural products discovery, phytoestrogens are identified after specific and distinctive experimental steps which generally include extraction of plant material, fractionation, separation, and isolation of the plant constituents, structure elucidation thereof and finally determination of biological properties of the purified compounds (Fig. 2A). However, the complicated nature of natural extracts related mainly to the variability and uniqueness of structural patterns hamper the discovery process. Moreover, the limited quantity of the starting material usually available and therefore of the final identified products hinders further the entire procedure which is significantly slow, laborious, expensive and suffers from reproducibility and sensitivity [49,50]. An important improvement of this approach comprises the so-called “bioguided isolation” which introduces a more rational and targeted concept throughout the classical procedure (Fig. 2B). By the successive biological evaluation of extracts, derived fractions, and finally purified compounds, the entire process is monitored and guided based on the determined activity, orienting therefore more selectively towards the discovery of actives. Even if this approach is much more time- and cost-effective, relatively highly specific, and more focused to the biological target, there are some specific restrictions which limit its use. Competition and synergism phenomena of constituents in extracts and fractions often result in false positive and negative results misleading the whole isolation procedure [51,52].

Recently, two new trends have been developed in order to circumvent these drawbacks aiming at the same time to upgrade, improve, and accelerate the natural product discovery process. As it is generally ascertained in modern science, both are technology driven and based on advanced methodologies using multidisciplinary methods in a high-throughput manner [52,53]. With or without incorporation of the traditional “bioguided” concept and employing state-of-the-art analytical platforms both seems to offer more targeted, reproducible and faster alternatives. The first approach (Fig. 2C) is generally more target-oriented (biology-based) sustaining the traditional steps of bioevaluation, however introducing a small-scale screening and libraries concept (miniatursised approach) in all phases involving most of the time high-throughput well plate formats [54,55]. Furthermore, analytical tools (HPLC-UV, HPLC-DAD, LC-MS) are incorporated mainly at the (micro)fractionation step enabling the correlation of constituents (peaks) with the activity resulting in a more focused isolation of candidate active metabolites [52,56].
On the other hand, the second approach (Fig. 2D) is more structure-oriented (chemistry-based) utilizing to a higher degree the power of analytical and elucidation platforms (LC-HRMS and HRMS/MS, high-field NMR) as well as dereplication strategies which consist in the rapid identification of already known molecules [57,58]. The identification of a secondary metabolite early in the discovery process (e.g., extract) accelerates drastically the entire progression of new actives discovery eliminating the repeated and laborious isolation steps. Employing spectral databases, informatics, intelligent correlation software, and hyphenated analytical apparatuses enables the prioritization of extracts/fractions leading to the active compounds selectively [59,60].

Overall, the above-mentioned work-flows are followed by the majority of phytochemistry research groups either alone or as combinations. However, besides the general discovery concepts, significant progress has been made in the individual steps of extraction, separation, isolation, and identification as well as dereplication. Technological advanced instrumentation, bioactives enrichment capabilities, affinity-based screening formats and devices are increasingly incorporated for accelerating not only the detection and purification of an active secondary metabolite but also for activity-dereplication purposes [59,60]. Thereafter, recent advances in extraction, isolation, analysis, characterization, and identification of phytoestrogens will be discussed as well as new approaches and methodologies for the dereplication of phytoestrogens.

**Recent Advances in the Discovery Work-Flow of Phytoestrogens**

**Extraction**

Traditionally, the strategy to isolate active compounds from plant material is initially to use maceration (shaking, stirring) or Soxhlet extraction using solvents of different polarities, individually or successively. However, such techniques are generally time- and solvent-consuming, suffering from reproducibility. In the last decade, alternative extraction techniques including PLE, SWE, SFE, UAE, and MAE, have emerged as energy saving technologies [61]. PLE, SWE, and SFE are based on the use of compressed fluids as extracting agents [62], whereas UAE and MAE use ultrasonic [63] and microwave energy, respectively [64], as the source of heating of the solvent-matrix solution (Fig. 3). In a general point of view, these modern approaches have been developed for more accurate and reproducible results and are mainly employed in combination with high-throughput analyses like UHPLC-MS/MS. These techniques have also been applied for the extraction of phytoestrogens, while in the literature, most studies concern the extraction of isoflavonoids from soybean and other soy derivatives, thoroughly reviewed by Rostagno et al. [65].

**Pressurized liquid extraction and subcritical water extraction**: PLE uses high pressure under an inert nitrogen atmosphere that allows the application of the extraction solvent at temperatures above their normal boiling point [62]. ASE is one of the most known types of PLE (Fig. 3A). The combination of elevated pressure and temperature provides a more efficient and reproducible extraction process due to a better solubility and mass transfer of targeted analytes and to a lower solvent viscosity. Other advantages of PLE over conventional extraction methods include automation, faster extraction rate, and lower solvent consumption. Zgórka in a recent work investigated PLE for the extraction of isoflavones in hydrolyzed extracts obtained from aerial parts of five *Trifolium* (clover) species [66]. Using a methanol-water solution (75:25, v/v) at 125 °C, PLE proved to be an efficient and accurate approach compared to UAE and conventional solvent extraction methods for the recovery of biochanin A, formononetin, daidzein, and genistein. The same author recently investigated the phytoestrogenic content of *Trifolium incarnatum* and other clover species using the same PLE procedure monitored by HPLC connected to UV and fluorescence detectors [67]. Generally, PLE extraction of isoflavonoids is usually performed with alcohol or hydroalcoholic solutions [65]. Nevertheless, Luthria et al. [68] recommend to systematically evaluate and optimize the extraction solvent composition for the accurate determination of phenolic compounds. Indeed, they studied the impact of cosolvents on the extraction rate of isoflavones from soybeans using a previously optimized solvent composition. Optimum recovery of 12 isoflavones using PLE with dimethyl sulphoxide:ethanol:water (5:70:25, v/v/v) as the solvent mixture at 100 °C was achieved. SWE is a PLE variant where water is heated at a temperature above 100 °C and maintained as liquid under high pressure. Pressurized hot water extraction has been utilized for the extraction of lignans and prenylated flavonoids. Ho et al. [69] investigated the potential of SWE for the extraction of secoisolariciresinol di glucoside from flaxseed. The maximum yield of lignans was obtained at 170 °C with a solvent to solid ratio of 100 mL/g at pH = 9.

Just recently, Gil-Ramírez et al. demonstrated that subcritical water at 150 °C is an alternative for the extraction of prenyflavonoids such as xanthohumol and isoaxanthohumol from hops [70]. Furthermore, the extraction method based on water has been proven to be more selective than similar processes based on subcritical ethanol or sequential extraction with different solvents. However, it is worthwhile to note that using a high temperature during the extraction process can induce degradation of compounds, especially thermolabile ones, and increase intensity of reactions such as hydrolysis and oxidation [71].

**Supercritical fluid extraction**: The supercritical state of organic compounds is obtained when the conditions of pressure and temperature reach values above their critical points. The main supercritical fluid used in extraction processes is carbon dioxide (CO₂, critical conditions = 30.9 °C and 73.8 bar) which is considered inexpensive and environmentally friendly and generally recognized as safe [72]. Although SFE is an interesting alternative technique with a lower solvent consumption, the main limitation of CO₂ is its low ability to dissolve polar molecules due to its high hydrophobicity. This complication can be overcome using coextracting solvents (e.g., methanol, ethanol) to increase the polarity of the CO₂ and, in general, of the supercritical fluid, enhancing its solvating power towards polar compounds.

SFE was recently applied for the extraction of isoflavonoids from soybeans at 55 °C, 100 bar, and with 7.5% of ethanol as the solvent modifier. SFE extract was characterized by a slightly lower isoflavones recovery compared to conventional extraction with organic solvents. However, SFE presents advantages such as a faster extraction time without utilization of organic solvents, a reduced sample handling, and absence of filtration or cleaning step [73]. A study performed by Bajer et al. [74] compared various extraction techniques for isolation of isoflavonoids in four plant species: *Matricaria recutita* (Asteraceae), *Rosmarinus officinalis* (Lamiaceae), *Foeniculum vulgare* (Apiaceae), and *Agrimonia eupatoria* (Rosaceae). Optimization of all techniques was first carried out using soybean flour, and analysis of the obtained extracts us-
ing HPLC-UV. Specifically, the authors compared various extraction techniques including SFE, PLE, UAE, ultrasonic homogenizing, MSPD, Soxhlet extraction, and solid SPE. The results demonstrated that the extraction yields of daidzein and genistein were higher with UAE and Soxhlet extraction, whereas apigenin and biochanin A were better extracted using SFE with 5% methanol as the cosolvent. However, when considering solvent and time consumptions, SFE appears to be the most appropriate method to extract phytoestrogens from the specific plant matrices. 

SFE extraction at 40°C, 150 bar, with 7.5% ethanol led to the isolation of resveratrol from the grape skin of Vitis vinifera in 15 min [75], while Casas et al. investigated the viability of SFE for the recovery of resveratrol from grape seeds, stems, skin, and pomace of a Spanish grape variety [76]. Optimum extraction of resveratrol from grape by-products was found at 400 bars, 35°C, with 5% of ethanol. The authors also mentioned that resveratrol was not extracted with low pressure (i.e., 100 bar) and suggested that utilization of a cosolvent was necessary to improve the extraction yield of resveratrol. Moreover, SFE has been used for the extraction of other polyphenols from grape skin, for example, (+)-catechin, (-)-epicatechin, rutin, quercetin. Similar conclusions regarding the extraction yield when pressure and modifier percentage are increased have been conducted also in other studies [77].

In the specific study, optimized conditions were established using 250 bars, 60°C, and 20% ethanol as the modifier. Finally, Choi et al. [78] proved that CO2 extraction was well adapted for extraction of aglycon lignans like arctigenin from Forsythia koreana (Oleanaceae). Even if pure CO2 extraction was effective, authors suggested to incorporate methanol as a modifier in order to enhance the extraction yield. Maximum yield was obtained at 340 bars, 80°C, and 20% methanol.

**Microwave-assisted extraction**: MAE is increasingly employed in the extraction of natural products because it is a cheap and rapid technique, and the extraction time and solvent consumption are considerably reduced [79]. MAE is based on the use of microwave energy to heat the solvent-matrix solution in order to desorb molecules from the matrix to the solvent (© Fig. 3B). The principle of heating the material by microwaves is due to two phenomena which intervene simultaneously: dipolar rotation and ionic conduction [64]. Du et al. compared MAE, PLE, UAE, and reflux for extraction of several isoflavonoids from the Chinese preparation *Radix puerariae* (Kudzu) [80]. It has been found that the extraction yield of minor compounds was not affected by the extraction method, but for the major phenolic compounds, such as puerarin, mirificin, daidzein, and daidzein, the extraction efficiency was higher with MAE and PLE compared to reflux and UAE. Furthermore, extraction efficiency by MAE was higher than with the other techniques, offering the same time speed and low solvent consumption. MAE was performed at 100°C for 2 min with 65% ethanol as the extraction solvent, an irradiation power of 600 W, and an extraction volume of 17 mL.

Recently, Terigar et al. [81] designed and optimized a continuous microwave extraction system, consisting of microwave units placed in a series for extraction of major isoflavones (genistin, genistein, daidzin, and daidzein) from soy flour. Using this apparatus, the major isoflavones were extracted efficiently in a short time. Nemes and Orsat [82] optimized by response surface methodology a fast MAE method for extracting secoisolariciresinol diglucoside from flaxseed. A particular point of this specific methodology a fast MAE method for extracting secoisolariciresinol diglucoside from flaxseed. A particular point of this specific method is the use of sodium hydroxide diluted in water, which improves the release of the glucoside lignan from the matrix. Finally, MAE utilization was reported by Yu et al. for the rapid extraction of arctigenin in *Saussurea medusa* (Asteraceae) [83]: optimization by an orthogonal experimental design conducted to a maximum content of arctigenin with one extraction cycle at 390 W for 20 min, using methanol as the solvent and a solvent/solid ratio of 50 mL/g.

**Ultrasound-assisted extraction**: UAE is a low-cost extraction method, applicable with any kind of solvent and simple to set up and maintain. Indeed, the UAE extraction may be performed in a very simple way by using an ultrasound bath or via an ultrasound probe combined to an agitator. Extraction of molecules by ultrasound is attributed to the phenomenon of cavitation produced in the solvent by the passage of ultrasonic waves. Cavitation bubbles
are produced and compressed during the application of ultrasound (Fig. 3C). The increase in the pressure and temperature caused by the compression leads to the collapse of the bubbles which results in high-speed jets [65]. In cases of plant matrix, this liquid jet drills the plant cell wall and allows the transfer of the molecule from the matrix to the solvent.

A chemotaxonomic study of the *Trifolium* genus was undertaken to establish chemical profiles and to determine the amounts of the phytoestrogens genistein, daidzein, formononetin, and biochanin A in the aerial parts of thirteen clover species native to Poland [84]. Isolation of active isoflavones was first undergone using SPE combined to UHPLC, and then *Trifolium* extracts were analyzed by HPLC coupled to UV and fluorescence detectors. Moreover, the thirteen clover species were discriminated based on their chromatographic profile and their flower color variability.

An UAE method was also optimized by experimental design approaches for the extraction of puerarin and isoflavones from *Pueraria lobata* (Fabaceae) [85]. The results demonstrated that the optimal extraction yield for puerarin and isoflavones was obtained with ethanol concentrations of 71.35 and 80.00%, extraction times of 49.08 and 55.00 min, and the solvent-to-mass ratio of 21.72 and 12.81, respectively. Additionally, Jiang et al. [86] investigated the aqueous root extracts of two Chinese medicinal herbs, *Pueraria thomsonii* and *P. lobata* (Fabaceae). Their extraction procedure included an UAE for 1 h in water, followed by fingerprinting and semipreparative isolation of compounds of interest using HPLC. The study led to the isolation of 6 isoflavones; among them, daidzein, daidzin, puerarin, and 5-hydroxy-puerarin which were isolated for the first time from *P. thomsonii*.

Finally, UAE seems to be well adapted to the extraction of estrogenic compounds from *Radix pueraria* extracts as it has been demonstrated above and also by Lee and Lin [87] who showed that extraction efficiency of puerarin, daidzin, and daidzein from this root was higher using UAE rather than the ASE method.

**Hyphenated extraction methods:** An emerging trend in the extraction of natural materials seems to be the combination of two extraction techniques like simultaneous UAE and MAE to improve extraction efficiency. In 2010, Lou et al. [88] investigated the effects of UAE for the extraction of phenolic compounds as well as the phytoestrogen arctiin from burdock leaves (*Arctium lappa*, Asteraceae). The optimized conditions were a microwave power of 500 W, an extraction time of 30 s, a solvent to solid ratio of 500 W, and two extraction cycles. Using this new methodology, the phenolic yield was improved with simultaneous reduction of extraction time compared to classic maceration. This phenomenon was correlated to a strong disruption of the microstructure of the leaves due to intense shaking and violent collapse of cavitation bubbles generated by ultrasound combined with the fast heating of microwaves. Characterization of the extract was then carried out by UHPLC–QTOF–MS.

Furthermore, recently a Chinese group developed an interesting extraction strategy combining UAE and MAE for viable extraction of flavonoids from *Spaltholus suberectus* (Fabaceae) [89]. This approach was found to be more efficient and faster than commonly used extraction methods with optimal conditions determined as: microwave power of 300 W, extraction time of 450 s, 70% methanol as the extraction solvent, solvent to solid ratio of 20 mL/g, ultrasound power of 50 W, extraction temperature of 80°C, and one extraction cycle. 38 compounds including flavonones (e.g., liquiritigenin and naringenin), chalcones (e.g., isoliquiritigenin), and isoflavones (i.e., ononin, daidzein, calycosin, genistein, formononetin, and biochanin A) were further identified by UHPLC–QTOF–MS.

Finally, MAE combined to sonication (MAS) has been well employed for the extraction of bioactive isoflavonoids (genistein, daidzein, and biochanin A) and trans-resveratrol from peanuts (*A. hypogaea*, Fabaceae) [90]. The MAS method has been shown to extract significantly higher amounts of the selected compounds compared to stirring, sonication, and Soxhlet extraction.

**Fractionation and isolation of phytoestrogens**

The progress in the fractionation and isolation step could be considered less striking compared to extraction and analysis. Generally, it is restricted to the technological upgrade of existing platforms, the increase of the automation degree, and optimization of former methodologies. Faster and more usable systems, which are combined with collection and autosampler devices, are available, while new packing materials have been developed and more widely employed [52,55,59]. However, amongst these improvements, special attention should be given to CCC.

**Countercurrent chromatography:** CCC is a chromatographic technique which is based on the continuous liquid–liquid partition of molecules in a biphasic solvent system for separation, fractionation, and purification of bioactive natural products [91]. Advances in the mechanical parts of the previously developed apparatuses resulted in highly efficient and robust performance. This resulted in a dynamic comeback of the technique and increasing elaboration of CCC in phytochemistry laboratories in recent years. The main advantage of CCC compared to other separation techniques is the absence of solid support which avoids irreversible adsorption and degradation of molecules during the purification step. The two main modern apparatuses encountered in the literature are HSCCC and CPC which are characterized by a hydrodynamic and a hydrostatic equilibrium system, respectively. In the bioguided fractionation approach, the application of CCC is really useful since it may be used as a fractionation technique to chemically simplify a crude extract and concentrate minor compounds without any degradation or adsorption of phytochemicals on a solid stationary phase. Moreover, CCC can be used as a purification technique for the isolation of active compounds. CCC has been widely applied, and several well-known representatives including isoflavonoids, prenylated flavonoids, stilbenoids, lignans, and flavonolignans have been isolated. Several reviews describing solvent systems and methodologies for the isolation of estrogenic compounds such as xanthones and ubiquitous flavonoids using CCC technologies are also available [92,93]. The reader can refer to Table 1 which resumes recent attempts related to the isolation of phytoestrogens using HSCCC or CPC, while only some representative applications will be discussed in this section.

Special attention can be given to the work published by Du and Li who first used the combination of HSCCC and the DPH (2,2-diphenyl-1-picrylhydrazyl) antioxidant assay to separate an ethanolic extract of *Macuna sempervirens* (Fabaceae) leaves into several fractions [94]. Then estrogenic activity of 12 flavonoids, further isolated by preparative HPLC, was evaluated using a luciferase assay. The main active flavonoids were formononetin, genistein, kaempferol, 20-hydroxy-biochanin A, and dihydroquercetin. Overall, isoflavones from different Fabaceae species have been mainly purified by HSCCC using different kinds of solvent systems. Recently, an ethyl acetate extract of the roots of licorice (*Glycyrrhiza glabra*) was separated into 51 fractions using CPC [95]. The authors characterized the derived fractions by
UHPLC-MS and screened for estrogenicity on both estrogen receptors. Active fractions were correlated to the presence of prenylated flavonoids in the mixture. Furthermore, fractions containing galbrene, an isoflavone considered one of the principle estrogenic molecules of the licorice root, exhibited strong estrogenic activity on the ERα. Another original application has been developed by Renault’s group for the purification of prenylated chalcones and flavanones from *Humulus lupulus* by CPC using an acetone-based solvent scale CPC as well as HSCCC, which has also been used for purification of stilbenoids especially from *V. vinifera*, flavonolignans from *Silybum marianum* (Asteraceae), and lignans from several sources.

### Analysis and characterization of phytoestrogens

A unique feature of botanical extracts is the substantial number of constituents, their structural diversity and extensive variability. These features are directly dependent on seasonal and geographical aspects, harvesting time, biotic or abiotic stress, and the extraction procedure. Thus, it is obvious that highly demanding analytical methods for the complete characterization thereof are required. To this end, chromatographic techniques and methodologies offering high resolution, reproducibility, sensitivity, and accuracy are needed for both qualitative and quantitative purposes. Moreover, fast, however, reliable methods are needed for the characterization of such complicated mixtures, appropriate for the quality control of botanical and/or medicinal preparations. Thus, in the last years, different strategies have been developed in liquid chromatography to decrease analysis time while sustaining efficiency and resolution.

### Table 1 Isolation of phytoestrogens by CCC and CPC from different plant materials.

<table>
<thead>
<tr>
<th>Purified molecules</th>
<th>Plant material</th>
<th>Solvent system</th>
<th>Instrumentation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daidzin, genistin, genistein, daidzein</td>
<td>Glycine max</td>
<td>Elution gradient from (1:2:1:1:5) to (1:2:1:3:0.5:5) Hex/EtOH/EtOH/MeOH/CH3COOH/H2O</td>
<td>HSCCC</td>
<td>[163]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>CHCl3/MeOH/H2O (4:3:2); CHCl3/MeOH/CH3COOH/H2O (4:3:0.5:2); MtBE/THF/MeOH/H2O 0.5% TFA (2:2:1:0.15:4)</td>
<td>HSCCC</td>
<td>[164]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Radix astragali</td>
<td>CHCl3/MeOH/CH3COOH/H2O (2:1:1:1)</td>
<td>HSCCC</td>
<td>[165]</td>
</tr>
<tr>
<td>Liquiritigenin, isoliquiritigenin</td>
<td>Glycyrrhiza uralensis</td>
<td>Hex/EtOH/MeOH/ACN/H2O (2:2:1:0.6:2)</td>
<td>CPC</td>
<td>[95]</td>
</tr>
<tr>
<td>Puerarin, 3′-methoxy-puerarin</td>
<td>Pueraria lobata</td>
<td>EA/MeOH/H2O (2:1:3)</td>
<td>HSCCC</td>
<td>[167]</td>
</tr>
<tr>
<td>Prenylated isoflavones</td>
<td>Glycyrrhiza uralensis</td>
<td>Hex/EtOH/MeOH/H2O (6.5:5.5:6:4)</td>
<td>HSCCC</td>
<td>[168]</td>
</tr>
<tr>
<td>Prenylated flavonoids</td>
<td>Glycyrrhiza glabra</td>
<td>Hex/MeOH/H2O (5:9:1)</td>
<td>CPC</td>
<td>[95]</td>
</tr>
<tr>
<td>Isoflavones, flavonols</td>
<td>Mucuna sempervirens</td>
<td>Stepwise elution gradient using Hex/EtOH/MeOH/EA</td>
<td>HSCCC + HPLC preparative</td>
<td>[94]</td>
</tr>
<tr>
<td>Xanthohumol</td>
<td>Humulus lupulus</td>
<td>Hex/EtOH/MeOH/H2O (5:5:4:3)</td>
<td>HSCCC</td>
<td>[169]</td>
</tr>
<tr>
<td>Xanthohumol, isoxanthohumol, 6,8-diprenylmatairesinol</td>
<td>Humulus lupulus</td>
<td>Hept/Toluene/Acetone/H2O (24.8:2.8:50:22.4)</td>
<td>CPC</td>
<td>[170]</td>
</tr>
<tr>
<td>Resveratrol, emodin,physcion</td>
<td>Polygonum cuspidatum</td>
<td>Elution gradient from (3.5:4:6) to (3.5:7:3) light petroleum/EtOH/MeOH/H2O</td>
<td>HSCCC</td>
<td>[171]</td>
</tr>
<tr>
<td>Resveratrol, viniferin, vitisin C</td>
<td>Vitis riparia × Vitis berlandieri</td>
<td>Back step and back gradient with Hept/EtOH/MeOH/H2O (2:1:2:1) and (5:6:5:6)</td>
<td>CPC-MS</td>
<td>[172]</td>
</tr>
<tr>
<td>Resveratrol, viniferin, piceatannol, vitisin C, amepelosin A</td>
<td>Vitis vinifera stems</td>
<td>Hept/EtOH/MeOH/H2O (1:2:1:2 and 5:6:5:6)</td>
<td>CPC</td>
<td>[173]</td>
</tr>
<tr>
<td>Resveratrol, stilbenoids, catechin, epicatechin, vitisin C</td>
<td>Vitis vinifera</td>
<td>Hex/EtOH/H2O (1:8:2:7) and (4:5:3:3)</td>
<td>CPC</td>
<td>[174]</td>
</tr>
<tr>
<td>Resveratrol, anthraglycoside A, anthraglycoside B</td>
<td>Polygonum cuspidatum</td>
<td>CHCl3/MeOH/H2O (4:3:2)</td>
<td>HSCCC</td>
<td>[175]</td>
</tr>
<tr>
<td>Resveratrol, arachidin-1, arachidin-3</td>
<td>Arachis hypogaea</td>
<td>Hept/EtOH/H2O (4:5:3:3)</td>
<td>CPC</td>
<td>[176]</td>
</tr>
<tr>
<td>Secoisolariciresinol diglucoside</td>
<td>Linum usitatissimum</td>
<td>MtBE/MeOH/ACN/H2O (1:3:1:5)</td>
<td>HSCCC</td>
<td>[177]</td>
</tr>
<tr>
<td>Arctigenin, matairesinol</td>
<td>Forsythia koreana</td>
<td>Hex/EtOH/MeOH/H2O (5:5:5:5)</td>
<td>CPC</td>
<td>[178]</td>
</tr>
<tr>
<td>Silychristin, silydianin, taxifolin, silybinin, iso-silybin</td>
<td>Silybum marianum</td>
<td>Hex/CHCl3/MeOH/CH3COOH/H2O 0.5% CH3COOH (0.5:11:10:6)</td>
<td>HSCCC</td>
<td>[179]</td>
</tr>
<tr>
<td>Silychristin, silydianin, silybinin, isosilybin</td>
<td>Silybum marianum</td>
<td>Hept/EtOH/H2O (1:4:3:4)</td>
<td>CPC</td>
<td>[180]</td>
</tr>
<tr>
<td>Silychristin, silydianin, silybinin, isosilybin</td>
<td>Silybum marianum</td>
<td>Hex/EtOH/MeOH/H2O (1:4:3:4)</td>
<td>HSCCC</td>
<td>[181]</td>
</tr>
</tbody>
</table>

High-speed countercurrent chromatography (HSCCC), high-performance liquid chromatography (HPLC); acetic acid (CH3COOH), acetoneitrile (ACN), butanol (BuOH), chloroform (CHCl3), ethanol (EtOH), ethyl acetate (EA), heptane (Hept), hexane (Hex), methanol (MeOH), methyl tert-butyl ether (MtBE), trifluoroacetic acid (TFA), tetrahydrofuran (THF), water (H2O).
ventional HPLC. Additionally, because of the stationary phase particle sizes of around two micrometers or smaller, the performance and the quality is maintained or even improved [57]. Especially the combination of UHPLC with newly developed stationary phases such as monolithic supports, fused-core technologies (superficially porous particles), as well as the usage of an elevated temperature in the mobile phase gave new potentials in the analysis of complicated mixtures and especially the quality control of botanical preparations [96].

From the detection and identification point of view, HPLC or UHPLC hyphenated to MS is nowadays extensively employed in natural products chemistry. Especially for UHPLC, this growing interest is based on the fast and efficient separation capabilities of UHPLC combined with the increase of sensitivity and structural information obtained by MS. Further structural information could be derived from the employment of HRMS analysers as well as MS/MS (or HRMS/MS) and MSn acquisitions revealing valuable fragmentation patterns [97,98].

In the area of phytoestrogens, numerous studies have been carried out incorporating these technologies, aiming to the profiling of natural extracts, detection of phytoestrogens, and dereplication. Recent studies are summarized in Table 2 showing the potential of such technologies in fast and efficient analysis and characterization of phytoestrogens, especially isoflavonoids, from diverse plant extracts. Monolithic supports (i.e., Chromolith Performance RP-18e columns) have been successfully utilized for fast separation of isoflavonoids from different soy extracts. In some cases, low back pressure of such columns even allowed linking in a series two chromatographic columns for the analysis of soy isoflavones in less than 10 min [99]. Compared to classical HPLC, the setup of one or two monolithic columns reduced analysis time without losing any separation efficiency [99,100–102]. Another alternative is the use of fused-core columns, which dramatically reduce the analysis time, however, generate high back pressure. For this purpose, core-shell (e.g., kinetex) columns have been used for rapid determination of isoflavones in Radix puerariae [103] and G. max [104] using classical HPLC equipment. Optimized and selective RP stationary phases have been introduced for more efficient and accelerated analysis. For instance, Klejdus et al. proposed to use a dC18 reversed-phase fast chromatographic column (Atlantis) for the determination of 10 isoflavones (e.g., genistin, daidzein, formononetin, and biochanin A) using HPLC-DAD [105]. The separation was achieved in less than 8 min with good peak efficiency. This fast analytical procedure was then applied for the quantification of the detected isoflavones in different varieties of soybeans and soy plant organs [105,106] and was found subsequently improved [107]. Recently, the same group traced eight common isoflavones in different samples of sea algae, freshwater algae, and cyanobacteria using fast LC-MS with cyan-o-silica columns (Zorbax SB-CN) [108]. Further investigations at an elevated temperature (i.e., 80°C) on the same column chemistry finally led to an ultrafast separation (<1 min) of ten isoflavonoids from different plant extracts [109].

Furthermore, the employment of UHPLC has been increasing rapidly the last years for fast profiling of phytoestrogens [110–114]. For instance, a short run of 2 min was achieved for the separation of isoflavonoids and phenolic acids from extracts of different Fabaceae plants [111]. Also, different classes of phytochemicals, including isoflavonoids, were separated in less than 10 min from the aerial parts of 57 Trifolium species allowing at the same time their chemi-taxonomic classification [112]. Du et al. employed a short 1.8 mm C18 column at 46°C for the analysis of isoflavonoids in Radix puerariae extracts, achieving a separation of 14 compounds in 8 min [103].

Structure dereplication, profiling and fingerprinting of phytoestrogens: Complete characterization of the metabolites from crude extracts of plants is a challenging task due to their complex nature. Thus, fast metabolite identification generally requires methods providing high chromatographic resolution and sensitivity, as previously discussed, in combination with high-throughput identification techniques [115]. The determination of the chemical identity of known metabolites early in the discovery process (dereplication) as well as the structure elucidation of novel compounds is generally performed using HPLC-MS and NMR. With recent developments in the analysis of plant metabolomes, such as UHPLC-HRMS, and highly sensitive NMR probes (micro-, cryoprobes), it is now possible to detect hundreds of compounds simultaneously, in an untargeted manner, even in low quantity [115]. Numerous data derived from HPLC-MS and/or HRMS analysis of plant extracts containing potent estrogenic compounds can be found in the literature. Special attention should be given to the untargeted metabolomic approaches, including metabolic fingerprinting and profiling which are widely used in several applications.

For instance, in order to profile 13 H. lupulus varieties, Farag et al. [116] developed an approach using NMR, UHPLC-qTOF-MS, and HRMS (FTICR) in parallel and in combination with pattern recognition analysis (e.g., PCA). Application of this platform led to the identification of 46 molecules belonging to various chemical classes including bitter acids, flavonoids (i.e., chalcones, flavones, and flavonols), terpenes, fatty acids, and sugars. Xanthohumol was identified as the major principal prenylflavonoid, and FT-MS analyses suggested the presence of new isoprenylated compounds in hop resins, while PCA loading plots revealed a difference in bitter acid composition among the 13 H. lupulus varieties. The same research team [25] has also recently undertaken a metabolic profiling and fingerprinting of root extracts of Glycyrrhiza species (G. glabra, G. uralensis, G. inflata, and G. echinata) of different geographical origins. As previously, they used a combination of HPLC-MS, NMR, and multivariate analysis to differentiate licorice species based on their chemical composition. More than 60 compounds including triterpene saponins, flavonoids, and coumarins were detected, and 46 were identified. Among them glycyrrhizin, 4-hydroxyphenyl acetic acid, and glycosidic conjugates of liquiritigenin/isoliquiritigenin were those which significantly contributed to the discrimination between species. One of their former studies also reported the utilization in parallel of HPLC-UV-ESI-MS and GC-MS for the identification of 35 polyphenols including 26 isoflavones, 3 flavones, 2 flavanones, 2 aurones, and a chalcone in Medicago truncatula (Fabaceae) root and cell culture [117]. Another study realized by an Italian group focused on the discrimination of G. glabra from different geographic areas [118]. A metabolite profiling using HPLC-MS/MS has been applied for quantification and characterization of saponins and phenolic compounds of liquorice extract. The analysis of G. glabra roots has revealed differences in the pattern of molecules according to the source, allowing to distinguish the origin of the plant material: for instance, Chinese and Italian roots were the richest in glycyrrhizic acid. Simons et al. [119] described a rapid UHPLC-MS/MS screening procedure based on the fragmentation of prenyl substituents of flavonoids in the positive ion mode for the accelerated identification of prenylated flavonoids from plant extracts. Application of this screening method, especially based on the detection of neutral losses of 42 and 56 mass...
units resulted in the identification of 70 prenylated flavonoids [e.g., isoflavones, flavanones] in three licorice root extracts without any purification or fractionation steps.

Untargeted LC-MS profiling has also been adapted for characterization of lignans. Hanhineva et al. [120] identified well-known (e.g., secoisolariciresinol and matairesinol) and novel lignans in rye bran (Secale cereal, Poaceae) by the use of UHPLC-qTOF-MS/MS.

Unquestionably, a technique which should be underlined is HPLC-SPE-NMR. It is a hyphenated technique that uses SPE as an interface between chromatography and NMR. Thus, analytes eluted in chromatographic bands are automatically trapped on SPE cartridges and subsequently eluted into an NMR flow-probe with a small amount of deuterated solvents [121]. HPLC-SPE-NMR is a state-of-the-art technique which was incorporated recently for the identification of secondary metabolites in natural extracts and/or fractions and has also been utilized for the discovery of phytoestrogens. For instance online HPLC-SPE-NMR has been used for the fast identification of isoflavones from Smirnowia iranica (Fabaceae) [122]. Seventeen metabolites represented in 10 chromatographic peaks were finally identified, and the authors mentioned that this approach can be easily applied for the high-throughput structural identification of phytoestrogens.

HPLC-SPE-NMR has also been used for the analysis of numerous structurally related phytoestrogens such as lignans present in Phyllanthus urinaria and P. myrtifolius (Phyllanthaceae) or stilbenes from Syagrus romanzoffiana (Arecaceae) [123, 124]. Furthermore, HPLC-SPE-NMR is often associated, online or offline, with MS devices comprising a very efficient and powerful tool for rapid identification of known natural products. For instance, Hamburger’s group achieved the analysis of German iris (Iris germanica, Iridaceae) rhizomes using a combination of HPLC–DAD–MS and semipreparative HPLC/off-line microprobe-NMR measurements [125]. The authors demonstrated the potential of this approach for the rapid phytochemical profiling of plant extracts by purification of 20 isoflavones, isoflavone glycosides, and acetovanillone via two successive chromatographic

---

Table 2: Fast and ultra-fast analyses of phytoestrogens from plant extracts.

<table>
<thead>
<tr>
<th>Molecule class</th>
<th>Plant material</th>
<th>Column</th>
<th>Instrumentation detection</th>
<th>Analysis time</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>&lt;10 min</td>
<td>[99]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>&lt;25 min</td>
<td>[100]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>26 min</td>
<td>[101]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>10 min</td>
<td>[102]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Radix puerariae</td>
<td>HPLC-DAD</td>
<td></td>
<td>11 min</td>
<td>[103]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>6 min</td>
<td>[104]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>8 min</td>
<td>[105]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>8 min</td>
<td>[106]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Trifolium pratense</td>
<td>HPLC-DAD</td>
<td></td>
<td>4 min</td>
<td>[107]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>&lt;1 min</td>
<td>[109]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>8 min</td>
<td>[108]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Glycine max</td>
<td>HPLC-DAD</td>
<td></td>
<td>&lt;1 min</td>
<td>[109]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Soy supplement</td>
<td>UHPLC-MS/MS</td>
<td></td>
<td>3.5 min</td>
<td>[110]</td>
</tr>
<tr>
<td>Phenolic acids</td>
<td>Trifolium pratense</td>
<td>UHPLC-UV</td>
<td></td>
<td>2 min</td>
<td>[111]</td>
</tr>
<tr>
<td>Phenolic acids</td>
<td>Glycine max</td>
<td>UHPLC-UV</td>
<td></td>
<td>8 min</td>
<td>[112]</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Radix puerariae</td>
<td>UHPLC-DAD</td>
<td></td>
<td>8 min</td>
<td>[80]</td>
</tr>
<tr>
<td>Phenolic acids</td>
<td>Trifolium species</td>
<td>UHPLC-DAD</td>
<td></td>
<td>10 min</td>
<td>[112]</td>
</tr>
<tr>
<td>Isoflavones,</td>
<td>Vigna radiata</td>
<td>UHPLC-ESI-MS</td>
<td></td>
<td>17 min</td>
<td>[113]</td>
</tr>
<tr>
<td>flavonoids,</td>
<td>Glycine max</td>
<td>Acquity BEH shield RP-18</td>
<td></td>
<td>&lt;20 min</td>
<td>[114]</td>
</tr>
<tr>
<td>phenolic acids,</td>
<td>Acquity BEH shield RP-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>coumarins,</td>
<td>Acquity BEH shield RP-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flavanones,</td>
<td>Acquity BEH shield RP-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>coumarins,</td>
<td>Acquity BEH shield RP-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>phenolic acids,</td>
<td>Acquity BEH shield RP-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

High-performance liquid chromatography (HPLC); ultrahigh-performance liquid chromatography (UHPLC); diode array detector (DAD), electrospray ionization (ESI), mass spectrometry (MS)
steps. Furthermore, four isoflavone glycosides including a new compound, iriflogenin-4′-O-[β-D-glucopyranosyl(1 → 6)β-D-glucopyranoside] were detected for the first time in *I. germanica* rhizomes.

### Biological Screening and Activity Dereplication Strategies

The biological profile of a given compound is strongly related to several factors and parameters such as ligand-receptor interactions, absorption, metabolism, excretion, etc. which finally define its pharmacological effect. In the following section, only techniques based on molecular interactions between receptors and ligands giving rise to preliminary structure-activity relationships are described.

**Off-line HPLC-based detection and isolation of phytoestrogens**

As has been mentioned before in this review, the final goal of a phytochemical process is the identification of novel active compounds usually against a predefined pharmacological target. Therefore different approaches and concepts have been developed to fulfill this goal ([Fig. 1](#Fig1)). Nowadays and in most cases, a combination of these approaches are actually followed, and the traditional ones usually are framed with modern tools and methods. Analytical profiling concepts are very often implemented according to the identity of the starting material, the complexity of the sample under investigation, and the biological objectives.

For instance, a classical isolation approach was followed for the investigation of *Erythrina poepiggiana* (Fabaceae) methanol extract, however monitored with HPLC-DAD during the entire isolation procedure. Indeed, several prenylated isoflavones have been isolated, and their affinity to ERα and ERβ was evaluated using a receptor binding assay [126]. Furthermore, their estrogenic properties in a receptor subtype-specific reporter gene assay were determined [127]. Among the eleven isolated isoflavones, 3′-isoprenylgenistein exhibited the most potent estrogenic activity with a 10-fold lower binding affinity to genistein.

Likewise, in a more bioguided manner, Boué et al. evaluated the estrogenic activity of seven legumes using an estrogen-dependent proliferation assay in MCF-7 breast cancer cells together with HPLC-DAD analyses [128]. Among seven plants tested, kudzu root (*Pueraria lobata*, Fabaceae) and red clover blossom were the most estrogenic extracts and also exhibited the highest competitive binding to ERβ. To identify active compounds in the kudzu root extracts the authors used HPLC fractionation. The eleven pooled fractions were then analyzed for estrogenic activity using MCF-7 cells and were characterized by HPLC-MS. The fraction containing puerarin induced the highest level of cell proliferation (75.5%) followed by daidzin fraction (67.5%). Estrogenic activity was also significant inside fractions containing individually genistin, daidzein, and genistein.

Recently, the estrogen-like activity of 30 preparations used in traditional Chinese medicine were explored using the *Arabidopsis thaliana* pER8:GUS bioassay in combination with silica column and HPLC fractions [129]. Extracts of *Cuscuta chinensis* Lam. (Convolvulaceae), *Caesalpinia sappan* L. (Fabaceae), *Acorus gramineus* Soland (Acoraceae), *Cyperus rotundus* L. (Cyperaceae), and *Artemisia capillaris* Thunb. (Asteraceae) were found as compounds with moderate to good estrogenic activities if compared to soybean extract. Subsequent bioactivity-guided fractionation of the *C. sappan* extract led to the isolation of new and known compounds including two homoisoﬂavonoids, 3-deoxysappanone B and 3′-deoxysappanone B, and one chalcone derivative, 3-deoxysappanchalcone, which all expressed a significant estrogenic activity. Furthermore, brazilein and brazilin, two characterized homoisoﬂavonoid derivatives, inhibited the ERE transcription induced by 17β-estradiol. The same Chinese group further investigated *C. chinensis* using the same methodology and found that kaempferol and isorhamnetin were the major compounds from the ethanolic extract and the key contributors to its estrogenic activity [130].

Another example comprises the phytochemical study of *Dalbergia parviflora* Roxb. (Fabaceae) which led to the isolation of eight new phenolic compounds together with 32 known compounds [131]. The further assessment of their estrogenic activity was carried out by measuring cell proliferation in the estrogen-responsive breast cancer cell lines MCF-7 and T47D. Isoflavones such as genistein, biochanin A, tectorigenin, and 22-zethoxyformononetin showed the strongest estrogenic activities, but the new isoflavane (35)-sativanone was less active. Finally, fractionation guided by estrogenic activity, using recombinant yeast and Ishikawa Var-I (consisting of stimulation of the activity of alkaline phosphatase by estrogen in a human endometrial adenocarcinoma cell line) bioassays of the roots of *Sophora flavescens* Ait. (Fabaceae) led to the isolation of kurarine [132]. This flavanone, isolated by combining SPE and HPLC, exhibited a 10000-fold weaker estrogenic activity than 17β-estradiol.

### Online LC-based biochemical detection

This screening platform consists of a direct hyphenation of an HPLC-DAD apparatus and/or MS with BCD. Plant extracts are initially separated by HPLC and then a post-column (bio)chemical assay determines the bioactivity of the individual analytes while DAD, MS, or NMR can be used simultaneously as detectors for the identification of the active compounds. To date, a range of online post-column assays has been developed for the detection of radical scavengers, enzyme inhibitors, and ligands binding to specific receptors. This screening technology was reviewed in detail recently by Shi et al. [133].

LC-BCD using human ER was first described by a Dutch team [134]. The biochemical assay was based on the competition of fluorescence-labeled ligands and bioactive natural products. Coumestrol, a known ligand for ER, was used as label to enhance fluorescence. Typically, compounds are separated on HPLC columns, afterwards ER is added, and the mixture is allowed to interact for 30 sec in a reaction coil. At a second step, coumestrol is added to saturate the free binding sites of the ER. When it binds to the ER, the fluorescence emission of coumestrol is characterized by a blue shift to 410 nm, whereas the fluorescence intensity of the unbound coumestrol is 4 times less (at the same excitation wavelength). The difference in fluorescence between free and bound coumestrol allows detection of the presence of estrogenic compounds in the starting mixture [134, 135]. When MS is used to identify compounds in parallel to BCD, a flow splitter is placed before reaction coils in order to send HPLC effluent towards both detection systems (i.e., fluorescence detector and mass spectrometer) ([Fig. 4](#Fig4)).

Schobel et al. [136] used LC-BCD-MS for the screening of plant extracts in order to determine their ERα and ERβ binding activity. After a preliminary screening of 9888 extracts with only BCD detection, the six most active extracts were analyzed by LC-BCD-MS. The number of active compounds present in the ex-
Among them, two techniques, namely FAC chemical profiling, and in new potential drug lead identification. To screen protein-ligand interactions using mass spectrometry. In the past few years, several technologies have been developed. Online MS-based protein-ligand affinity determination has been demonstrated by van Elswijk's group study [137] in which three weak estrogenic compounds (i.e., luteolin, quercetin, and kaempferol) were identified. As a starting material, the hydrolyzed pomegranate (*Punica granatum*, Punicaceae) peel extract was used. Both determination of biological activity and chemical characterization were performed in parallel during a single chromatographic run. A more complex screening approach combining RP-HPLC online to an ERα affinity detection system using fluorescence polarization has been recently described by Reinen et al. [138]. This method is a sensitive screening platform for measuring the ERα binding affinities of individual components in mixtures. Even if it has not been applied to plant extracts yet, it has been demonstrated that it is possible to separate and detect ERα ligands (i.e., coumestrol, coumarol, and zearalenone) from a standard mixture. Overall, LC-BCD seems to dramatically improve the determination of bioactivity in complicated mixtures and could be applied as a screening method. Moreover, it provides, at the same time, structural information of the active molecules without the need of laborious isolation steps and biological evaluation of individual compounds.

**Online MS-based protein-ligand affinity determination**

In the past few years, several technologies have been developed to screen protein-ligand interactions using mass spectrometry. These new methods, well reviewed by Jonker et al. [139], have become of great interest in natural extract screening, bioactive chemical profiling, and in new potential drug lead identification. Among them, two techniques, namely FAC-MS and ultrafiltration MS, have been employed for detection of phytoestrogens. As a general rule, these methods involve several steps before detection of bioactive compounds including: (1) protein-ligand complexation, (2) separation of non-bound compounds from the protein-ligand mixtures, (3) release of the active ligand by elution-dissociation, and (4) mass spectrometry detection of eluted ligands [139].

**Frontal affinity chromatography coupled to MS:** First introduced by Schriemer et al. [140], FAC-MS is based on the continuous infusion of samples containing potential ligands through a stationary phase containing the immobilized target protein. The eluting compounds are monitored and identified by MS following their *m/z* values. In detail, ligands are retained according to their affinity for the target protein and detected as characteristic breakthrough curves (Fig. 5). All non-retained compounds will break through in the void volume, whereas active molecules will bind to the target protein. At some point, the target will start to be saturated by the ligand producing a rise in the chromatographic trace, which is terminated when the target is totally saturated (measured concentration equal to initial one). Thus, the observed breakthrough time reflects the relative binding strengths of the ligand (single or in mixture) therefore allowing the identification of active molecules [139, 141, 142].

Several studies have shown the power of this technique for the investigation of molecular interactions between estrogen receptors and known ligands. For instance, Chan et al. [143] use chromatographic assays involving immobilized ERβ in microaffinity columns to study the binding affinity of nafoxidine, tamoxifen, and 17β-estradiol. Moaddel and coworkers [144] report the immobilization of human ER onto a stationary phase with a silica backbone. This specific column was then connected to MS, and six known ER ligands, including genistein, were analyzed. The obtained ligand binding affinities were in accordance with previous works.

Another online and more robust FAC-MS platform has been described by Ng et al., using two columns containing immobilized human ERβ, while the first column is being regenerated, the other is being used. The overall system is coupled to an API 3000 triple-quadrupole MS equipped with an ESI interface. The total runtime was 30 min for the screening of 100 compounds. The authors described the automated FAC-MS as a moderate primary HTS system and expected because of its wide extrapolation that in a continuous 24-h operation, 10 000 ligands could be potentially tested [145].

**Pulsed ultrafiltration-MS:** Originally developed by van Breemen et al. [146], PUF-MS is an online combination of ultrafiltration extraction and MS detection which facilitates the identification of potential ligands in complex mixtures such as plant extracts. Basically, when a mixture of compounds is injected into the ultrafiltration unit which contains a biological macromolecular receptor, the ultrafiltration membrane retains macromolecules and ligand–receptor complexes; however, it allows the unbound (non-binders) low molecular molecules to escape from the chamber (washing). At the elution phase, the bound ligands (binders) are dissociated from the receptor and eluted from the chamber by buffer adjustment (Fig. 6). The eluent from the ultrafiltration cell is then either analyzed using online MS or is forwarded to an HPLC column coupled with an MS detector [139, 147].

Using ERα and/or ERβ as receptors, this approach can be easily used for fast identification of phytoestrogens from complex sam-

---

**Fig. 4** Typical online LC-BCD-MS system used for biodetection and characterization of estrogenic compounds. 1. Separation of a complex mixture on an HPLC column; 2. splitting the flow to mass spectrometry and reaction coil system; 3. binding of ER and potential ligands in the reaction coil; 4. binding of fluorescent coumestrol and free estrogen receptors in the reaction coil; 5. detection and quantification of the receptor-ligand complex by fluorescence. (Color figure available online only.)
Liu et al. [148] evaluated the estrogenic potential of eight plant extracts using an ER competitive binding assay for the determination of estrogenicity along with screening and identification of the active molecules by ultrafiltration LC-MS. Among the eight botanical preparations, red clover (T. pratense, Fabaceae) as well as chasteberry (Vitex agnus-castus, Lamiaceae) and hops (H. lupulus, Cannabaceae) showed a significant binding affinity to ERα and ERβ. Subsequent screening of the red clover extract by ultrafiltration LC-MS demonstrated that genistein was the most active ERβ binder followed by daidzein and biochanin A. Further investigations by the same group [149], using a combination of ER binding and ultrafiltration LC-MS, led to the identification of linoleic acid as an estrogenic compound in chasteberries. They also screened the binding affinity to ERα and ERβ of red clover and hop extracts by ultrafiltration LC-MS [150]. Results led to the identification of genistein, daidzein, formononetin, and biochanin A in red clover, as well as 8-prenylnaringenin and isoxanthohumol in hops, verifying their approach.

More recently, Choi and van Breemen [151] described a slightly different screening assay for ligands to the estrogen receptor based on magnetic microparticles and LC-MS. This assay, like ultrafiltration LC-MS, is compatible with a wide range of buffers and prevents the introduction of non-binders into the LC-MS. In this variant, a solution containing potential ligands is firstly incubated with ER immobilized on magnetic particles, and then the ligand-receptor complexes are isolated magnetically. After washing to remove unbound compounds, the ligands are released with methanolic solution and characterized by LC-MS. The authors applied this approach for the screening of red clover (T. pratense, Fabaceae) and hop (H. lupulus, Cannabaceae) extracts. Phytochemicals with the highest affinities for ERβ in red clover were genistein and daidzein followed by formononetin and biochanin A. In the hops extract, 8-prenylnaringenin was found as the most potent binder, while isoxanthohumol, xanthohumol, and 6-prenyl-naringenin also exhibited activity but were less potent.

Finally, Onorato and Hellion [152] explored the binding affinity to ERβ of triterpenes from black cohosh (Cimicifuga racemosa, Ranunculaceae), a plant species known for its positive effects for the treatment of postmenopausal symptoms, combining off-line affinity ultrafiltration and LC-MS. However, they did not observe a significant binding affinity among the three triterpene glycosides and the aglycons tested.

**Online magnetic bead protein affinity LC-MS assay:** This recent methodology introduced by Jonker et al. [153] employs magnetic nanoparticles on which the target protein is immobilized and the subsequently bound ligands are analyzed by SPE-LC-MS. After in-
cubation of the analyte mixture with polyhistidine tag target proteins, the protein-ligand complexes are mixed with Co(II)-PAB and then injected into a magnetic trapping device. Bioactive ligands that bound to the protein–Co(II)-PABs are retained in the magnetic field while unbound compounds are removed by washing. Bound ligands are then eluted using a pH shift and analyzed online by LC-MS. This approach was applied in a screening assay employing ERα and a mixture of possible ligands including coumestrol. The authors demonstrated that this screening approach was able to easily distinguish binders and non-binders from a compound mixture and also binders with a weak affinity. Furthermore, the methodology is rapid, allowing for the screening of 40 compounds in 45 min and providing simultaneously information on the affinity and chemical characteristics of active compounds.

**Miniaturization**

*Microchip: Particular attention should be paid to the development of microfabricated devices consisting of miniaturized systems with dimensions in the range of micrometers to millimeters. Miniaturization leads to high-throughput devices combining faster analysis times and parallel reactions, as well as reduced consumption of expensive enzymes, proteins, and substrates [154]. Binding of ER-ligand complexes to ERE induces gene activation and is an important step in estrogen-induced biological effects. For this purpose, microchips have been developed to investigate the effects of some phytoestrogens such as genistein, daidzein, and coumestrol on the interaction of ERα and ERβ to ERE [155, 156]. Although effective, these methods have never been adapted to the characterization of phytoestrogens from complex mixtures and seem restricted until today to the investigation of the particular effect of known molecules. However, Imura et al. [157] quite recently opened a new area for development for these microchips. Indeed, they realized a microbioassay system comprising a micro-intestine, micro-liver, and the target components. The microchips are composed of slide glass and PDMS sheets with two microchannels mimicking a human intestinal lumen and a blood vessel. Different kinds of cells (i.e., Caco-2, MCF-7, and HepG2) have been introduced into the microchips. By using this system, the authors studied intestinal absorption, hepatic metabolism, and bioactivity towards target cells of compounds such as 17β-estradiol and soy isoflavones.

**Biosensors:** Biosensor technology offers a useful alternative to facilitate detection of bioactive substances. A biosensor is an analytical device composed of a sensitive biological element (e.g., microorganisms, cellular receptors, enzymes, antibodies) in contact with a physicochemical transducer. The transduction system, generally consisting of electrochemical, optical, or piezoelectric devices, transforms the signal resulting from the interaction of the analyte with the biological element into a measurable signal. [158]. The need for a convenient method to detect phytoestrogens has led researchers toward the development of specific biosensors. Andreescu and Sadik [159] reported the use of a simple amperometric tyrosinase-based biosensor for the measurement of phenolics and specifically phytoestrogens. The sensor responses have been evaluated for the detection of phenolic mixtures including the natural phytoestrogens resveratrol, genistein, and quercetin. The authors demonstrated that it was possible to detect phytoestrogens with phenolic groups inside the mixture using a biosensor based on tyrosinase activity. Recently, Woo’s Korean team developed a FRET biosensor probe of ERα conjugated to a nanoparticle for the detection of phytoestrogens [160]. FRET is a technique for measuring interactions between two fluorescent molecules currently used to monitor ligand–protein interactions. In this case, the FRET sensor was built based on the ligand-binding domain of ERα. The nanoparticle–based FRET sensor permitted to quantify and differentiate antagonists such as stilbenoids (i.e., resveratrol) from agonists such as isoflavones (i.e., daidzein and genistein) basing on signal intensities. Two other biosensors detecting estrogenic activity can be found in the literature, nonetheless the authors did not focus on phytoestrogen but their works are related to the ER receptor. Briefly, Liang et al. [161] developed an efficient yeast-based system able to screen and evaluate the selectivity of ligands for ER subtypes. This biosensor included a reporter gene plasmid and an ER expression plasmid. After incubation, the agonistic effects of chemicals (e.g., genistein) and their receptor affinities were assessed by measuring the β-galactosidase activity. On the other hand, Carmon et al. [162] described a biosensor that detects estrogenic substances using a quartz crystal microbalance with a genetically engineered construct of the hormone-binding domain of the ERα. The receptor was immobilized to a piezoelectric quartz crystal forming a uniform orientation on the crystal surface. This allowed the rapid monitoring of the presence of ligands that bind to the ER as well as the characterization of those ligands. However, although effective, all these biosensors need further investigation to enable the application of this technology to the identification of estrogenic compounds from more complex samples such as food and plant extracts.

**Conclusion**

Indubitably, phytoestrogens may represent valuable agents to manage menopausal complaints and encompass a core scientific topic not only in molecular biology but also in natural products chemistry. Recently, various technological advances have emerged, and novel methodologies have been developed facilitating considerably the discovery and biological evaluation of estrogen-like natural compounds. Nevertheless, many aspects related mostly to the rational, sensitive, and targeted disclosure of bioactives require additional scientific effort. Even if significant progress has been made in the area of natural product chemistry, discussed in the first part of this review, it seems insufficient. The power of other disciplines should be also incorporated and exploited in combination. A crucial scientific space where phytochemistry meets molecular biology, medicine, bioinformatics, to name just a few, needs to be covered. Briefly, some attempts have been mentioned mainly in the second part of this review, however much more multidisciplinary concepts are needed in order to not only reveal the pharmacological potential of phytoestrogens but also to make it available to users.

**Conflict of Interest**

The authors sincerely declare that they have no conflicts of interest.
none reductase 2 and aromatase inhibitory activities. J Med Chem 2009; 52: 1873–1884


Herrero M, Mendiola JA, Cifuentes A, Ibáñez E. Supercritical fluid extraction: Recent advances and applications. J Chromatogr A 2010; 1217: 2495–2511

Seo J, Yoo J, Suh J. Comparison of supercritical fluid extraction and solvent extraction of isoflavones from soybeans. J Liq Chromatogr Relat Technol 2009; 32: 923–932


Pascual-Martí MC, Salvador A, Chafier A, Berna A. Supercritical fluid extraction of resveratrol from grape skin of Vitis vinifera determined by HPLC. Talanta 2001; 54: 735–740


Choi Y, Kim J, Yoo KP. High performance liquid chromatography-electrospray ionization MS-MS analysis Forsythia koreana leaves, fruits, and stems. Enhancement of the efficiency of extraction of arctigenin by use of supercritical-fluid extraction. Chromatographia 2003; 57: 73–79

Michel T, Destanda E, Elifurk C. Evaluation of a simple and promising method for extraction of antioxidants from sea buckthorn (Hippophaé rhamnoïdes L.) berries: Pressurised solvent-free microwave assisted extraction. Food Chem 2011; 126: 1380–1386


Terigar BG, Balasubramanian S, Bolder D, Xa Z, Lima M, Sabllov CM. Continuous microwave-assisted isoflavone extraction system: design and performance evaluation. Bioprocess Technol 2010; 11: 2466–2471

Nemes S, Orsat V. Microwave-assisted extraction of secoisolariciresinol diglucoside-method development. Food Bioprocess Technol 2011; 4: 1219–1227


114 Shi SY, Zhang YP, Jiang XY, Chen XQ, Huang KL, Zhou HH. Coupling HPLC to on-line, post-column (bio)chemical assays for high-resolution


135 van Elswijk DA, Irth H. Analytical tools for the detection and characterization of biologically active compounds from nature. Phytorev Chem 2002; 1: 427–439


144 Zhang T, Cao X, Han X. Preparation of national certified reference materials of active compounds from natural products by CCC. J Liq Chromatogr Relat Technol 2003; 26: 3497–3509


146 Chuang YJ, Huang JW, Mukamba H, Tsai ML, Li CW, Chen SH. Electrophoretic mobility shift assay on poly(ethylene glycol)-modified glass microchips for the study of estrogen responsive element binding. Electrophoresis 2006; 27: 4148–4165


149 Andreasons S, Sadik OA. Correlation of analyte structures with biosensor responses using the detection of phenolic estrogens as a model. Anal Chem 2003; 75: 662–669


153 Shinomiya K, Kabasawa Y, Nakazawa H, Ito Y. Countercurrent chromatographic separation of soybean isolavones by two different types of coil plate centrifuges with various two-phase solvent systems. J Liq Chromatogr Relat Technol 2003; 26: 1565–1576


157 Zhang T, Cao X, Han X. Preparation of national certified reference materials of active compounds from natural products by CCC. J Liq Chromatogr Relat Technol 2003; 26: 1565–1576


This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.


179 *Liu H, Yuan Q, Li CF, Huang TX.* Isolation and purification of silychristin, silydianin and taxifolin in the co-products of the silybin refined process from the silymarin by high-speed counter-current chromatography. Proce Biochem 2010; 45: 799–804


Please note: This article was changed according to the following erratum on March 28, 2013: © Fig. 2 was changed.