Available online at www.sciencedirect.com
ScienceDirect
Phytochemistry 69 (2008) 1250–1260
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Hydroxycinnamic acids are ester-linked directly to glucosyl moieties within the lignan macromolecule from flaxseed hulls

Karin Struijs, Jean-Paul Vincken, René Verhoef, Alphons G.J. Voragen, Harry Gruppen *

Laboratory of Food Chemistry, Wageningen University, P.O. Box 8129, 6700 EV Wageningen, The Netherlands

Received 18 June 2007; received in revised form 13 November 2007
Available online 9 January 2008

Abstract

In flaxseed hulls, lignans are present in an oligomeric structure. Secoisolariciresinol diglucoside (SDG), ester-linked to hydroxy-methyl-glutaric acid (HMGA), forms the backbone of this lignan macromolecule. The hydroxycinnamic acids *p*-coumaric acid glucoside (CouAG) and ferulic acid glucoside (FeAG) are also part of the lignan macromolecule. However, their position and type of linkage are still unknown. The aim of this study was to investigate how CouAG and FeAG are linked within the lignan macromolecule from flaxseed hulls.

Fragments of the lignan macromolecule were obtained by partial saponification. After isolation of the fragments by preparative RP-HPLC, several key structures were identified by MS and NMR.

Within the lignan macromolecule, CouAG is attached to the C-6 position of a glucosyl moiety of SDG. FeA is linked to the C-2 position of a glucosyl moiety of SDG. FeAG is ester-linked within the lignan macromolecule with its carboxyl group, but it remains unclear whether FeAG links to the C-2 or C-6 position of SDG. Attachment of HMGA to the glucosyl moiety of CouAG or FeAG was not observed. The results clearly show that within the lignan macromolecule, the hydroxycinnamic acids are linked directly via an ester bond to the glucosyl moiety of SDG.

Keywords: Flaxseed hull; Lignan; Lignan macromolecule; Secoisolariciresinol diglucoside (SDG); *p*-Coumaric acid glucoside (CouAG); Ferulic acid glucoside (FeAG); Hydroxy-methyl-glutaric acid (HMGA)

1. Introduction

Lignans are phenolic compounds, which are widely distributed in plants. Especially in seeds and nuts, high concentrations can be found (Milder et al., 2005). Concentrations up to 3% (w/w) have been reported in flaxseeds, making flax one of the richest edible sources of lignans (Eliasson et al., 2003; Milder et al., 2005). Secoisolariciresinol diglucoside (SDG) is the most important lignan in flaxseed (Milder et al., 2005). After ingestion, SDG is converted into the mammalian lignans enterodiol and enterolacton, which exhibit several health beneficial effects (Thompson et al., 1996; Vanharanta et al., 1999; Ward et al., 2001).

In contrast to most plants in which free lignans are present, the lignans in flaxseeds are incorporated into an oligomeric structure (Kamal-Eldin et al., 2001; Westcott and Muir, 1996), which is referred to as lignan macromolecule. This lignan macromolecule is also reported in flaxseed hulls (Struijs et al., 2007).

Although a variety of lignans (Bakke and Klosterman, 1956; Liggins et al., 2000; Meagher et al., 1999; Sicilia et al., 2003), flavonoids (Qiu et al., 1999), and (hydroxy)cinnamic acids (Dabrowski and Sosulski, 1984; Klosterman et al., 1955; Westcott and Muir, 1996) have been identified in flaxseed extracts, only a small number of these constituents are related to the lignan macromolecule. The most abundant lignan in flaxseeds, secoisolariciresinol

* Corresponding author. Tel.: +31 317 482888; fax: +31 317 484893.
E-mail address: harry.gruppen@wur.nl (H. Gruppen).

0031-9422/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2007.11.010

diglucoside (SDG), is esterified to hydroxy-methyl-glutaric acid (HMGA), thereby forming the backbone of the lignan macromolecule (Kamal-Eldin et al., 2001). *p*-Coumaric acid glucoside (CouAG) and ferulic acid glucoside (FeAG) are thought to be part of the lignan macromolecule, since they are only observed after alkali treatment of flaxseed extracts containing lignan macromolecule (Johnsson et al., 2002; Struijs et al., 2007; Westcott and Muir, 1996). Also the release of caffeic acid glucoside from a similar extract has been reported (Westcott and Muir, 1996). Recently, it was shown that the flavonoid herbacetin diglucoside (HDG) is also part of the macromolecule (Struijs et al., 2007). In a similar way as SDG, HDG is attached within the lignan macromolecule via ester-linkages with HMGA.

Until now, it is unclear how the hydroxycinnamic acid glucosides are linked within the lignan macromolecule. In-line with the attachment of SDG and HDG, hydroxycinnamic acid glucosides might be linked within the lignan macromolecule via ester-linkage of their glucosyl moieties to HMGA. Another possibility is that they are ester-linked directly with their carboxyl group to a glucosyl moiety of SDG.

The aim of the present study is to identify the linkage of the hydroxycinnamic acid glucosides CouAG and FeAG within the lignan macromolecule from flaxseed hulls.

2. Results

For the isolation of structural elements of the lignan macromolecule from flaxseed hulls containing CouAG or FeAG, the procedure of partial saponification was followed as previously described (Struijs et al., 2007). It should be noted that during saponification, transesterification between ethanol and the lignan macromolecule can take place so that ethanolates can be formed (Ford et al., 2001; Johnsson et al., 2002).

Fragments of the lignan macromolecule formed during partial saponification were separated by preparative RP-HPLC. Fractions were collected as indicated by the marks in the RP-HPLC profile (Fig. 1). All fractions were analyzed on analytical RP-HPLC/MS. The MS data were screened for *m*/*z*-ratios corresponding to fragments composed of monomeric constituents of the lignan macromolecule being *p*-coumaric acid (CouA), CouAG, ferulic acid (FeA), FeAG, HDG, secoisolariciresinol (SECO), SDG, HMGA, their ethanolates and combinations thereof. Only those fractions with high signal to noise ratios on MS, high UV-signals, or annotations matching the previously mentioned compositional criteria, were studied further. On this basis, 11 fractions were selected as indicated by the numbers 1–11 in Fig. 1.

In Table 1 the *m*/*z*-ratios of the fractions 1–11 with their corresponding annotations are listed. Also the MS/MS data were supporting for these annotations. The number of compounds present in each fraction might be overestimated, as peaks were overlapping (Fig. 1) and fragmentation as a result of ionization during MS analysis could occur. Some fragments were solely annotated based on their *m*/*z*-ratios (no MS/MS data obtained), but the likeliness of the presence of these fragments allowed annotation as such.

In accordance with the identification of the monomeric constituents after full saponification (Struijs et al., 2007), fractions 1–4 were annotated as the monomeric constituents of the lignan macromolecule. In fractions 5–8, high intensity *m*/*z*-ratios corresponding to fragments consisting of SDG and HMGA were found. In fractions 6–11,

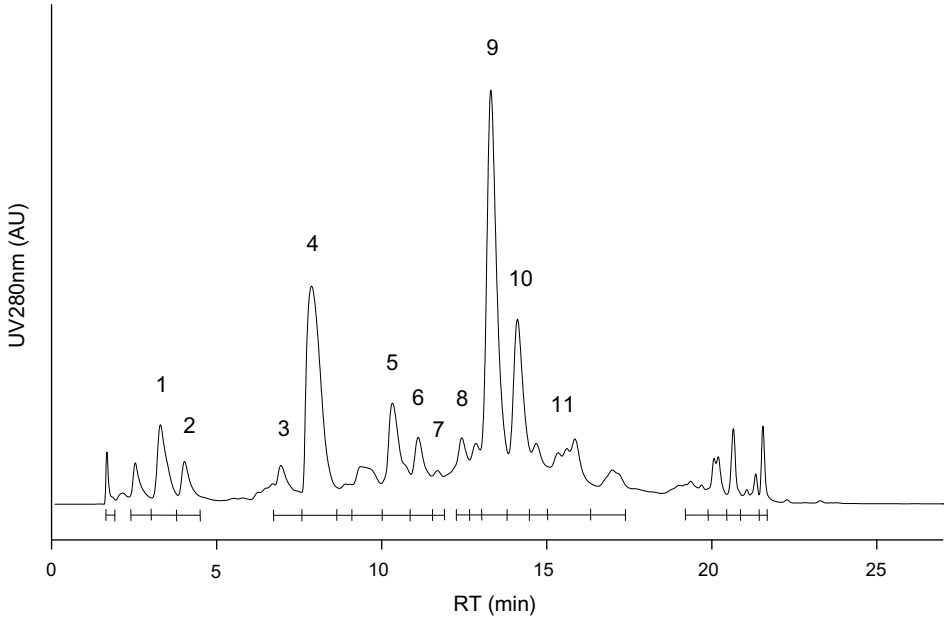


Fig. 1. Preparative RP-HPLC profile of partially saponified lignan macromolecule. The collected fractions are indicated. Annotations of the numbered peaks are listed in Table 1.

Table 1
Annotations of the fragments obtained by partial saponification of the lignan macromolecule as determined by RP-HPLC/MS (negative mode) and MS/MS detection

Fraction	Annotation	m/z $[M-H^+]^-$	MS/MS
1	**CouAG**	**325.0**	**163.1**; **204.8**
2	**FeAG**	**355.0**	**178.1**; **193.0**
3	**HDG**	**625.2**	**463.1**
4	**SDG**	**685.3**	**361.4**; **523.2**
	HDG + HMGA	769.2	463.1; 607.1; 625.2; 707.1
5	**SDG + HMGA**[ab]	**829.1**	**523.3**; **667.1**; **685.1**
6	**SDG + HMGA**[a]	**829.1**	**523.1**; **667.2**; **685.2**
	HDG + HMGA + EtOH[a]	797.2	463.1; 505.2; 625.2; 667.1
	SDG + CouA	831.2	523.3; 667.2; 686.2
	SDG + 2HMGA[a]	**973.2**	**685.2**; **829.2**; **871.1**
	SDG + CouAG	993.0	709.0
	SDG + HMGA + CouAG[a]	1137.1	ND
7	**SDG + 2HMGA**	**973.2**	**685.3**; **829.1**; **871.1**
	HDG + HMGA + EtOH	797.2	463.2; 607.2; 625.2; 779.0
	SDG + HMGA	829.2	667.2; 685.2; 727.0
	SDG + HMGA + CouAG	1137.2	974.8
	SDG + HMGA + FeAG	1167.2	861.1; 1004.9
	SDG + HMGA + HDG	1437.3	463.1; 625.2; 667.2; 1275.3
8	**SDG + 2HMGA**[a]	973.1	685.2; 829.1; 871.1
	SDG + FeA[ab]	861.1	343.3; 505.1; 667.1; 685.1
	SDG + HMGA + CouAG[a]	1137.2	831.2; 873.2; 913.2; 975.2
	SDG + HMGA + FeAG[a]	1167.2	861.2; 1005.1
	SDG + HMGA + HDG	1437.2	463.0; 667.1; 1275.1
	SDG + 2HMGA + HDG[a]	1581.0	ND
9	**CouAG + EtOH + acetate**[a]	**412.9**	**191.1**; **250.1**; **352.7**
	SDG + FeA + 520[ca]	**1363.1**	**861.2**; **1057.1**; **1219.1**
	HDG + HMGA + EtOH[a]	797.1	ND
	SDG + FeA[a]	861.2	523.2; 667.2; 685.2
	SDG + HMGA + CouAG	1137.1	614.6
	SDG + CouAG + FeA[a]	1169.0	831.2; 861.3; 1007.3
	SDG + 2HMGA + CouAG[a]	1281.3	831.1; 975.1; 1119.2; 1137.2
10	**SDG + HMGA + FeA**[a]	**1005.1**	**829.3**; **861.2**
	FeAG + EtOH + acetate[a]	442.9	221.1; 281.0; 382.5
	SDG + FeA[ab]	861.2	505.2; 667.1; 685.0
	SDG + HMGA + CouA[a]	975.1	667.1; 829.0; 831.2; 873.2
11	SDG + 2HMGA + EtOH	1001.1	685.4; 727.3; 973.3
	2SDG + HMGA	1497.1	ND
	2SDG + 2HMGA	1641.0	ND
	SDG + CouA	831.2	325.1; 523.2; 667.2; 685.2
	SDG + FeA	861.1	361.2; 505.3; 523.2; 699.2

Fractions 5, 6, 8, 9, and 10 were subjected to semi-preparative separation. Fragments indicated with [a] were found in both preparative and semi-preparative separations. MS/MS data were collected after preparative separation. MS/MS data collected after semi-preparative separation are marked with [b]. Bold numbers are the predominant m/z-ratios. ND = no MS/MS data available. [c]The m/z-ratio of 861.2 $[M-H^+]^-$ in the MS/MS spectrum of the fragment with an m/z-ratio of 1363.1 $[M-H^+]^-$ points at the presence of SDG + FeA in this fragment. The rest of the fragment remained unidentified.

fragments of the lignan macromolecule with m/z-ratios corresponding to SDG plus FeA(G) or CouA(G), and SDG + HMGA plus FeA(G) or CouA(G) were annotated. Especially these fragments were of interest for the identification of linkage types between SDG and CouA(G) or FeA(G). Therefore, they were analyzed further by NMR. A fragment of HDG + HMGA + EtOH, which shows the linkage of HDG within the lignan macromolecule, has been described before (Struijs et al., 2007).

Since the fractions putatively consisting of SDG and hydroxycinnamic acids were not base line separated, fractions 5, 6, 8, 9, and 10 were subjected to a semi-preparative purification step before NMR analysis. After semi-preparative purification, fractions 5*, 6*, 8*, 9*, and 10* were obtained, and analyzed on analytical RP-HPLC/MS. The abundance of fraction 7 was found to be too low and fraction 11 too complex for further purification.

The analytical RP-HPLC profiles of the five fractions obtained by semi-preparative RP-HPLC, showed that, except for some small shoulders, the peaks were pure based on HPLC-UV signals (data not shown). However, MS analysis revealed several *m/z*-values per fraction. In Table 1, the putative compounds, which were still present after semi-preparative purification, are indicated ([a]). It is remarkable that despite the semi-preparative purification, several *m/z*-ratios (e.g. $m/z = 829.1$ $[M-H^+]^-$ and $m/z = 1137.2$ $[M-H^+]^-$) were found back in several fractions. No further attempts to obtain more pure fractions were performed. The five fractions purified by semi-preparative RP-HPLC were analyzed by NMR to confirm the annotations made based on MS data and to identify the linkage types between the hydroxycinnamic acids and SDG. In Table 2 the structural elements identified by NMR are listed, including the numbering used throughout the text.

In fraction 5* based on the COSY and HMBC spectra, the structural element SDG + HMGA was identified. Chemical shifts corresponded to those of SDG and HMGA reported in the literature (Kamal-Eldin et al., 2001; Struijs et al., 2007). Besides, a cross-peak between C-α and H-6″a/b corresponded to the coupling of the carboxyl group of HMGA to one of the glucosyl units of SDG. Neither the C-ε nor the C-6‴ showed up- or downfield shifts of proton and carbon signals, nor a cross-peak indicating a linkage. This showed that the HMGA also carried a free carboxyl group (C-ε) and that the second glucosyl moiety of SDG had a free C-6 position (C-6‴). These data corresponded to the structural element of SDG + HMGA as shown in Table 2.

The HMBC spectrum of fraction 6* is shown in Fig. 2. Two structures were identified in this fraction. The first structure identified in the HMBC spectrum was HDG + HMGA + EtOH, as identified before (Struijs et al., 2007). Second, a structural element consisting of SDG + CouAG was found. The chemical shifts of this fragment are given in Table 3. The hydroxyl group of CouA linked via a glycosidic bond to glucose, showing the presence of CouAG (in Fig. 2 cross-peak indicated as CouA 4*, Glc 1‴″). The chemical shifts of this coupling were as reported in the literature (Johnsson et al., 2002; Struijs et al., 2007). The carboxyl group of CouAG (C-9*) coupled to the C-6″ of a glucosyl residue of SDG as shown by a cross-peak between C-9* and the downfield shifted H-6″a/b (in Fig. 2 cross-peak indicated as CouA 9, Glc 6″a/b). Cross-peaks indicating linkage of CouAG to another position of SDG were not found. These results showed that CouAG is linked with its carboxyl group to the C-6 position of a glucosyl moiety of SDG (see Table 2). Also in fraction 8* some low intensity cross-peaks indicating SDG + CouAG were present. No indications for linkage of CouA to another constituent were obtained by NMR. The annotation of fragments containing CouA based on MS analysis are most likely explained by occurrence of fragmentation during MS analysis.

Based on the proton-NMR spectra of fraction 9* it was confirmed that CouAG is linked within the macromolecule via its carboxyl group. This fraction consisted predominantly of CouAG + EtOH (see Table 2). The ethanol-ester was present on the carboxyl group (C-9) of CouAG as shown by the slight upfield shift of H-8. Besides, the chemical shifts of the ethanolate group were comparable to the ones found for FeAG + EtOH as described later. Furthermore, the integrals of all proton signals of this spin system fitted well with the proposed structure (see spectral data in Section 4).

A fragment annotated based on the MS data as SDG + HMGA + CouAG (Table 1, fraction 6–8), suggested that the structural element of SDG + CouAG, as identified by NMR, could be elongated by HMGA. However, the MS/MS data (Table 1) were not conclusive about which glucosyl moiety carried HMGA.

In fraction 8*, cross-peaks pointing at the presence of two main structural elements were identified: SDG + 2HMGA and SDG + FeA (see Fig. 3 for HMBC spectra, Table 3 for chemical shifts, and Table 2 for the chemical structures). For the structural element of SDG + 2HMGA, again chemical shifts corresponding to SDG and HMGA were observed. A cross-peak between C-α and H-6′a/b was present in the HMBC spectrum showing the coupling between the carboxyl group of HMGA and the C-6 of the glucosyl residue of SDG (in Fig. 3, cross-peak indicated as HMGAα, Glc 6′a/b). The structural element SDG + 2HMGA (see Table 2) was identified since its spin system showed only one glucose anomeric signal (Glc C-1′) with downfield shifted H-6′a/b protons, both showing cross-peaks in the HMBC spectrum to the C-α of HMGA. Additional proof for this symmetrical molecule was the slightly upfield shifted H-9a compared to a glucosyl residue of SDG which is not substituted with HMGA.

In fraction 8*, also chemical shifts corresponding to FeA were found (see Table 3). In the HMBC spectrum (Fig. 3) no coupling of glucose to the C-4* position of FeA was observed. The coupling of SDG to FeA was identified by a cross-peak between C-9* and H-2″ showing an ester-linkage between the carboxyl group of FeA and the C-2 position of a glucosyl residue of SDG (in Fig. 3, cross-peak indicated as FeA 9*, Glc 2″). The identification of a structural element with FeA was surprising, since in the RP-HPLC/MS data of the fully saponified lignan macromolecule only very low intensity *m/z*-values annotated as FeA + EtOH were found (data not shown).

Also in fraction 10*, some low intensity peaks were present indicating the presence of SDG + FeA with FeA linked to the C-2 position of glucose. In agreement with these weak NMR signals was the annotation of the fragment of SDG + HMGA + FeA based on MS results of fraction 10*. The MS/MS data showed that FeA or HMGA was split off, so that both FeA and HMGA formed a terminal group of this fragment. In the HMBC spectrum of fraction 10* (data not shown) also a cross-peak of HMGA coupled

Table 2
Structural elements of the lignan macromolecule identified by NMR

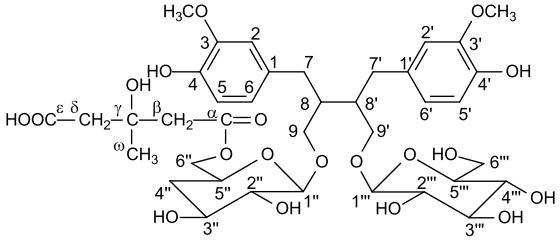
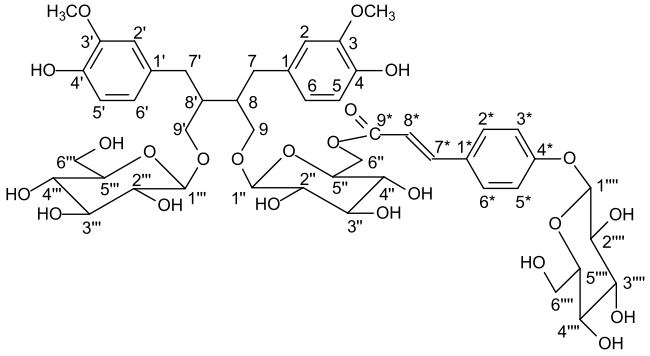
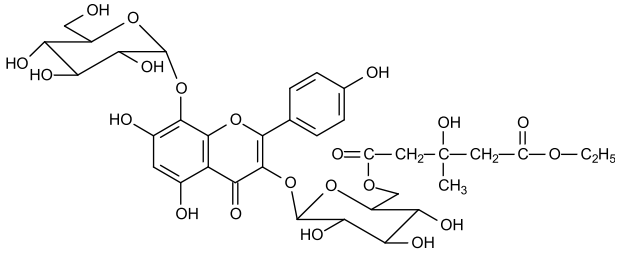
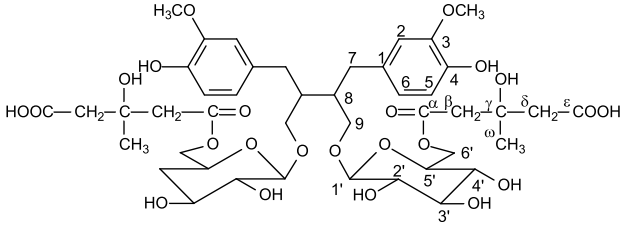
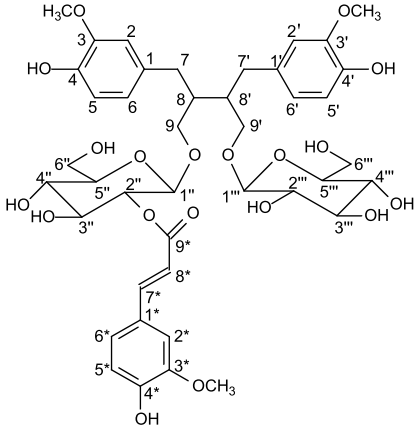
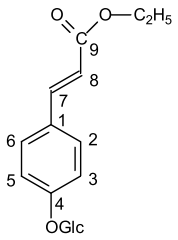
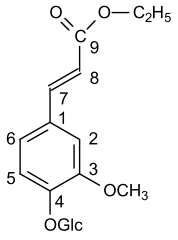
RP-HPLC fraction	Annotation	Chemical structure
5*	SDG + HMGA	
6*	SDG + CouAG	
6*	HDG + HMGA + EtOH	
8*	SDG + 2HMGA	
8*	SDG + FeA	

Table 2 (*continued*)

RP-HPLC fraction	Annotation	Chemical structure
9*	CouAG + EtOH	
10*	FeAG + EtOH	

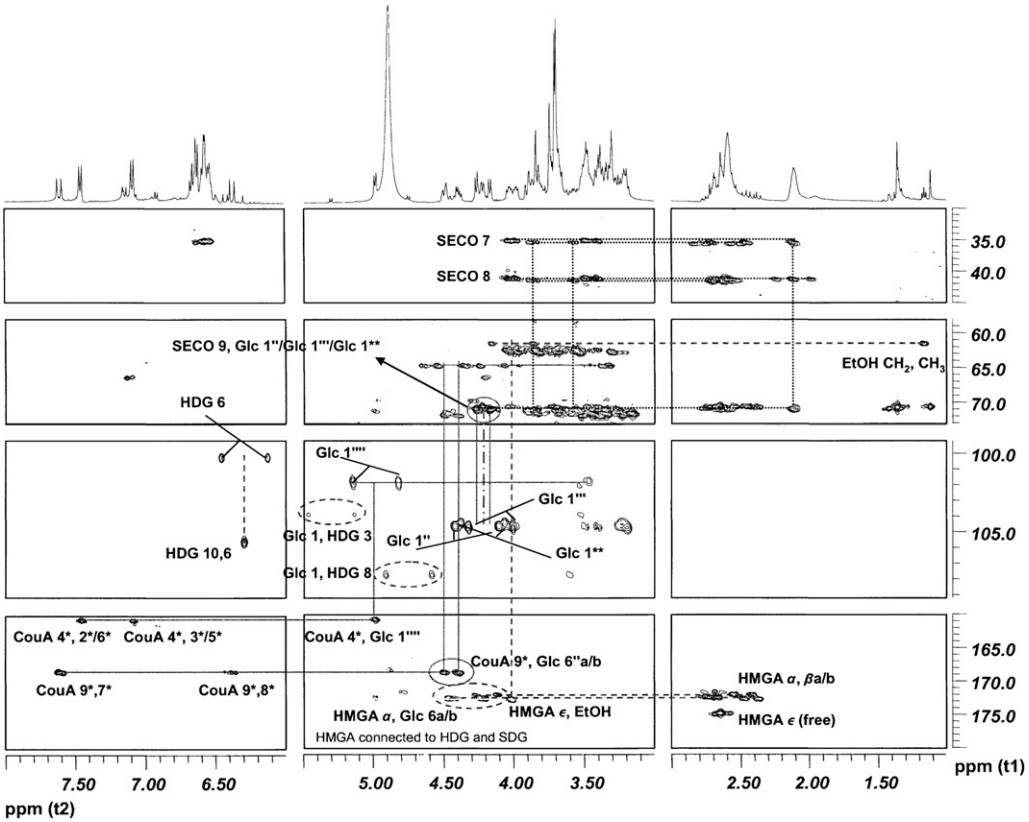


Fig. 2. HMBC spectrum of fraction 6* of the lignan macromolecule. Cross-peaks of SDG + CouAG (CouAG and linkage to glucose: (—); SDG: (·····)), and HDG + HMGA + EtOH (------) are indicated. Numbering corresponds to numbers given in Table 2. Cross-peaks are indicated as the coupling of carbon to proton. For example, the notation of CouA 9*, Glc 6″ indicates the coupling of the C-9* to the H-6″, the notation Glc‴ indicates a single bond coupling of C-1‴ to H-1‴ of Glc. Due to overlapping signals, it could not be determined to which compound Glc1** (-·-·-·-) is attached.

to the C-6 position of a glucosyl unit of SDG was identified. However, due to overlapping peaks, it was not possible to determine if other groups, e.g. FeA, were attached to the same glucosyl moiety. Combining the NMR and MS data resulted in two possible configurations of SDG + HMGA + FeA as schematically shown below:

Table 3
NMR chemical shifts of SDG + CouAG, SDG + 2HMGA and SDG + FeA

SDG + CouAG[a]				SDG + 2HMGA[a]				SDG + FeA[a]			
	^{1}H δ	mult., J (Hz)[b]	^{13}C δ		^{1}H δ	mult., J (Hz)[b]	^{13}C δ		^{1}H δ	mult., J (Hz)[b]	^{13}C δ
SECO				SECOa				SECOb			
1	–	–	133.812	1	–	–	133.81	1	–	–	133.81
2	6.58		113.593	2	6.580		113.37	2	6.486		113.37
3	–	–	148.813	3	–	–	148.59	3	–	–	148.59
4	–	–	145.117	4	–	–	145.11	4	–	–	145.11
5	6.63	(*d*) 7.75	115.767	5	6.616		115.76	5	6.604		115.76
6	6.54	(*d*) 7.29	122.941	6	6.474		122.72	6	6.474		122.72
7a	2.599		35.543	7a	2.616		35.12	7a	2.629		36.19
7b	2.599			7b	2.569			7b	2.522		
8	2.115		41.196	8	2.169		41.19	8	1.955		41.63
9a	3.983	(*dd*) 8.55, 5.02	71.198	9a	3.981		70.98	9a	3.957		71.41
9b	3.403			9b	3.475			9b	3.381		
OCH_3	3.734		56.415	OCH_3	3.700		56.19	OCH_3	3.700		56.19
1′	–	–	133.812	Glc				1′	–	–	133.81
2′	6.58		113.593	1′	4.245	(*d*) 7.78	104.46	2′	6.486		113.37
3′	–	–	148.813	2′	3.252		75.33	3′	–	–	148.59
4′	–	–	145.117	3′	3.322		77.94	4′	–	–	145.11
5′	6.63	(*d*) 7.75	115.767	4′	3.322		71.85	5′	6.604		115.76
6′	6.54	(*d*) 7.29	122.941	5′	3.369		75.33	6′	6.474		122.72
7′a	2.599		35.543	6′a	4.416		64.89	7′a	2.628		35.76
7′b	2.599			6′b	4.241			7′b	2.499		
8′	2.115		41.413					8′	1.955		41.63
9′a	4.027	(*dd*) 9.74, 5.69	70.981	HMGA				9′a	4.087		71.20
9′b	3.499			α	–	–	172.51	9′b	3.428		
OCH_3	3.734		56.415	βa	2.65		46.19	OCH_3	3.700		56.19
				βb	2.65						
Glc				γ	–	–	70.76	Glc			
1″	4.27	(*d*) 7.71	104.679	δa	2.64		46.19	1″	4.504	(*d*) 8.01	102.94
2″	3.28		78.155	δb	2.64			2″	4.912		75.33
3″	3.4-3.15		78.155	ε	–	–	174.90	3″	3.653		76.2
4″	3.381		72.068	ω	1.358		27.93	4″	3.465		71.85
5″	3.527		75.329					5″	3.371		78.15
6″a	4.493	(*dd*) 11.92, 1.92	64.893					6″a	3.907		
6″b	4.391	(*dd*) 11.89, 6.14						6″b	3.759		62.93
Glc								Glc			
1‴	4.17	(*d*) 7.71	104.896					1‴	4.174	(*d*) 7.68	104.89
2‴	3.21		75.329					2‴	3.230		75.33
3‴	3.4-3.15		78.155					3‴	3.369		78.15
4‴	3.31		71.633					4‴	3.416		71.63
5‴	3.4-3.15		78.155					5‴	3.218		77.72
6‴a	3.84		62.937					6‴a	3.830		62.72
6‴b	3.687							6‴b	3.700		
CouA								FeA			
1*	–	–	129.898					1*	–	–	127.66
2*	7.46	(*d*) 8.81	130.985					2*	6.94	(*d*) 1.51	111.77
3*	7.09	(*d*) 8.83	117.941					3*	–	–	149.22
4*	–	–	160.77					4*	–	–	150.46
5*	7.09	(*d*) 8.83	117.941					5*	6.76	(*d*) 8.05	122.938
6*	7.46	(*d*) 8.81	130.985					6*	6.91	(*dd*) 8.20, 1.89	123.98
7*	7.61	(*d*) 16.02	146.204					7*	7.52	(*d*) 15.84	147.28
8*	6.37	(*d*) 15.58	116.854					8*	6.3	(*d*) 15.97	116.44
9*			168.597					9*	–	–	168.47
								OCH_3	3.8		56.411
Glc											
1⁗	4.98	(*d*) 7.05	101.853								
2⁗	3.517		75.112								
3⁗	3.4-3.15		77.938								
4⁗	3.4-3.15		71.416								
5⁗	3.235		77.938								
6⁗a	3.91		62.502								
6⁗b	3.71										

[a] Annotations are given in Table 2.
[b] Due to overlapping signals, not all *J*-values could be determined.

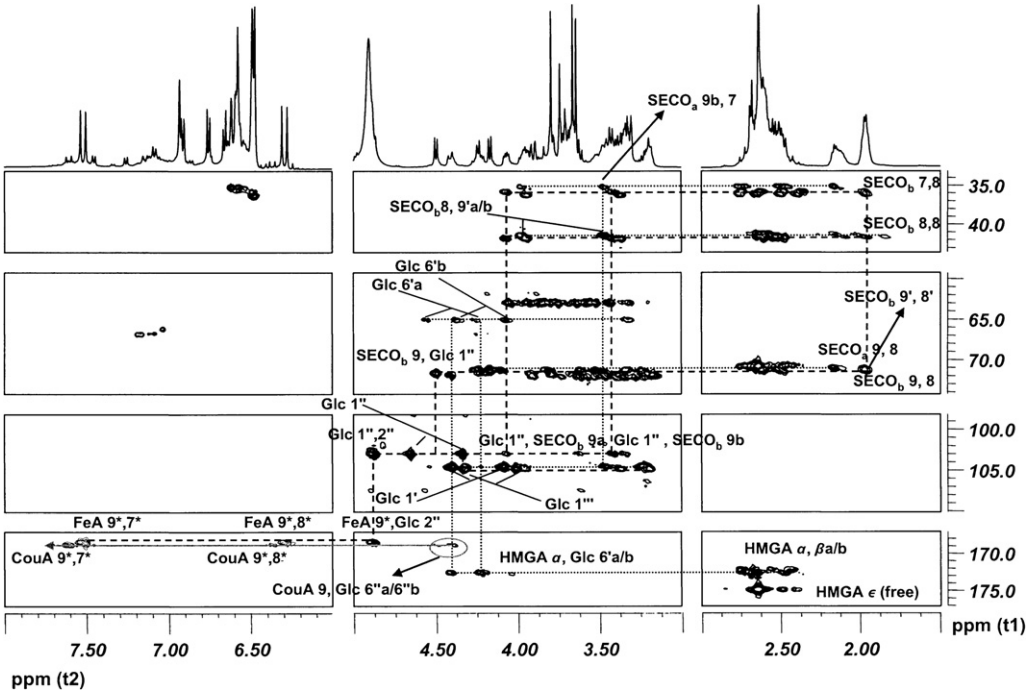
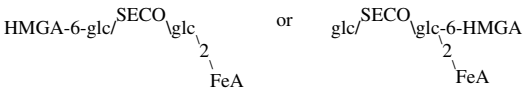


Fig. 3. HMBC spectrum of fraction 8* of the lignan macromolecule. Cross-peaks of SDG + 2HMGA (·····), SDG + FeA (------) and SDG + CouAG (–––) are indicated. Numbering corresponds to numbers given in Table 2. Cross-peaks are indicated as the coupling of carbon to proton. For example, the notation of FeA 9*, Glc 2″ indicates the coupling of the C-9* to the H-2″, the notation Glc‴ indicates a single bond coupling of C-1‴ to H-1‴ of Glc.



In fraction 10* a structural element was identified for FeAG (see spectral data in Section 4). A cross-peak between C-9 and CH_2 of EtOH led to the identification of FeAG + EtOH (see Table 2) and showed that also FeAG was ester-linked within the lignan macromolecule via its carboxyl group. In none of the fractions cross-peaks showing the coupling between FeAG and SDG were found.

In the MS data (Table 1) of fraction 8*, a structural element annotated as SDG + HMGA + FeAG was found. This structure might show the linkage of FeAG to the lignan macromolecule. The MS/MS data showed that glucose and glucose + HMGA could be split off. However, these data were not decisive about the position of HMGA. Therefore, it remains to be determined at which position FeAG links to SDG.

It is remarkable that based on MS analysis fragments with FeAG were annotated, and that these structural elements were not found with NMR. This might be explained based on the sensitivity to ionization. The fragment identified by MS carries a HMGA moiety. The free carboxyl moiety of this fragment is already in a deprotonated, charged state or is very easily ionized. Therefore, it is likely to give a large response on MS (Cech and Enke, 2001). NMR signals are not dependent on such phenomena, so they are proportional with the amounts present in the sample. The present data suggest that the fragment SDG + HMGA + FeAG was present in such low amounts that the NMR signals of this fragment could not be distinguished from the noise.

3. Discussion

The aim of this study was to determine how CouAG and FeAG are linked within the lignan macromolecule. Several structural elements of the lignan macromolecule from flaxseed hulls were identified. CouAG is attached directly via its carboxyl group to the C-6 of a glucosyl moiety of SDG. FeA is found to be linked with its carboxyl group to the C-2 position of a glucosyl moiety of SDG. These fragments show that the hydroxycinnamic acid (glucosides) are linked directly to SDG and not via the linker HMGA.

During biosynthesis, the attachment of HMGA to SDG is mediated by coenzyme A (CoA)-activated HMGA (Ford et al., 2001). It was recently shown that also HDG could be a target of HMGA coupling (Struijs et al., 2007). Since linkage of HMGA to SDG and HDG occurred on the glucosyl moiety, it was expected that HMGA could also link to other molecules carrying a glucosyl moiety. The present study shows that hydroxycinnamic acids are not linked within the lignan macromolecule via HMGA. Based on the finding that hydroxycinnamic acid (glucosides) are linked directly to SDG, it is suggested that the CoA-activated hydroxycinnamic acids, identified as intermediates of the monolignol biosynthesis (Boerjan et al., 2003), are

the driving force behind the linkage of the hydroxycinnamic acids to SDG.

The new structural elements SDG + CouAG and SDG + FeA are linked within the lignan macromolecule. Structures in which HMGA is linked to these elements were annotated based on MS analysis, providing indications on how they are integrated in the lignan macromolecule. These (tentative) structural elements are shown in Table 4.

Chain elongation of structural elements of SDG + CouAG or SDG + FeAG (Table 4, elements 3 and 5) might in theory occur via linkage of HMGA to the glucosyl moieties of FeAG or CouAG, or via linkage of HMGA to a glucosyl moiety of SDG. By NMR, no cross-peaks between the glucosyl moieties of CouAG or FeAG and HMGA (C-α or C-ε) were observed. Furthermore, no fragments of SDG + 2HMGA + (CouAG or FeAG) with HMGA linked to both ends of the fragment could be annotated. These observations point to HMGA being linked to a glucosyl moiety of SDG, and CouAG and FeAG being the terminal units of the lignan macromolecule (Table 4, elements 4 and 6). Chain elongation of SDG + FeA (Table 4, element 7) is only possible via linkage of HMGA to one of the glucosyl moieties of SDG (see also Section 2 and Table 4, elements 8).

FeA is linked to the C-2 position of a glucosyl moiety of SDG. However, it is less clear how FeAG is linked with the lignan macromolecule. FeAG is found to be linked within the lignan macromolecule via its carboxyl group. There are two possible positions at which FeAG can be attached to SDG: similar to FeA to the C-2 position of a glucosyl moiety or similar to CouAG to the C-6 position (Table 4, element 5). Various examples of linkage types of hydroxycinnamic acids, but not of hydroxycinnamic acid glucosides, to sugar moieties have been described. In arabinan, FeA links to the C-2 position of arabinose (Colquhoun et al., 1994; Ishii and Tobita, 1993), whereas for anthocyanidins attachment of FeA and CouA to the C-6 position of the glucosyl moiety was shown (Fossen et al., 2005; Matsufuji et al., 2003). So also based on the literature, no indications can be found about the position of attachment of FeAG to SDG.

Table 4
Overview of (tentative) structural elements of the lignan macromolecule from flaxseed hulls

Element	Structure
1	**HMGA-6-glc-SECO-glc**
2	**HMGA-6-glc-SECO-glc-6-HMGA**
3	**glc-SECO-glc-6-CouA-glc**
4	*HMGA*-**6-glc-SECO-glc-6-CouA-glc**
5	**glc-SECO-glc-*n*-FeA-glc**
6	*HMGA*-**6-glc-SECO-glc-*n*-FeA-glc**
7	**glc-SECO-glc-2-FeA**
8	*HMGA*-**6-glc-SECO-glc-2-FeA**
	glc-SECO-glc(-2-FeA)-6-*HMGA*

Parts of the structures in italics are tentative. n = unidentified but most likely represents linkage to C-2 or C-6 position of the glucosyl moiety.

In conclusion, CouAG and FeA(G) link within the lignan macromolecule from flaxseed hulls via ester-linkage of their carboxyl groups to glucosyl moieties of SDG. Attachment of HMGA to the glucosyl moiety of CouAG or FeAG is not observed. Therefore, it is suggested that CouAG and FeAG are terminal units of the lignan macromolecule.

4. Experimental

4.1. Lignan extraction from flaxseed hulls

Lignan macromolecule was extracted from flaxseed hulls, kindly provided by Frutarom Netherlands B.V. (Veenendaal, The Netherlands). The extraction procedure was as described in Struijs et al. (2007). In short, flaxseed hulls were defatted by soxhlet extraction. The lignan macromolecule was extracted from the defatted hulls by a three-step sequential extraction with 63% (v/v) aq. EtOH for 4 h at room temperature under continuous stirring. The extracts and the hulls were separated by filtration. The three extracts were combined and the EtOH was evaporated. The concentrated extract was lyophilized resulting in the lignan macromolecule extract.

4.2. Saponification of the lignan macromolecule

Lignan macromolecule extract of 2 mg/ml in 63% (v/v) aq. EtOH was partially saponified with 2 mM NaOH to obtain lignan macromolecular fragments. Saponification was performed at room temperature while stirring. A reaction volume of 2.5 l was used. After 24 h, the reaction was stopped by lowering the pH to 6.5–7.0 with glacial HOAc.

4.3. Sample clean up of saponified lignan macromolecule

Low molecular weight polar material was removed from the partially saponified samples by solid phase extraction (SPE; SepPak Vac, 20 cc/5 g, C18 cartridge, Waters) following the procedure previously described in Struijs et al. (2007).

4.4. Analytical reversed phase HPLC coupled on-line to mass spectrometry (RP-HPLC/MS)

Samples, which were collected after (semi)-preparative RP-HPLC (see below), were analyzed on an analytical X-Terra C18 MS column (Waters; 3.5 μm particle size, 4.6 × 150 mm) following the procedure described previously (Struijs et al., 2007).

The molecular masses and MS/MS-fragmentation patterns of the lignans and fragments of the lignan macromolecule were determined on a Thermo Finnigan LCQ Classic equipped with a ESI probe in the negative mode coupled on-line to the analytical RP-HPLC.

4.5. Purification by (semi)-preparative RP-HPLC

Fragments of the lignan macromolecule obtained by partial saponification were purified by (semi)-preparative RP-HPLC. For purification of 100–600 mg fragments of the lignan macromolecule, an X-Terra C18 MS column of 50 × 100 mm (Waters; 5 μm particle size, OBD) with an X-Terra C18 MS guard column (Waters; 5 μm particle size, 19 × 10 mm) was used. For purification of 100 mg fragments of the lignan macromolecule or less, the samples were separated on a semi-preparative X-Terra C18 MS column (Waters; 5 μm particle size, 29 × 150 mm, OBD) with an X-Terra C18 MS guard column (Waters; 5 μm particle size, 19 × 10 mm). The separation protocols used were similar as described previously (Struijs et al., 2007).

4.6. Nuclear magnetic resonance (NMR)

NMR spectra were recorded on a Bruker AMX-500 spectrometer located at the Wageningen NMR Centre as described previously (Struijs et al., 2007). ^{1}H and ^{13}C proton decoupled spectra were recorded. All 2D COSY spectra were acquired using the double quantum filtered (DQF) method with a standard pulse sequence delivered by Bruker. For the 2D HMBC spectrum a standard gradient enhanced 2D HMQC pulse sequence delivered by Bruker, was changed into a HMBC sequence.

4.7. Spectral data

trans-CouAG + EtOH (fraction 9*): ^{1}H NMR (500.13 MHz, CD_3OD) δ 7.54 (2H, *d*, $J = 8.74$ Hz, H-2/6), 7.11 (2H, *d*, $J = 8.74$ Hz, H-3/5), 7.62 (1H, *d*, $J = 15.95$ Hz, H-7), 6.39 (1H, *d*, $J = 16.05$ Hz, H-8), 4.97 (1H, *d*, $J = 7.28$ Hz, H-1′), 3.5–3.3 (4H, *m*, H-2′/3′/5′/4′), 3.90 (1H, *dd*, $J = 12.13$, 2.02 Hz, 1H, H-6a′), 3.70 (1H, *dd*, $J = 12.09$, 5.53 Hz, 1H, H6b′), 4.23 (2H, *q*, $J = 7.14$ Hz, CH_2), 1.31 (3H, *t*, $J = 7.17$ Hz, CH_3).

trans-FeAG + EtOH (fraction 10*): ^{1}H NMR (500 MHz, CD_3OD) δ 7.62 (1H, *d*, $J = 15.93$ Hz, H-7), 7.24 (1H, *d*, $J = 1.24$ Hz, H-2), 7.15 (1H, *dd*, $J = 8.47$, 1.53 Hz, H-6), 7.17 (1H, *d*, $J = 8.45$ Hz, H-5), 6.43 (1H, *d*, $J = 16.00$ Hz, H-8), 4.97 (1H, *d*, $J = 7.38$ Hz, H-1′), 3.89 (3H, *s*, CH_3), 3.89 (1H, *dd*, $J = 11.90$, 1.89 Hz, H-6a′), 3.69 (1H, *dd*, $J = 11.90$, 5.31 Hz, H6b′), 3.51 (2H, H-2′/3′), 3.40 (1H, H-4′), 3.45 (1H, H-5′), 4.23 (2H, *q*, $J = 7.1$ Hz, CH_2), 1.31 (3H, *t*, $J = 7.1$ Hz, CH_3).

^{13}C NMR (125.77 MHz, CD_3OD) δ 130.6 (C-1), 112.5 (C-2), 151.0 (C-3), 150.1 (C-4), 117.4 (C-5), 123.5 (C-6) 145.9 (C-7), 117.5 (C-8), 169.0 (COOD), 56.8 (OCH_3), 102.2 (C-1′), 74.9 (C-2′), 78.3 (C-3′), 71.3 (C-4′), 77.9 (C-5′), 62.5 (C-6′), 61.6 (CH_2), 14.69 (CH_3).

References

Bakke, J.E., Klosterman, H.J., 1956. A new diglucoside from flaxseed. Proc. North Dakota Acad. Sci. 10, 18–21.

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

Cech, N.B., Enke, C.G., 2001. Practical implications of some recent studies in electrospray ionization fundamentals. Mass Spectrom. Rev. 20, 362–387.

Colquhoun, I.J., Ralet, M.C., Thibault, J.F., Faulds, C.B., Williamson, G., 1994. Structure identification of feruloylated oligosaccharides from sugar-beet pulp by NMR-spectroscopy. Carbohydr. Res. 263, 243–256.

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.

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.

Fossen, T., Rayyan, S., Holmberg, M.H., Nateland, H.S., Andersen, O.M., 2005. Acylated anthocyanins from leaves of *Oxalis triangularis*. Phytochemistry 66, 1133–1140.

Ishii, T., Tobita, T., 1993. Structural characterization of feruloyl oligosaccharides from spinach-leaf cell-walls. Carbohydr. Res. 248, 179–190.

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., Clagett, C.O., 1955. The constitution of linocinnamarin. J. Am. Chem. Soc. 77, 420–421.

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

Matsufuji, H., Otsuki, T., Takeda, T., Chino, M., Takeda, M., 2003. Identification of reaction products of acylated anthocyanins from red radish with peroxyl radicals. J. Agric. Food Chem. 51, 3157–3161.

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.

Milder, I.E.J., Arts, I.C.W., van de Putte, B., Venema, D.P., Hollman, P.C.H., 2005. Lignan contents of Dutch plant foods: a database including lariciresinol, pinoresinol, secoisolariciresinol and matairesinol. Br. J. Nutr. 93, 393–402.

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.

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.

Struijs, K., Vincken, J.-P., Verhoef, R., van Oostveen-van 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.

Thompson, L.U., Seidl, M.M., Rickard, S.E., Orcheson, L.J., Fong, H.H.S., 1996. Antitumorigenic effect of a mammalian lignan precursor from flaxseed. Nutr. Cancer 26, 159–165.

Vanharanta, M., Voutilainen, S., Lakka, T.A., van der Lee, M., Adlercreutz, H., Salonen, J.T., 1999. Risk of acute coronary

events according to serum concentrations of enterolactone: a prospective population-based case-control study. Lancet 354, 2112–2115.

Ward, W.E., Yuan, Y.V., Cheung, A.M., Thompson, L.U., 2001. Exposure to purified lignan from flaxseed (*Linum usitatissimum*) alters bone development in female rats. Br. J. Nutr. 86, 499–505.

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