Phytochemistry 70 (2009) 262–269

Contents lists available at ScienceDirect

Phytochemistry

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

The chain length of lignan macromolecule from flaxseed hulls is determined by the incorporation of coumaric acid glucosides and ferulic acid glucosides

Karin Struijs [a], Jean-Paul Vincken [a], Timo G. Doeswijk [b], Alphons G.J. Voragen [a], Harry Gruppen [a,*]

[a] Laboratory of Food Chemistry, Wageningen University, P.O. Box 8129, 6700 EV Wageningen, The Netherlands
[b] Biometris, Wageningen University, P.O. Box 100, 6700 AC Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 23 September 2008
Received in revised form 6 December 2008
Available online 18 January 2009

Keywords:
Flaxseed
Lignan
Lignan macromolecule
Secoisolariciresinol
p-Coumaric acid glucoside
Ferulic acid glucoside
Extinction coefficient
Chain length

ABSTRACT

Lignan macromolecule from flaxseed hulls is composed of secoisolariciresinol diglucoside (SDG) and herbacetin diglucoside (HDG) moieties ester-linked by 3-hydroxy-3-methylglutaric acid (HMGA), and of *p*-coumaric acid glucoside (CouAG) and ferulic acid glucoside (FeAG) moieties ester-linked directly to SDG. The linker molecule HMGA was found to account for 11% (w/w) of the lignan macromolecule. Based on the extinction coefficients and RP-HPLC data, it was determined that SDG contributes for 62.0% (w/w) to the lignan macromolecule, while CouAG, FeAG, and HDG contribute for 12.2, 9.0, and 5.7% (w/w), respectively.

Analysis of fractions of lignan macromolecule showed that the higher the molecular mass, the higher the proportion of SDG was. An inverse relation between the molecular mass and the proportion (%) CouAG + FeAG was found. Together with the structural information of oligomers of lignan macromolecule obtained after partial saponification, it is hypothesized that the amount of CouAG + FeAG present during biosynthesis determines the chain length of lignan macromolecule.

Furthermore, the chain length was estimated from a model describing lignan macromolecule based on structural and compositional data. The average chain length of the lignan macromolceule was calculated to be three SDG moieties with CouAG or FeAG at each of the terminal positions, with a variation between one and seven SDG moieties.

1. Introduction

Lignans are phenolic compounds, which, by definition, are composed of two propyl-benzene moieties, which are linked by a bond between the 8 and the 8′ position (Moss, 2000). The propyl-benzene structures, coniferyl alcohols, are biosynthesized through the phenylpropanoid pathway and stereospecifically dimerized into the lignan pinoresinol (PINO) (Davin et al., 1997).

Subsequently, reduction, oxidation, dehydrogenation and addition reactions lead to the formation of a broad range of lignans, which are present in both aglyconic and in glycosylated forms (Ford et al., 2001; Hano et al., 2006; von Heimendahl et al., 2005; Youn et al., 2005). Lignans are of interest because of their potential (anti)-estrogenic and antioxidant properties (Eklund et al., 2005; Wang, 2002).

Secoisolariciresinol diglucoside (SDG) is the most abundant lignan present in flaxseed (Johnsson et al., 2000). It is synthesized from PINO by the action of pinoresinol–lariciresinol reductase (PLR) (von Heimendahl et al., 2005), followed by glucosylation by an UDPG:glucosyltransferase (Ford et al., 2001). The formation of SDG takes place in the outer layer of the seed (Hano et al., 2006). Therefore, the concentration of SDG found in flaxseed hulls is higher than that of whole seeds (Madhusudhan et al., 2000). The lignans in flaxseeds (hulls) are part of an oligomeric structure, called lignan macromolecule. Within this lignan macromolecule, SDG is ester-linked by 3-hydroxy-3-methylglutaric acid (HMGA) (Kamal-Eldin et al., 2001; Klosterman and Smith, 1954). Also herbacetin diglucoside (HDG), ferulic acid glucoside (FeAG) and *p*-coumaric acid glucoside (CouAG) are part of the lignan macromolecule (Johnsson et al., 2002; Struijs et al., 2007). The chemical structures of the constituents of lignan macromolecule are shown in Fig. 1.

Analysis of seeds at different developmental stages showed that hardly any free SDG is present in developing seeds. Almost directly after its formation, SDG is incorporated within the lignan macromolecule (Ford et al., 2001; Hano et al., 2006). Oligomers of SDG and HMGA are formed by ester-linkage with CoA-activated HMGA (Ford et al., 2001). Molecules consisting of one or two SDG moieties with one, two or three HMGA moieties have been identified (Ford et al., 2001). Just like SDG, CouAG and FeAG are incorporated in an alkali labile structure in early stages of seed development (Ford et al., 2001; Hano et al., 2006).

Only limited data about the composition and size of lignan macromolecule are available. In 2001, a model of lignan

* Corresponding author. Tel.: +31 317 483211; fax: +31 317 484893.
E-mail address: harry.gruppen@wur.nl (H. Gruppen).

0031-9422/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2008.12.015

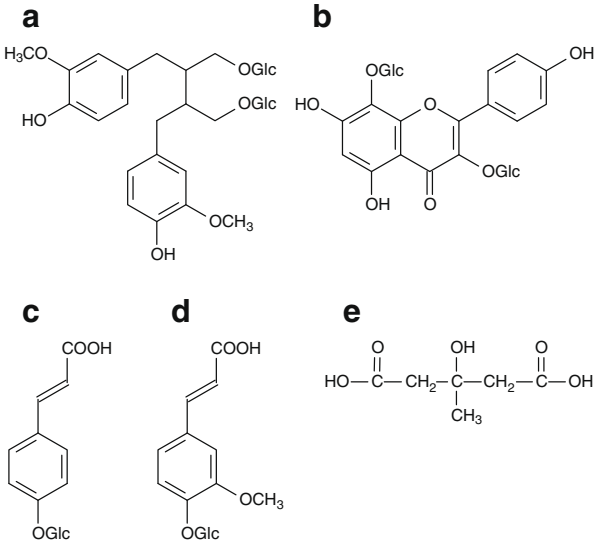

Fig. 1. The chemical structures of the constituents of lignan macromolecule from flaxseed hulls. a = SDG, b = HDG, c = CouAG, d = FeAG and e = HMGA.

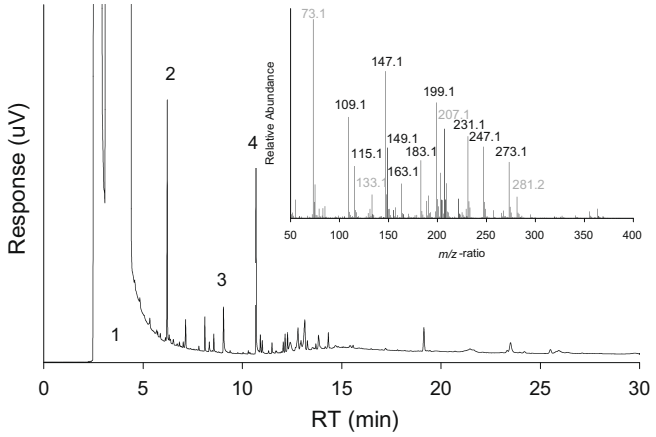

Fig. 2. GC profile of fully saponified lignan macromolecule after derivatisation. 1 = pyridine, 2 = derivatizing agent, 3 = glutaric acid (internal standard), 4 = HMGA. In the insert the GC-MS spectrum of peak 4. *m*/*z*-ratios ($[M + H]^+$) marked gray are also present in other fractions eluting at different retention times and are, therefore, not originating from HMGA.

macromolecule consisting of five SDG units being esterified with four HMGA residues was published (Kamal-Eldin et al., 2001). However, in later studies, variation in the proportions of constituents was observed between fractions of lignan macromolecule separated by reversed-phase solid-phase extraction (Johnsson et al., 2002; Strandas et al., 2008), indicating that lignan macromolecule represents an array of molecules with varying composition. In addition, it has been shown that the flavonoid HDG links, just like SDG, within lignan macromolecule via HMGA (Struijs et al., 2007), while CouAG and FeAG are linked directly to the glucosyl moieties of SDG through their carboxylic groups (Struijs et al., 2008).

More precise data about the composition and size of lignan macromolecule could give direction to further research on the biosynthesis of lignan macromolecule. The aim of present research is to investigate the composition and size of lignan macromolecule from flaxseed hulls in order to be able to determine the correlation between the composition and the size of the molecules.

2. Results and discussion

2.1. Quantification of the constituents of lignan macromolecule

An extract of lignan macromolecule from flaxseed hulls was obtained by extraction with 63% aq. EtOH (Struijs et al., 2007). Monomeric units of lignan macromolecule were obtained by treating the lignan macromolecular extract with 75 mM NaOH (Struijs et al., 2007). The components described to be part of lignan macromolecule from flaxseed hulls, can be divided into two groups: the phenolic constituents (see later) and the linker molecule HMGA. Fig. 2 shows the GC profile of HMGA from fully saponified lignan macromolecule. To confirm that peak 4 corresponded to HMGA, it was analyzed by GC–MS. In the insert of Fig. 2, the MS spectrum of peak 4 (RT = 10.9 min) is shown. By comparing the spectrum with the spectrum of an authentic standard of HMGA and with literature values (Ford et al., 2001), this peak was identified as HMGA. The concentration HMGA was calculated as 111 ± 13 μg/mg macromolecule. This number corresponds with the amount determined on the basis of the recovered weights of HMGA isolated from lignan macromolecule as reported previously (Prasad, 2004).

Besides the 11.1% (w/w) of HMGA, lignan macromolecule consisted of phenolic compounds, mainly SDG, CouAG, FeAG and HDG. To be able to determine the proportion of each of these phenolic constituents, they were purified by preparative RP-HPLC (Struijs et al., 2007), and their molar extinction coefficients were determined, as these data have not been reported in literature.

The molar extinction coefficients at 280 nm of purified SDG, CouAG, FeAG and HDG are given in Table 1. In this table, also the purity of the compounds, based on NMR signals is given (see Struijs et al., 2007 for NMR). The purity of the compounds was found to be close to 90% or higher.

On a molar basis, SDG was found to be the most abundant constituent, followed by CouAG, FeAG, and HDG as shown in Table 2. Recalculation of the molar proportions of the phenolic compounds into weight ratios showed that lignan macromolecule consists of 62.0% (w/w) SDG, 5.7% HDG, 12.2% CouAG and 9.0% FeAG. These

Table 1
Molar extinction coefficients of the four main phenolic constituents of lignan macromolecule.

Constituent	Purity	Molar extinction coefficient at 280 nm (M^{-1} cm^{-1})	SD absolute (%)	λ_{max} (nm)
CouAG	92%	19474	1043 (5.4%)	297
FeAG	98%	12133	292 (2.4%)	291 + 316
HDG	89%	10347	158 (1.5%)	271
SDG	100%	5838	201 (3.4%)	281

Table 2
Molar proportions and weight percentages of the constituents of lignan macromolecule.

Constituent	Molar proportions[a] of phenolic constituent (%)	±SD	Weight percentage[b] (%, w/w)	±SD
CouAG	23.1	±2.8	12.2	±1.5
FeAG	15.6	±0.9	9.0	±0.5
HDG	5.6	±0.4	5.7	±0.4
SDG	55.7	±2.0	62.0	±2.2
HMGA			11.1	±1.3

[a] Determined based on peak area of the RP-HPLC profiles (not shown; see Struijs et al. (2007) and the molar extinction coefficient (Table 1).
[b] The weight percentage (%, w/w) of HMGA was determined by GC. The weight percentage (%, w/w) of phenolic constituents were calculated from molar proportions.

values correspond with literature values where different methods were used (Eliasson et al., 2003; Prasad, 2004).

The proportion of CouAG was higher than the proportion of FeAG. This might be explained by the biosynthetic pathway of these compounds. During monolignol biosynthesis, CouA is formed earlier than FeA (Boerjan et al., 2003). Besides, it was found that incorporation of CouA and FeA in lignan macromolecule occurs at early stages of seed development (Ford et al., 2001). Based on these literature data and the results obtained in the present study, it is suggested that when CouA is glucosylated and/or it is incorporated within lignan macromolecule, it is not available anymore for conversion into FeA and FeAG. This might explain the lower proportion of FeAG in lignan macromolecule compared to CouAG.

2.2. Molecular mass of lignan macromolecule

Since precise data on the molecular mass of lignan macromolecule are lacking, MALDI-TOF MS analysis was performed (Fig. 3). The spectrum obtained by MALDI-TOF MS showed a complex pattern of peaks, which showed variation in composition and size of the molecules of lignan macromolecule. The spectrum showed four groups of peaks, each group differing 812 (m/z; 812 = 830-water) from each other, corresponding to a difference of SDG + HMGA. Within each group, a repetitive pattern of five clusters (A–E) of peaks was observed. The differences in m/z-ratios between the clusters varied. An enlargement of the first group, in which the individual peaks in the five clusters are marked, is shown in the insert of Fig. 3. Each of the five clusters showed a unique peak pattern. For example, in cluster A, two peaks were observed while in cluster E six peaks were distinguished. The differences in m/z-ratio between the peaks mostly corresponded to 12, 14, 16 or 18. The difference in m/z-ratio of 16 was tentatively assigned as the difference between the sodium versus the potassium adduct.

In Table 3 the peaks of the first group are listed together with their molecular masses and their (tentative) assignments. Peaks were assigned by making combinations of SDG, HMGA, CouA(G), FeA(G), and HDG or other phenolic compounds described to be present in flaxseed: the lignans matairesinol (MAT) (Liggins et al., 2000), isolariciresinol (iso-LARI) (Meagher et al., 1999), pinoresinol (PINO) (Meagher et al., 1999) and lariciresinol (LARI) (Sicilia et al., 2003), the flavonoid kaempferol diglucoside (Qiu et al., 1999), and the hydroxycinnamic acids caffeic acid glucoside (Westcott and Muir, 1996) and sinapic acid (Dabrowski and Sosulski, 1984).

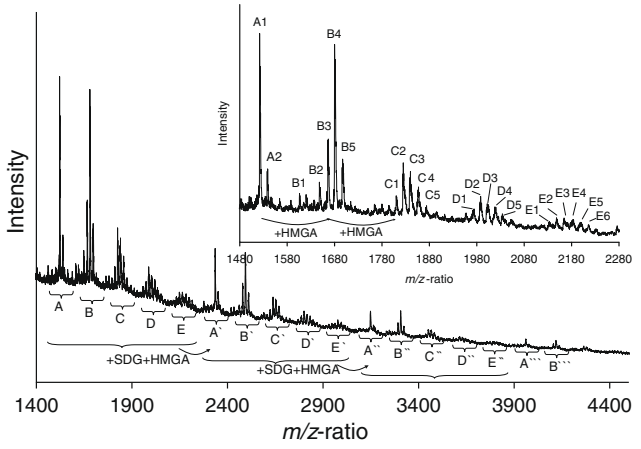

Fig. 3. MALDI-TOF MS spectrum of lignan macromolecule measured in the positive reflector mode. Four groups of five clusters (A–E) of peaks are shown. Each group represents an elongation of the macromolecular chain by SDG + HMGA compared to the previous group. An enlargement of 1st group is depicted in the insert. Peak assignments are listed in Table 3.

Table 3
Assignments of m/z-ratios found the first group of peaks in the MALDI-TOF MS spectrum of lignan macromolecule as shown in Fig. 3.

peak	m/z-ratio $[M+Na]^+$	possible assignments $[M+Na]^+$	partially assigned peaks[a]
A1	1521.6	2SDG + HMGA	
A2	1537.6	2SDG + HMGA $[M + K]^+$[b]	
B1	1605.6	SDG + 2HMGA + HDG	1521.6 + HMGA
B2	1647.7		
B3	1665.7	2SDG + 2HMGA	
B4	1679.8		
B5	1695.7		
C1	1809.9	2SDG + 3HMGA	1521.6 + 2 HMGA
			1647.9 + Glc
C2	1823.8		1679.8 + HMGA
			1647.7 + FeA
C3	1837.9		
C4	1855.9		1679.8 + FeA
C5	1873.8		
D1	1971.9		1647.7 + 2Glc
			1647.7 + CafAG
	1973.9	2SDG + 2HMGA + CouAG	1665.7 + CouAG
D2	1985.9	2SDG + 3HMGA + FeA	1823.8 + Glc
			1809.9 + FeA
			1647.7 + FeAG
	1987.9		1971.9 + 16
D3	1999.9		1837.9 + Glc
			1823.8 + FeA
			1855.9 + HMGA
	2001.9	2SDG + 3HMGA + FeA $[M + K]^+$[b]	1985.9 + 16
	2003.9	2SDG + 2HMGA + FeAG	1665.7 + FeAG
D4	2017.9		1855.9 + Glc
			1679.8 + FeAG
	2019.9	2SDG + 2HMGA + FeAG $[M + K]^+$[b]	2003.9 + 16
D5	2031.9		1855.9 + FeA
	2033.9		2017.9 + 16
E1	2131.0		1823.8 + CouAG
E2	2148.1	2SDG + 3HMGA + FeAG	1985.9 + Glc
			1809.9 + FeAG
			1823.8 + CafAG
E3	2164.0 (av)		
E4	2184.0 (av)		
E5	2199.9 (av)	2SDG + HMGA + 2FeAG	2003.9 + FeAG
E6	2215.9 (av)	2SDG + HMGA + 2FeAG $[M + K]^+$[b]	2148.1 + 16

(av) = most dominant isotope is not the first isotope peak, assignments calculated based on average molecular weights in stead of monoisotopic molecular weights.

[a] Differences in m/z-ratios between peaks could be identified as constituents of lignan macromolecule.

[b] Mostly fragments were assigned as sodium adducts $[M + Na]^+$ but in some occasions the potassium adduct $[M + K]^+$ was found. The potassium adduct shows a difference of 16 with the sodium adduct.

Peak A1, for example, was assigned as 2SDG + HMGA and peak A2 as the potassium-adduct of this same molecule. In the B-cluster, peak B3 was assigned as A1 elongated with one HMGA moiety, and in the C-cluster, peak C1 was assigned as A1 plus two HMGA moieties. Several peaks within the D and E clusters were identified as molecules composed of SDG and HMGA to which CouA, FeA or their glucosides were attached.

Most other peaks were partially assigned by looking at differences in m/z-ratios between peaks and assigning the differences as components present in flaxseeds. For example, peak E1 (m/z = 2131 $[M–H]^-$) was partially assigned as peak C2 (m/z = 1823.8 $[M–H]^-$) plus CouAG (m/z = 325 $[M–H]^-$). Differences in m/z-ratios with unidentified peaks corresponding to HMGA, glucose, FeA(G), CafAG, or CouAG were found, suggesting that these peaks were derivatives of lignan macromolecule.

Based on the peaks identified and the difference in m/z-ratios between the groups, it was shown that lignan macromolecule is

a collection of molecules of various sizes and compositions. Molecules ranging from 2SDG+1HMGA (M_r = 1.5 kDa) up to at least 5SDG + 6HMGA (M_r = 4.3 kDa) were identified by MALDI-TOF MS as being part of lignan macromolecule. It still remains to be identified what the maximum molecular mass is because the MALDI-TOF MS signal was declining with increasing molecular mass.

2.3. Subunit composition of fractions of lignan macromolecule

To investigate the variation in composition and size in more detail, lignan macromolecule was fractionated by gel permeation chromatography (GPC) as indicated in Fig. 4. The total lignan macromolecule eluted as one broad peak, in accordance with previous results (Shukla et al., 2002). Reinjection of the collected fractions on the same GPC column (insert in Fig. 4) showed that fractions were obtained, which differed in molecular mass.

The composition of the collected GPC-fractions was determined by analyzing the fractions on analytical RP-HPLC after full saponification (Fig. 5). In Fig. 6 the molar proportions of the four main phenolic constituents per GPC-fraction are shown. In the first five fractions, SDG was the predominant constituent, whereas in the fraction 7 and 8 CouAG + FeAG were predominating. With increasing retention times on GPC (thus a decrease molecular weight) the relative molar proportion of SDG decreased from 73 to 27%, whereas the proportion of CouAG increased from 12 to 52%. FeAG also showed an increase in proportion (from 10 to 18%), but this increase was smaller than that for CouAG. These changes in ratios between SDG, CouAG and FeAG in various fractions of lignan macromolecule are in agreement with literature results (Strandas et al., 2008).

The increase in proportions of CouAG + FeAG provided additional information about the structure of the molecules of the lignan macromolecule. When CouAG and FeAG would have been randomly distributed over a molecule, an increase in chain length is not expected to result in a change in ratio of SDG:(CouAG + FeAG). Since an inverse relation between molecular mass and proportion of CouAG + FeAG was observed, it is suggested that CouAG and FeAG were the terminating units. The molar ratio of SDG:(CouAG + FeAG) of 1:2 found in the fraction with the lowest molecular mass (fraction 8), also points in the direction of the presence of CouAG or FeAG on each terminal position of SDG. Also previous research showed indications for such a structure (Struijs et al., 2008).

These data suggest that chain elongation of lignan macromolecule stops after incorporation of a CouAG or FeAG moeity. It is hypothesized that the chain length of the molecules of lignan macromolecule is controlled by the concentrations free CouAG and FeAG present during biosynthesis. In early stages of flaxseed development, high concentrations of free CouAG and FeAG are detected, while in later stages, their concentration was lower (Hano et al., 2006). In addition, it has been observed that higher oligomers of SDG and HMGA are formed in later stages of seed development (Ford et al., 2001). These observations suggest that in early stages of seed development molecules with short chain lengths are terminated by the, in those stages, highly abundant CouAG and FeAG moieties. In later stages, when CouAG and FeAG are less abundant, it is suggested that longer chains are biosynthesized.

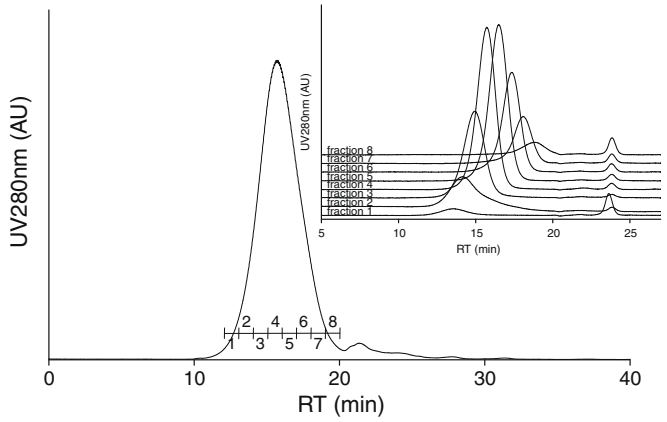

Fig. 4. GPC profile of lignan macromolecule. The collected fractions are indicated. In the insert, the GPC profiles of the collected fractions are shown reinjected on the same column.

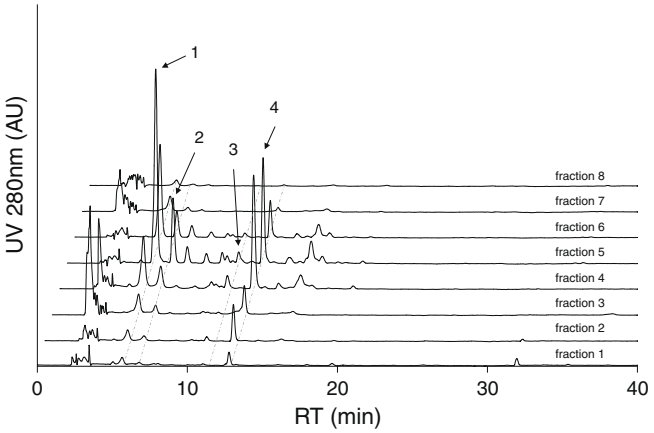

Fig. 5. RP-HPLC profiles of the saponified fractions of lignan macromolecule collected after GPC. 1 = CouAG, 2 = FeAG, 3 = HDG, 4 = SDG.

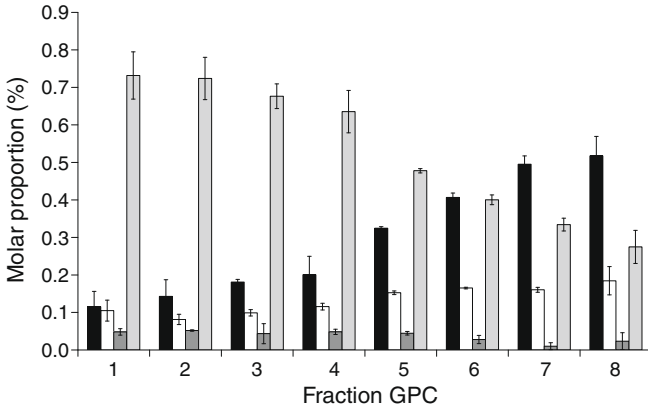

Fig. 6. Relative molar proportions of CouAG (black), FeAG (white), HDG (dark grey) and SDG (grey) in the fractions collected by GPC after saponification.

2.4. Schematic representation of lignan macromolecule

Taking all results into account, lignan macromolecule can be defined as a collection of molecules with various molecular masses, which are composed of one or more SDG moieties each linked to none, one or two HMGA moieties, and to which CouAG or FeAG residues can be linked. A schematic representation of lignan macromolecule is shown in Fig. 7a. SDG (or HDG), CouAG, FeAG (CouAG and FeAG are indicated in Fig. 7a as HCAG), and HMGA are the building blocks of lignan macromolecule, represented in Fig. 7a by a circle, a square and a parallelogram, respectively. SDG (or HDG) moieties link to each other via HMGA (Struijs et al., 2007), while CouAG and FeAG link directly to SDG (Struijs

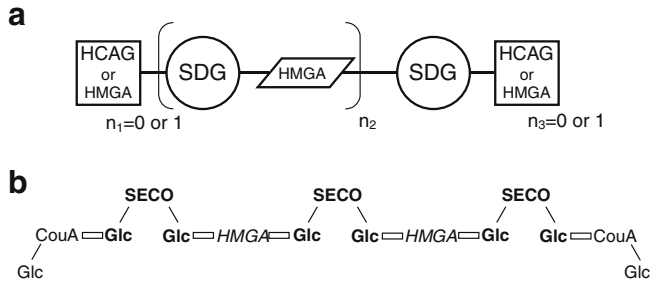

Fig. 7. (a) Schematic representation of lignan macromolecule from flaxseed hulls. The circles represent the backbone moieties. The backbone consists mainly of SDG. One out of eleven backbone units is HDG. The parallelogram represents the linker molecule HMGA and the squares represent the terminal units, which can be a CouAG, FeAG (HCAG = CouAG or FeAG) or HMGA. $n_1 + n_2 + n_3 \geqslant 1$. Modeled values for n_2+1 can be found in Table 4. (b) An example of a representative of the average lignan macromolecule is given, in which $n_1 = n_3 = 1$ and HCAG = CouAG, $n_2 = 2$, / = glycosidic bond, and ▭ = ester bond.

et al., 2008). Also FeA is a constituent of lignan macromolecule (Struijs et al., 2008), but present in such low amounts that it is not accounted for.

The schematic representation should comply with the structures assigned after MALDI-TOF MS analysis (Table 3: 2SDG + HMGA; $n_1 = n_3 = 0$, $n_2 = 1$ up to 5SDG + 5HMGA; $n_2 = 4$, n_1 or $n_3 = 1$). Based on these MALDI-TOF MS data, it was expected that a combination of SDG and HMGA is the basic unit of the molecules of lignan macromolecule. However, also the composition of the GPC-fractions should be taken into account. The ratio of 1:2 between SDG and CouAG (+FeAG) showed that CouAG–SDG–CouAG is a possible member of lignan macromolecule, which led to the suggestion that CouAG and/or FeAG are present at the terminal positions (as discussed above). Therefore, the representation (Fig. 7a) should also reflect the possibility of CouAG and FeAG at the terminal positions.

By definition, the smallest molecules being part of lignan macromolecule are SDG + HMGA or SDG + CouAG (or SDG + FeAG). For these molecules the following is true: $n_2 = 0$ and n_1 or $n_3 = 1$. For larger molecules, CouAG, FeAG or HMGA can or can not be present at the terminal position so the value for n_1 and n_3 can be either 0 or 1. The value for n_2 will vary with the chain length.

2.5. Average chain length of lignan macromolecule

Experimentally, the chain length could not be measured and the value of n_2 remained unidentified. To be able to estimate n_2, a model was developed which predicts chain length based on the compositional (current study) and structural data (Struijs et al., 2008).

A lignan macromolecular chain is described as a combination of the different building blocks:

$$\mathrm{LM} = (n_2 + 1)\mathrm{SDG} + (n_2 + a)\mathrm{HMGA} + b(\mathrm{CouAG} + \mathrm{FeAG}) \quad (1)$$

in which LM = the total lignan macromolecule, (n_2+1) = number of SDG units, a = number of terminal units HMGA, b = number of terminal units CouAG + FeAG. The following constraints are active: $n_2 \geqslant 0$, $a \geqslant 0$, $b \geqslant 0$, $a + b \geqslant 2$. In Eq. (1), $a + b$ are the total terminal units of HMGA plus CouAG + FeAG. In Fig. 7a, they are represented by $n_1 + n_3$.

Assuming that the molar proportion of SDG (y_{SDG}), the molar proportion of CouAG + FeAG ($y_{(CouAG+FeAG)}$), and the molar proportion of HMGA (y_{HMGA}) are forming the total lignan macromolecule, the following is true:

$$y_{SDG} + y_{HMGA} + y_{(CouAG+FeAG)} = 1$$

in which,

$$\begin{aligned} y_{SDG} &= (n_2 + 1)/((n_2 + 1) + (n_2 + a) + b) \\ y_{HMGA} &= (n_2 + a)/((n_2 + 1) + (n_2 + a) + b) \\ y_{(CouAG+FeAG)} &= b/((n_2 + 1) + (n_2 + a) + b) \end{aligned} \quad (2)$$

The following equations describe the molar proportions of SDG, HMGA and CouAG + FeAG in lignan macromolecule as a set of linear equations.

$$\begin{aligned} y_{SDG} &= (n_2 + 1)(2y_{SDG} - 1) + a y_{SDG} + b y_{SDG} \\ y_{HMGA} - 1 &= (n_2 + 1)(2y_{HMGA} - 1) + a(y_{HMGA} - 1) + b y_{HMGA} \\ y_{(CouAG+FeAG)} &= 2(n_2 + 1)y_{(CouAG+FeAG)} + a y_{(CouAG+FeAG)} \\ &\quad + b(y_{(CouAG+FeAG)} - 1) \end{aligned} \quad (3)$$

For the average (or total) lignan macromolecule, the molar proportions of SDG, HMGA, and CouAG + FeAG were recalculated from Table 2: $y_{SDG} = 0.433$, $y_{HMGA} = 0.294$ and $y_{(CouAG+FeAG)} = 0.273$. By trying to solve this set of equations, they appeared to be dependent. This dependency is most likely explained by the relationship between the amounts of SDG and HMGA. Solving the solutions resulted in a continuum of solutions only restricted by the constraints. In Fig. 8 the possible solutions are plotted. According to Fig. 8, the average lignan macromolecule can be described as follows:

$$\begin{aligned} &1 \leqslant n_2 + 1 \leqslant 3.1 \\ &0.6 \leqslant b \leqslant 2.0 \\ &0 \leqslant a \leqslant 0.7 \end{aligned}$$

The average chain length was modeled as being maximally 3 SDG moieties ($n_2 = 2.1$) with a CouAG or FeAG moiety on most terminal positions. This average chain length is lower than the chain length of 5 SDG units determined by Kamal-Eldin et al. (2001). In that study, CouAG or FeAG were not considered as being terminal units.

2.6. Chain length distribution of lignan macromolecule

Information about the chain length distribution can be obtained from the fractions obtained by GPC. In Table 4, the molar proportions (%) of SDG and CouAG + FeAG per GPC-fraction are given. Since no data about the content of HMGA in the fractions are available, lignan macromolecule should now be described by a combination of only SDG and CouAG + FeAG.

$$\mathrm{LM} = (n_2 + 1)\mathrm{SDG} + b(\mathrm{CouAG} + \mathrm{FeAG}) \quad (1b)$$

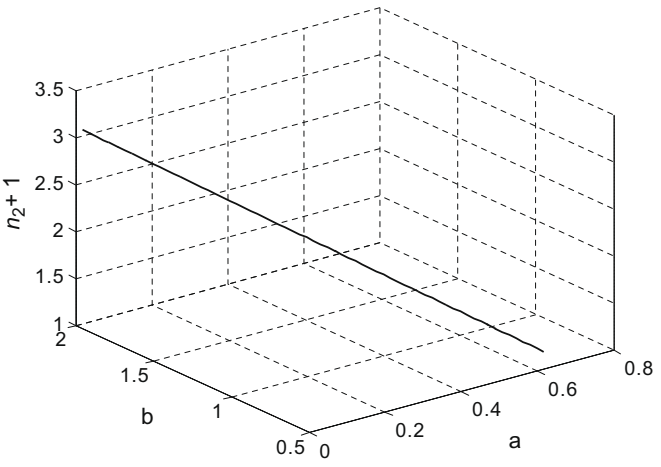

Fig. 8. Plot describing the average lignan macromolecule in terms of n_2+1 (amount of units SDG), b (amount of terminal CouAG + FeAG units), and a (amount of terminal HMGA units).

Table 4
Characteristics of the GPC-fractions of lignan macromolecule.

	GPC fraction								Weighted average
	1	2	3	4	5	6	7	8	
Peak area (%)	2	5	16	27	26	15	7	2	100
SDG (mol%)[a]	78	77	72	68	53	47	34	31	59
CouAG + FeAG (mol%)[a]	22	23	28	32	47	53	66	69	41
n_2+1[b]	1–7.1	1–6.7	1–5.1	1–4.3	1–2.3	1–1.8	1	1	1–2.9
b[c]	0.3–2	0.3–2	0.4–2	0.5–2	0.9–2	1.1–2	2	2	0.7–2

[a] Molar fractions of SDG and CouAG + FeAG are derived from Fig. 6 and are used to calculate n_2+1 and b by solving Eq. (3b). SDG represents the molar fraction (mol%) of SDG plus HDG. HMGA is not accounted for.
[b] $n_2 + 1$ is the amount of SDG (plus HDG) units per macromolecule.
[c] b is the amount (CouAG+FeAG) at the terminal positions.

and

$$z_{SDG} + z_{(CouAG+FeAG)} = 1 \quad (2b)$$

in which (n_2+1) = number of units SDG, b = number of terminal CouAG + FeAG moeities, and z are the molar ratios of SDG and CouAG + FeAG. The value for n_2 should be $\geqslant 0$, and $0 \leqslant b \leqslant 2$. z_{SDG} and $z_{(CouAG+FeAG)}$ can be rewritten as follows by combining Eqs. (1b) and (2b):

$$\begin{aligned} z_{SDG} &= (n_2+1)/((n_2+1)+b) \\ z_{(CouAG+FeAG)} &= b/((n_2+1)+b) \end{aligned} \quad (3b)$$

or combined:

$$(n_2+1) = z_{SDG}/z(CouAG + FeAG)^{*}b \quad (4b)$$

Also these equations were found to be dependent, giving a series of solutions limited by the constraints. In Table 4, the (dependent) ranges for n_2 (units of SDG) and b (amount of terminal CouAG + FeAG units) per GPC-fraction are given. Also the weighted average proportions of SDG and CouAG + FeAG (= (peak area * mol% per fraction)/(total area * mol% of all fractions)) were calculated and used to model the weighted average values for n_2 and b. The weighted average composition of all GPC-fractions resulted in a chain length of maximally 3 SDG units (n_2 = 1.9). This corresponds very well to the chain length found for the average lignan macromolecule.

For fractions 7 and 8, the Eq. (3b) can only be solved when b = 2. For the other fractions, b (and n_2) can vary. The calculated maximal chain length of lignan macromolecule from flaxseed hulls was found to be 7 SDG (or HDG; n_2 = 6.1) moieties with two CouAG or FeAG residues as terminal units (b = 2), but longer chain length can not be excluded. These data fit well with the maximum found chain length of at least 5 SDG units as identified by MALDI-TOF MS.

3. Conclusions

The average lignan macromolecule consists of three backbone units (SDG or HDG; n_2 = 2). The chain length varies between one and seven SDG (or HDG) moieties ($0 \leqslant n_2 \leqslant 6$) per molecule. An example of an average lignan macromolecule is shown in Fig. 7b.

The results obtained indicate that CouAG and FeAG are most likely present at the terminal positions. The amount of CouAG and FeAG seem to determine the chain length of lignan macromolecule during biosynthesis.

4. Experimental

4.1. Extraction lignan macromolecule from flaxseed hulls

Lignan macromolecule was extracted from flaxseed hulls as described previously (Struijs et al., 2007). In short, 400 g of flaxseed hulls, kindly provided by Frutarom Netherlands B.V. (Veenendaal, The Netherlands), were extracted with 3 l of *n*-hexane. Lignan macromolecule was extracted from the defatted hulls by a three-step sequential extraction with 63% (v/v) aq. EtOH at room temperature under continuous stirring. The extracts and the hulls were separated by filtration on a 595 round paper filter (Schleicher&Schuell). From the combined extracts the EtOH was evaporated and the concentrated extract was lyophilized yielding lignan macromolecule.

4.2. Saponification of lignan macromolecule

To release monomeric constituents, solutions of (fractions of) lignan macromolecule were fully saponified in 75 mM NaOH. The incubation volume was 1 ml, the concentration lignan macromolecule was 2 mg/ml or less. Saponification was performed at room temperature while stirring. After 24 h, the reaction was stopped by lowering the pH to 6.5–7.0 with HOAc.

4.3. Sample clean up of (saponified) lignan macromolecule

Prior to analysis, low molecular weight polar material was removed from lignan macromolecule extract or saponified samples by solid phase extraction (SPE; SepPak Vac, 20 cc/5 g, C18 cartridge, Waters). Prior to loading samples onto the SPE cartridge, the EtOH concentration in the reaction mixture was reduced by diluting the samples with 25 ml water.

After activation of the SepPak cartridge first with MeOH and followed by water, the diluted saponified lignan macromolecule solution was loaded onto the cartridge. After a wash step with 1 column volume of water, lignans were eluted from the cartridge with 3 column volumes of MeOH of which the first column volume was discarded. The MeOH eluate was evaporated under a stream of air to a final volume of 1 ml.

4.4. Reversed phase HPLC coupled on-line to mass spectrometry (RP-HPLC-MS)

Samples were analyzed on an X-Terra C18 MS column (Waters; 3.5 μm particle size, 4.6 × 150 mm) with an X-Terra C18 MS guard column (Waters; 3.5 μm particle size, 4.6 × 10 mm) run on a Thermo Separation Products HPLC system as previously described (Struijs et al., 2007). The gradient (Struijs et al., 2007) was run at 0.7 ml/min at room temperature. Samples of 20 μl were injected being dissolved in MeOH.

The molecular masses of the constituents of the macromolecule were determined on a Thermo Finnigan LCQ Classic coupled on-line to the analytical RP-HPLC. 10% of the flow from the analytical RP-HPLC was directed to the MS. The MS was equipped with an ESI injector. Spectra were obtained in the negative ion mode over an m/z range of 150–2000. The capillary temperature was 270 °C,

the ion spray voltage was set on 4.50 kV, the capillary voltage was −7.00 V, and helium was used as sheath gas. MS/MS analysis was performed with a normalized collision energy of 27%.

4.5. Gel permeation chromatography (GPC)

To obtain fractions of lignan macromolecule with different molecular masses, aliquots of 100 µl containing 2 mg lignan macromolecule (cleaned over SPE)/ml 63% (v/v) EtOH were injected on a similar HPLC system as described for RP-HPLC, equipped with a Tricorn Superdex Peptide 10/300 GL column (Amersham Bioscience; 10 × 300–310 mm, bed volume = 24 ml, optimum separation range 100–7000 Da). Fractions (0.8 ml) were collected during three subsequent runs. The column was run isocratically at a flow rate of 0.8 ml/min with 40% (v/v) aq. ACN + 0.1% (v/v) TFA as eluent. The eluate was monitored at 280 nm. The included volume (V_{inc} = 16.5 ml) was determined with water. No further calibration of the Superdex Peptide column was performed since no suitable calibration compounds were available.

The same procedure was used to determine the molecular weight distribution of the GPC-fractions isolated from the parental lignan macromolceule. The injection volume was 50 µl.

4.6. Determination molar extinction coefficients of CouAG, FeAG, HDG and SDG

CouAG, FeAG, HDG and SDG were purified by preparative RP-HPLC as described previously (Struijs et al., 2007). Based on stock-solutions of about 10 mg/ml, dilution series in MeOH were made. The absorbance at 280 nm of these dilutions was measured against MeOH in a 1 ml quartz cuvet. The molar extinction coefficients (ϵ) were calculated using Abs = $\epsilon^{*} l^{*} c$ in which Abs = absorbance at 280 nm, l = light path = 1 cm, c = concentration (M).

4.7. Determination HMGA by GC-MS

Lignan macromolecule (2 mg/ml in EtOH) was fully saponified. Aliqouts (0.5 ml) were spiked with 50 µg glutaric acid (GA) dissolved in water (50 µl, 1 mg/ml) as internal standard. A calibration curve was made with a concentration range of HMGA (Aldrich) starting from a 1 mg/ml stock-solution. The volume was increased to 5 ml with water and the pH was adjusted to pH 1.5–2.0 with HCl. The samples were evaporated to dryness under a stream of air at 50 °C. To remove residual water, 100 µl acetone was added and evaporated. Reaction vials were purged with N_2.

To obtain TMSi derivatives, to each dried sample 1 ml of HMDS:TMCS:pyridine (3:1:9; SIL-PREP Kit, Alltech) was added. After 2 h the derivatisation reaction was complete and the derivatized samples were analyzed on a GC 8000TOP system with an AS800 autosampler and FID-detection (Thermo Finnigan) equipped with a DB-1 column (J&W; 30 × 0.25 × 1.0). Injection volume was 1 µl (splitless injection). The temperature program was a follows: t = 0 min, 80 °C; t = 0–12 min, 80–325 °C (linear increase); t = 10–42 min, 325 °C. Helium was used as a carrier gas at 100 kPa. The FID-detector was set on 350 °C.

GC-MS data were obtained by analyzing the derivatized samples on a Trace GC (Thermo Finnigan) coupled to a Polaris Q mass spectrophotometer. The settings of the GC were similar as describe above. MS data were collected by electron impact, in full scan mode between 5 and 25 min of the GC run, in the positive ion mode over a mass range of 50–400 Da. The temperature of the ion source was 250 °C. The auto-tune function was used to optimize the MS settings.

4.8. Calculation of the molar proportions and the weight percentages

The molar proportions of the four phenolic compounds were calculated by dividing the molar equivalents of the individual compounds by the total molar equivalents. The molar equivalents were determined by multiplying the molar extinction coefficients (Table 1) by the peak area under the RP-HPLC profiles of fully saponified lignan macromolecule (chromatograms not shown). The weight percentages were calculated based on the measured value for HMGA and the molar proportions of the phenolic constituents.

4.9. MALDI-TOF MS of lignan macromolecule

For the determination of the molecular mass distribution, unsaponified lignan macromolecule was analyzed on a Ultraflex MALDI-TOF MS (Bruker Daltronics GmbH). As matrix 2,5-dihydroxybenzoic acid dissolved in 50% aq. ACN (15 mg/ml) was used. Two µl of sample mixture consisting of 10 µl 2 mg/ml unsaponified lignan macromolecule cleaned over SPE, 10 µl matrix-solution, and 1 µl 1 mM NaOAc pH 5.0 was spotted on a gold plate. The MALDI-TOF MS was calibrated with a mixture of maltodextrins (mass range 365–4092 Da). The system was used in the positive reflector mode.

References

Boerjan, W., Ralph, J., Baucher, M., 2003. Lignin biosynthesis. Annu. Rev. Plant Biol. 54, 519–546.

Dabrowski, K.J., Sosulski, F.W., 1984. Composition of free and hydrolyzable phenolic acids in defatted flours of ten oilseeds. J. Agric. Food Chem. 32, 128–130.

Davin, L.B., Wang, H.B., Crowell, A.L., Bedgar, D.L., Martin, D.M., Sarkanen, S., Lewis, N.G., 1997. Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. Science 275, 362–366.

Eklund, P.C., Langvik, O.K., Warna, J.P., Salmi, T.O., Willfor, S.M., Sjoholm, R.E., 2005. Chemical studies on antioxidant mechanisms and free radical scavenging properties of lignans. Org. Biomol. Chem. 3, 3336–3347.

Eliasson, C., Kamal-Eldin, A., Andersson, R., Aman, P., 2003. High-performance liquid chromatographic analysis of secoisolariciresinol diglucoside and hydroxycinnamic acid glucosides in flaxseed by alkaline extraction. J. Chromatogr. A 1012, 151–159.

Ford, J.D., Huang, K.S., Wang, H.B., Davin, L.B., Lewis, N.G., 2001. Biosynthetic pathway to the cancer chemopreventive secoisolariciresinol diglucoside-hydroxymethyl glutaryl ester-linked lignan oligomers in flax (*Linum usitatissimum*) seed. J. Nat. Prod. 64, 1388–1397.

Hano, C., Martin, I., Fliniaux, O., Legrand, B., Gutierrez, L., Arroo, R.R.J., Mesnard, F., Lamblin, F., Laine, E., 2006. Pinoresinol–lariciresinol reductase gene expression and secoisolariciresinol diglucoside accumulation in developing flax (*Linum usitatissimum*) seeds. Planta 224, 1291–1301.

Johnsson, P., Kamal-Eldin, A., Lundgren, L.N., Aman, P., 2000. HPLC method for analysis of secoisolariciresinol diglucoside in flaxseeds. J. Agric. Food Chem. 48, 5216–5219.

Johnsson, P., Peerlkamp, N., Kamal-Eldin, A., Andersson, R.E., Andersson, R., Lundgren, L.N., Aman, P., 2002. Polymeric fractions containing phenol glucosides in flaxseed. Food Chem. 76, 207–212.

Kamal-Eldin, A., Peerlkamp, N., Johnsson, P., Andersson, R., Andersson, R.E., Lundgren, L.N., Aman, P., 2001. An oligomer from flaxseed composed of secoisolariciresinoldiglucoside and 3-hydroxy-3-methyl glutaric acid residues. Phytochemistry 58, 587–590.

Klosterman, H.J., Smith, F., 1954. The isolation of beta-hydroxy-beta-methylglutaric acid from the seed of flax (*Linum-usitatissimum*). J. Am. Chem. Soc. 76, 1229–1230.

Liggins, J., Grimwood, R., Bingham, S.A., 2000. Extraction and quantification of lignan phytoestrogens in food and human samples. Anal. Biochem. 287, 102–109.

Madhusudhan, B., Wiesenborn, D., Schwarz, J., Tostenson, K., Gillespie, J., 2000. A dry mechanical method for concentrating the lignan secoisolariciresinol diglucoside in flaxseed. Food. Sci. Technol. 33, 268–275.

Meagher, L.P., Beecher, G.R., Flanagan, V.P., Li, B.W., 1999. Isolation and characterization of the lignans, isolariciresinol and pinoresinol, in flaxseed meal. J. Agric. Food Chem. 47, 3173–3180.

Moss, G.P., 2000. Nomenclature of lignans, neolignans (IUPAC Recommendations 2000). Pure Appl. Chem. 72, 1493–1523.

Prasad, K., 2004. Lignan complex derived from flaxseed as hypercholesterolemic and anti-atherosclerotic agent. International Patent US 6,673,773 B2.

Qiu, S.-X., Lu, Z.-Z., Luyengi, L., Lee, S.K., Pezzuto, J.M., Farnsworth, N.R., Thompson, L.U., Fong, H.H.S., 1999. Isolation and characterization of flaxseed (*Linum usitatissimum*) constituents. Pharm. Biol. 37, 1–7.

Shukla, R., Hilaly, A., Moore, K.M., 2002. Process for obtaining lignan, pharmaceutical compositions and uses thereof. International Patent WO 02/077003.

Sicilia, T., Niemeyer, H.B., Honig, D.M., Metzler, M., 2003. Identification and stereochemical characterization of lignans in flaxseed and pumpkin seeds. J. Agric. Food Chem. 51, 1181–1188.

Strandas, C., Kamal-Eldin, A., Andersson, R., Aman, P., 2008. Composition and properties of flaxseed phenolic oligomers. Food Chem. 110, 106–112.

Struijs, K., Vincken, J.-P., Verhoef, R., n Casteren, W.H.M., Voragen, A.G.J., Gruppen, H., 2007. The flavonoid herbacetin diglucoside as a constituent of the lignan macromolecule from flaxseed hulls. Phytochemistry 68, 1227–1235.

Struijs, K., Vincken, J.-P., Verhoef, R., Voragen, A.G.J., Gruppen, H., 2008. Hydroxycinnamic acids are ester-linked directly to glucosyl moeities within the lignan macromolecule from flaxseed hulls. Phytochemistry 69, 1250–1260.

von Heimendahl, C.B.I., Schafer, K.M., Eklund, P., Sjoholm, R., Schmidt, T.J., Fuss, E., 2005. Pinoresinol–lariciresinol reductases with different stereospecificity from *Linum album* and *Linum usitatissimum*. Phytochemistry 66, 1254–1263.

Wang, L.Q., 2002. Mammalian phytoestrogens: enterodiol and enterolactone. J. Chromatogr. B 777, 289–309.

Westcott, N.D., Muir, A.D., 1996. Process for extracting and purifying lignans and cinnamic acid derivatives from flaxseed. International Patent WO 9630468.

Youn, B.Y., Moinuddin, S.G.A., Davin, L.B., Lewis, N.G., Kang, C.H., 2005. Crystal structures of apo-form and binary/ternary complexes of *Podophyllum* secoisolariciresinol dehydrogenase, an enzyme involved in formation of health-protecting and plant defense lignans. J. Biol. Chem. 280, 12917–12926.