PERGAMON

Phytochemistry 59 (2002) 663–671

PHYTOCHEMISTRY

www.elsevier.com/locate/phytochem

Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*

Michael Reichelt, Paul D. Brown[1], Bernd Schneider, Neil J. Oldham, Einar Stauber, Jim Tokuhisa, Daniel J. Kliebenstein[2], Thomas Mitchell-Olds, Jonathan Gershenzon*

Max-Planck Institute for Chemical Ecology, Beutenberg Campus, Winzerlaer Strasse 10, D-07745 Jena, Germany

Received 1 October 2001; received in revised form 19 December 2001

Abstract

The spectacular recent progress in *Arabidopsis thaliana* molecular genetics furnishes outstanding tools for studying the formation and function of all metabolites in this cruciferous species. One of the major groups of secondary metabolites in *A. thaliana* is the glucosinolates. These hydrophilic, sulfur-rich glycosides appear to serve as defenses against some generalist herbivores and pathogens, and as feeding and oviposition stimulants to specialist herbivores. To help study their biosynthesis and role in plant-insect interactions, we wanted to determine the complete glucosinolate content of *A. thaliana*. In previous studies, 24 glucosinolates had been identified from ecotype Columbia. We reinvestigated Columbia as well as additional ecotypes and mutant lines, and identified 12 further glucosinolates, including five novel compounds. Structures were elucidated by MS and NMR spectroscopy of their desulfated derivatives, and by enzymatic cleavage of the attached ester moieties. Four of the novel glucosinolates are benzoate esters isolated from the seeds. In all but one of these compounds, esterification is on the glucose moiety rather than the side chain, a very unusual feature for glucosinolates. Among additional glucosinolates identified were the first non-chain elongated, methionine-derived glucosinolate from *A. thaliana* and the first compounds that appear to be derived from leucine.

Keywords: *Arabidopsis thaliana*; Brassicaceae; Crucifers; Structure elucidation; Glucosinolates; Benzoic acid esters; Benzoyloxy glucosinolates; Sinigrin; Myrosinase; Methionine

1. Introduction

Glucosinolates are sulfur-rich plant secondary metabolites whose basic skeleton consists of a β-thioglucose residue, an *N*-hydroxyiminosulfate moiety and a variable side chain (Halkier, 1999). More than 120 different glucosinolates have been identified (Fahey et al., 2001), almost exclusively from species of the Brassicaceae (the mustard family) and several closely related families. These substances play prominent roles in plant-herbivore and plant-pathogen interactions (Chew, 1988; Halkier, 1999), and are attracting increasing attention for their anticarcinogenic properties (Hecht, 2000).

One glucosinolate-producing species, *Arabidopsis thaliana* (L.) Heynh., has become a model system for studying many aspects of plant biology and biochemistry. The availability of extensive mutant collections, mapping tools and the full genomic sequence for this taxon provide extraordinary resources for studying the formation and function of all metabolites present, including glucosinolates. As part of our interest in the biosynthesis of glucosinolates in *A. thaliana* and their role in plant defense, we wanted to determine the complete glucosinolate profile of this species.

Previous work on this species has identified 24 glucosinolates (Hogge et al., 1988b; Agerbirk et al., 2001) which can be divided into three classes based on their likely amino acid precursor: methionine, phenylalanine or tryptophan (Table 1). The methionine-derived glucosinolates are the largest and structurally most variable class, with methylthioalkyl, methylsulfinylalkyl, alkenyl and hydroxyalkyl side chains being found. In the seeds, the hydroxyalkyl glucosinolates often occur as benzoate esters (Hogge et al., 1988b; Graser et al., 2001). The

* Corresponding author. Tel.: +49-3641-571300; fax: +49-3641-571302.

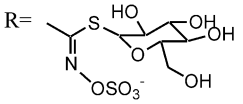
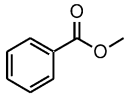
E-mail address: gershenzon@ice.mpg.de (J. Gershenzon).

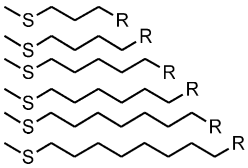
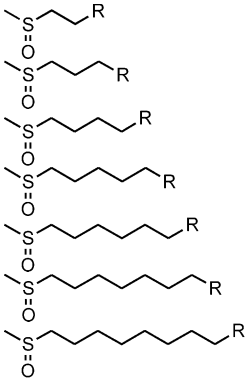
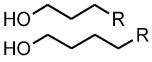
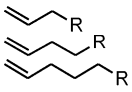
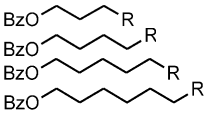
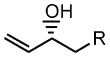
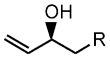
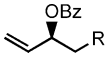
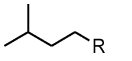
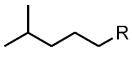
[1] Current address: Department of Environmental Studies, Trinity Western University, Langley, British Columbia V2Y 1Y1, Canada.

[2] Current address: Department of Vegetable Crops, University of California, Davis, CA 95616, USA.

0031-9422/02/$ - see front matter
PII: S0031-9422(02)00014-6

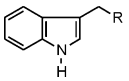
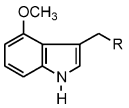
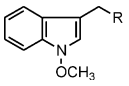
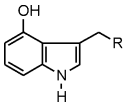
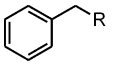
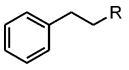
Table 1
Glucosinolates identified from *Arabidopsis thaliana* (Columbia and other ecotypes)

R=  , BzO= 

Systematic name	Common name	Structure
Glucosinolates derived from methionine		
With methylthioalkyl side chains		
3-Methylthiopropyl glucosinolate	Glucoiberverin	
4-Methylthiobutyl glucosinolate	Glucoerucin	
5-Methylthiopentyl glucosinolate	Glucoberteroin	
6-Methylthiohexyl glucosinolate	Glucolesquerellin	
7-Methylthioheptyl glucosinolate		
8-Methylthiooctyl glucosinolate		
With methylsulfinylalkyl side chains		
2- Methylsulfinylethyl glucosinolate[a,b]		
3-Methylsulfinylpropyl glucosinolate	Glucoiberin	
4-Methylsulfinylbutyl glucosinolate	Glucoraphanin	
5-Methylsulfinylpentyl glucosinolate	Glucoalyssin	
6-Methylsulfinylhexyl glucosinolate	Glucohesperin	
7-Methylsulfinylheptyl glucosinolate	Glucoibarin	
8-Methylsulfinyloctyl glucosinolate	Glucohirsutin	
With hydroxyalkyl side chains		
3-Hydroxypropyl glucosinolate		
4-Hydroxybutyl glucosinolate		
With alkenyl side chains		
2-Propenyl glucosinolate (allyl)[a,c]	Sinigrin	
3-Butenyl glucosinolate	Gluconapin	
4-Pentenyl glucosinolate[a,c]	Glucobrassicanapin	
With other side chains		
3-Benzoyloxypropyl glucosinolate	Glucomalcomiin	
4-Benzoyloxybutyl glucosinolate		
5-Benzoyloxypentyl glucosinolate		
6-Benzoyloxyhexyl glucosinolate		
(2*R*)-2-Hydroxy-3-butenylglucosinolate[a,c]	Progoitrin	
(2*S*)-2-Hydroxy-3-butenylglucosinolate[a,c]	Epiprogoitrin	
(2*S*)-2-Benzoyloxy-3-butenylglucosinolate[a,c]		
With modifications on the glucose moiety		
6′-Benzoyl-4-methylthiobutyl glucosinolate[a]		see Fig. 1
6′-Benzoyl-4-methylsulfinylbutyl glucosinolate[a]		see Fig. 1
6′-Benzoyl-4-benzoyloxybutyl glucosinolate[a]		see Fig. 1
Glucosinolates derived from leucine		
3-Methylbutyl glucosinolate[a,c]		
4-Methylpentyl glucosinolate[a,c]		

(*continued on next page*)

Table 1 (*continued*)

Systematic name	Common name	Structure
Glucosinolates derived from tryptophan		
Indol-3-ylmethyl glucosinolate	Glucobrassicin	
4-Methoxyindol-3-ylmethylglucosinolate	4-Methoxyglucobrassicin	
1-Methoxyindol-3-ylmethylglucosinolate	Neoglucobrassicin	
4-Hydroxyindol-3-ylmethylglucosinolate	4-Hydroxyglucobrassicin	
Glucosinolates derived from phenylalanine		
Benzyl glucosinolate[a,c]	Glucotropaeolin	
2-Phenethyl glucosinolate	Gluconasturtiin	

[a] Identified for first time in this study. Other compounds first identified by Hogge et al. (1988b) or Agerbirk et al. (2001).
[b] Isolated from mutant TU1.
[c] Isolated from ecotypes other than Columbia.

structural diversity of the methionine-derived glucosinolates is increased further by the elongation of their side chains with 1–6 additional methylene groups.

Here we report the isolation of 12 new glucosinolates from *A. thaliana*. Of these, five (**3**–**7**) are novel, methionine-derived compounds, all but one of which are benzoate esters. Previous investigations have focused on the Columbia ecotype (Col-0). We have reinvestigated this ecotype (three of the five novel glucosinolates (**3**–**5**) were isolated from Columbia), as well as searching for novel constituents in additional ecotypes and mutants. The complete distribution of glucosinolates in 39 *Arabidopsis* ecotypes has been reported separately (Kliebenstein et al., 2001a).

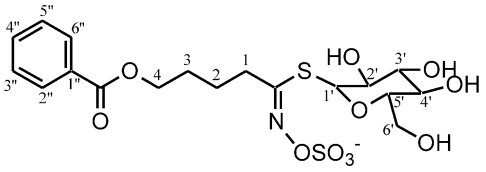
In the present study, structure elucidation (UV, MS and NMR) was carried out on the desulfated derivatives of the isolated glucosinolates (**d1**–**d7**), because on-column desulfation is a very efficient step in separating glucosinolates from other components of the plant extracts (Thies, 1979). However, the naturally occurring compounds are the sulfated glucosinolates (**1**–**7**) (Fig. 1).

2. Results and discussion

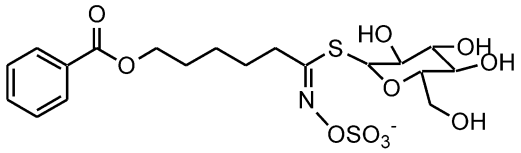
2.1. Novel benzoic acid glucosinolate esters from seeds

Glucosinolates were extracted from seeds of the ecotype Columbia (Col-0) and separated from other constituents of the extract by on-column desulfation on an anion-exchange matrix. The HPLC-UV chromatogram revealed many previously reported desulfated glucosinolates (Hogge et al., 1988b; Agerbirk et al., 2001), identified by comparison of their retention times and UV spectra to those of standards. Especially prominent, were compounds with benzoyloxy substituents, characterized by their distinctive UV spectra (absorption maxima at 195 and 230 nm, shoulder at 275 nm). The most abundant of this group were desulfo-3-benzoyloxypropyl and desulfo-4-benzoyloxybutyl (**d1**) glucosinolate. Four other desulfated benzoyloxyalkyl glucosinolates, whose identities were unclear from the chromatogram, were purified by preparative HPLC and subjected to MS and NMR analyses. Three proved to be the novel substances **d3**–**d5** (described below), while one was desulfo-5-benzoyloxypentyl glucosinolate (**d2**), previously reported by Hogge et al. (1988b) based on the occurrence of the corresponding isothiocyanate hydrolysis products following treatment with myrosinase. Spectral data for **d2** are also presented here since this was the first isolation of this compound.

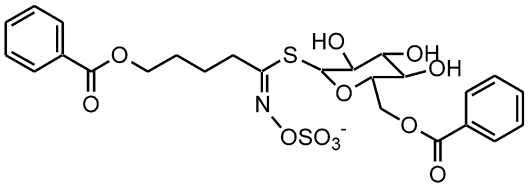
Atmospheric pressure chemical ionization MS in a positive ion mode resulted in an m/z of 430, representing the expected $[M+H]^+$ ion of **d2** ($C_{19}H_{28}NO_8S$). Collision induced dissociation (CID) of m/z 430 yielded product ions at m/z 268, 146, 137 and 105 compared to fragments of m/z 254, 132 and 105 for **d1**. The ^1H NMR spectrum of **d2** closely resembled that of **d1** except for an additional multiplet for a more shielded methylene group at δ 1.46, which was assigned to H_2–3 (Table 2).



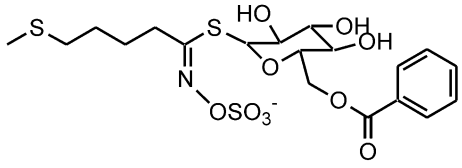
1



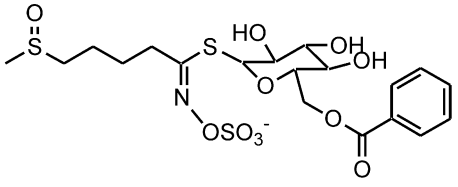
2



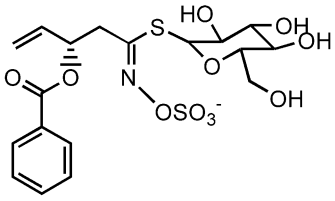
3



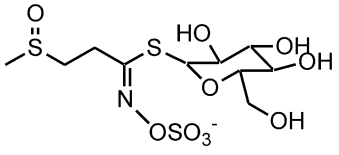
4



5



6



7

Characteristic signals of the benzoyl moiety appeared at δ 7.97 (H-2″/H-6″), 7.66 (H-4″) and 7.53 (H-3″/H-5″). The chemical shift of H_2–5 (δ 4.27) indicated attachment of the benzoyloxy unit to the terminal position of the alkyl chain. The ^1H–^1H COSY spectrum confirmed the assignment of ^1H signals of the benzoyl ring and the aliphatic chain, and was also used to identify the glucose signals including the hydroxyl protons. As a final confirmation of its structure, **d2** was treated with esterase to yield desulfo-5-hydroxypentyl glucosinolate, identified by comparison of its HPLC retention time and UV spectrum with that of a standard.

The desulfated derivative of compound **3** showed the typical UV spectrum of a benzoyloxyalkyl glucosinolate. However, the parent ion (m/z 520 for $[M+H]^+$) did not correspond to that of any known compound of this type. The product ion spectrum of m/z 520 showed a fragment at m/z 105 indicative of a benzoyl moiety, but did not show any fragment at $[M+H\text{-}162]^+$ representing the loss of the glucose moiety, as seen in the mass spectrum of all typical desulfated glucosinolates. Instead, a strong fragment at m/z 254 was present which could be attributed to the loss of glucose and a benzoyl moiety together $[M+H\text{-}266]^+$, along with prominent fragments at m/z 267, 249 and 132. The fragments at m/z 254 and 132 were also observed for **d1** (Graser et al., 2001) and indicated the presence of a 4-benzoyloxybutyl side chain. The ^1H NMR spectrum of **d3** was very similar to that of **d1**, but the additional set of benzoyl signals in the aromatic part of the spectrum indicated the occurrence of a second benzoyl moiety (Table 2). The absence of a signal for one of the four sugar hydroxyl groups provided evidence for the attachment of the additional benzoyl function to the glucose unit. The downfield shift of signals for H-6′a and H-6′b by nearly one full ppm indicated attachment at C-6′, which was confirmed by corresponding ^{13}C NMR shift differences (observed in the HMQC spectra) between the benzoylated and nonbenzoylated glucose moieties of compounds **d3** and **d1**, respectively (Table 3). Thus, **3** is 6′-benzoyl-4-benzoyloxybutyl glucosinolate.

As compound **3** has two benzoic acid groups that might require cleavage before metabolic activation by myrosinase (Rask et al., 2000), it was interesting to test the behavior of its desulfated derivative (**d3**) upon esterase treatment. Compound **d3** completely disappeared from the HPLC-UV chromatogram in the first minute after application of porcine liver esterase, and was replaced by an unknown substance assigned by MS/MS experiments to be desulfo-6′-benzoyl-4-hydroxybutyl glucosinolate ($[M+H]^+$ at m/z 416, fragments at m/z 267, 249, 150, 105). After one hour of esterase treatment, the HPLC-UV trace finally showed a signal for desulfo-4-hydroxybutyl glucosinolate, indicating

Table 2
^{1}H NMR spectral data of desulfonated derivatives of compounds **1–5** and **7**[a]

	H	d1	d2	d3	d4	d5	d7
Aliphatic moiety	1	2.61 (*m*)	2.55 (*m*)	2.44 (*m*) / 2.58 (*m*)	2.37 (*m*)	2.61 (*m*)	3.04 (*m*)
	2	1.73 (*m*)	1.64 (*m*)	1.55 (*m*) / 1.66 (*m*)	1.48 (*m*)	1.53 (*m*)	3.08 (*m*) / 3.25 (*m*)
	3	1.76 (*m*)	1.46 (*m*)	1.57 (*m*)	1.40 (*m*)	1.51 (*m*)	
	4	4.30 (*t, J* = 7.7 Hz)	1.74 (*m*)	4.16 (*m*)	2.33 (*m*)	2.63 (*m*)	
	5	–	4.27 (*t, J* = 6.6 Hz)				
	$S(O)CH_3$					2.46 (*s*)[b]	2.68 (*s*)[b]
	SCH_3				1.96 (*s*)[b]		
	N–OH	10.93 (*s*)	10.90 (*s*)	10.98 (*s*)	10.95 (*s*)	10.98 (*br s*)	n.d.[c]
Glucose moiety	1′	4.70 (*d, J* = 9.9 Hz)	4.67 (*d, J* = 9.9 Hz)	4.81 (*d, J* = 9.9 Hz)	4.79 (*d, J* = 9.9 Hz)	4.80 (*d, J* = 9.5 Hz)	5.04 (*d, J* = 9.8 Hz)
	2′	3.04 (*dd, J* = 9.9, 9.2 Hz)	3.03 (*dd, J* = 9.9, 9.2 Hz)	3.09 (*m*)	3.08 (*m*)	3.09 (*m*)	3.44 (*dd, J* = 9.8, 9.8 Hz)
	3′	3.21 (*dd, J* = 9.2, 8.4 Hz)	3.21 (*dd, J* = 9.2, 8.4 Hz)	3.28 (obscured by HDO)	3.28 (obscured by HDO)	~3.3 (obscured by HDO)	3.55 (*dd, J* = 9.8, 9.8 Hz)
	4′	3.08 (*dd, J* = 8.4, 9.2 Hz)	3.09 (*dd, J* = 8.4, 9.2 Hz)	3.20 (*m*)	3.19 (*m*)	3.21 (*m*)	3.44 (*dd, J* = 9.8, 9.8 Hz)
	5′	3.18 (*m*)	3.17 (*m*)	3.65 (*ddd, J* = 8.8, 7.7, 1.2 Hz)	3.64 (*ddd, J* = 8.8, 7.7, 1.5 Hz)	3.66 (*m*)	3.57 (*dd, J* = 9.8, 2.2 Hz)
	6′a	3.78 (*br d, J* = 11.7 Hz)	3.66 (*dd, J* = 11.7,5.1 Hz)	4.51 (*dd, J* = 11.7,1.2 Hz)	4.62 (*dd, J* = 11.7,1.5 Hz)	4.61 (*dd, J* = 11.4,5.1 Hz)	3.89 (*dd, J* = 12.5,2.2 Hz)
	6′b	3.67 (*dd, J* = 11.7, 6.2 Hz)	3.39 (*dd, J* = 11.7, 6.0 Hz)	4.23 (*dd, J* = 11.7,7.7 Hz)	4.22 (*dd, J* = 11.7,7.7 Hz)	4.24 (*dd, J* = 11.4,7.7 Hz)	3.69 (*dd, J* = 12.5,5.9 Hz)
	2′-OH	5.36 (*br s*)	5.34 (*d, J* = 5.9 Hz)	5.49 (*d, J* = 5.9 Hz)	5.46 (*d, J* = 5.9 Hz)	5.49 (*d, J* = 7.0 Hz)	n.d.
	3′-OH	4.99 (*br s*)	4.99 (*d, J* = 5.5 Hz)	5.27 (*brs*)	5.26 (*d, J* = 5.1 Hz)	5.40 (*d, J* = 5.1 Hz)	n.d.
	4′-OH	5.09 (*br s*)	5.10 (*d, J* = 5.1 Hz)	5.38 (*d, J* = 4.8 Hz)	5.37 (*d, J* = 5.5 Hz)	5.48 (*d, J* = 5.1 Hz)	n.d.
	6′-OH	4.58 (*br s*)	4.57 (*t, J* = 5.8 Hz)	–	–	–	n.d.
Benzoyl moiety	2″/6″	7.96 (*dd, J* = 8.4, 1.5 Hz)	7.97 (*d, J* = 8.4 Hz)	7.93 (*d, J* = 8.4 Hz) / 7.95 (*d, J* = 8.4 Hz)	7.96 (*d, J* = 8.4 Hz)	7.95 (*dd, J* = 8.4, 1.5 Hz)	
	4″	7.66 (*tt, J* = 7.3, 1.5 Hz)	7.66 (*t, J* = 7.7 Hz)	7.60 (*t, J* = 7.7 Hz) / 7.66 (*t, J* = 7.7 Hz)	7.68 (*t, J* = 7.3 Hz)	7.68 (*t, J* = 7.3 Hz)	
	3″/5″	7.53 (*dd, J* = 7.3, 8.4 Hz)	7.53 (*dd, J* = 7.7, 8.4 Hz)	7.50 (*dd, J* = 8.4, 7.7 Hz)/ 7.53 (*dd, J* = 8.4, 7.7 Hz)	7.55 (*dd, J* = 8.4,7.3 Hz)	7.56 (*d, J* = 8.4, 7.3 Hz)	

[a] Chemical shifts (δ) were recorded at 500.13 MHz with DMSO-d_6 (**d1**–**d5**) and D_2O (**d7**) as solvents and TMS (**d1**–**d5**) and TMSP-2,2,3,3-d_4 (**d7**) as internal standards.
[b] Three proton intensity.
[c] n.d. = Not detected.

Table 3
^{13}C NMR spectral data of desulfonated derivatives of compounds **1**, **3**, **4**, and **7**[a]

	C	d1	d3	d4	d7
Aliphatic moiety	0	n.d.[b]	n.d.	n.d.	n.d.
	1	31.2	30.5	30.7	30.1
	2	23.7	22.9	25.8	54.2
	3	28.0	27.3	27.5	
	4	64.7	64.1	32.7	
	SCH_3			14.2	
	$S(O)CH_3$				41.4
Glucose moiety	1′	81.9	81.2	81.2	86.0
	2′	73.2	72.5	72.5	76.8
	3′	78.7	77.6	77.5	81.8
	4′	70.3	69.7	69.9	73.9
	5′	81.0	77.2	77.2	84.6
	6′	61.5	64.5	64.4	65.4
Benzoyl moiety	COO	n.d.	n.d.	n.d.	
	1″	n.d.	n.d.	n.d.	
	2″/6″	129.3	128.8 / 128.8	128.7	
	4″	133.3	133.0 / 133.0	133.1	
	3″/5″	129.1	128.4 / 128.4	128.5	

[a] ^{13}C NMR data were obtained from HMQC spectra. Chemical shifts (δ) were recorded at 125.76 MHz with DMSO-d_6 (**d1**–**d4**) and D_2O (**d7**) as solvents and TMS (**d1**–**d4**) or TMSP-2,2,3,3-d_4 (**d7**) as internal standards.
[b] n.d. = Not detectable by means of HMQC.

that cleavage of the ester bond in the aliphatic side chain is much faster than that in the glucose moiety.

From its UV spectrum, **d4** was also likely to contain a benzoyl moiety. However, the product of esterase cleavage of **d4** was not a desulfated hydroxyalkyl glucosinolate but desulfo-4-methylthiobutyl glucosinolate. Mass spectrometry of **d4** showed a molecular species ($[M+H]^+$) at m/z 446. The product ion spectrum of m/z 446 showed some of the same fragments as the product ion spectrum of the $[M+H]^+$ of compound **d3**, m/z 267, 249 and 105, suggesting a similar benzoate esterification of the glucose moiety. However, the fragments from the 4-benzoyloxybutyl side chain (m/z 254 and 132) were absent. Instead, a fragment at m/z 180 was found indicative of a 4-methylthiobutyl side chain. This fragment, which represents the aliphatic side chain remaining after glucose cleavage, was also observed during the MS/MS analysis of a standard sample of desulfo-4-methylthiobutyl glucosinolate. The ^{1}H NMR spectrum of **d4** exhibited signals for an aliphatic chain of four methylene groups, a glucose moiety and, in contrast to **d3**, only one benzoyl unit (Table 2). The chemical shifts of the glucose signals of **d4**, especially those of H-5′, H-6′a, and H-6′b, closely resembled those of **d3**, indicating attachment of the benzoyl unit to C-6 of the sugar. The chemical shift of the methylene signal H_2–4 at δ 2.33, which was assigned by means of a ^{1}H–^{1}H COSY spectrum, and the methyl signal (δ 1.96) were indicative of a methylthiobutyl chain. ^{13}C NMR data (Table 3), obtained from an HMQC spectrum of **d4**, also confirmed that **4** is 6′-benzoyl-4-methylthiobutyl glucosinolate.

The desulfated derivative of compound **5** (**d5**) also had a UV spectrum indicating the presence of a benzoyl moiety, but eluted significantly earlier than **d3** or **d4** in reversed-phase HPLC. Mass spectrometry of **d5** gave a molecular species ($[M+H]^+$) at m/z 462. The MS/MS experiment showed fragments at m/z 267 and 105, which suggested benzoylation on the glucose moiety, in analogy to **d3** and **d4**. The base peak was at m/z 196, consistent with a 4-methylsulfinylbutyl side chain, also indicated by the formation of desulfo-4-methylsulfinylbutyl glucosinolate upon treatment of **d5** by esterase. ^{1}H NMR (Table 2) and ^{1}H–^{1}H COSY experiments on **d5** proved the structure of compound **5** as 6′-benzoyl-4-methylsulfinylbutyl glucosinolate. In comparison with the ^{1}H NMR spectrum of **d4**, major differences were found only in the chemical shifts of the S-methyl signal (δ 2.46 vs. 1.96) and the methylene signal of H_2–4 (δ 2.63 vs. 2.33), which are both shifted downfield due to the oxidation of the thiol function to a sulfoxide.

In the course of analyzing the seeds of 39 different ecotypes of *A. thaliana* by HPLC-diode array detection (Kliebenstein et al., 2001a), several ecotypes (e. g., Pog-0, Tac) were observed to have a compound (**d6**) that eluted only 0.2 min earlier than **d1** and also gave a benzoyloxy-type UV spectrum. The occurrence of this new glucosinolate (**6**) in the sample was correlated with the occurrence of (2*S*)-2-hydroxy-3-butenyl glucosinolate. In fact, esterase treatment of isolated **d6** gave desulfo-(2*S*)-2-hydroxy-3-butenyl glucosinolate suggesting that the structure of **6** is (2*S*)-2-benzoyloxy-3-butenyl glucosinolate. The mass spectrum (molecular species at m/z 414) and CID-MS/MS (fragments at m/z 252 and 130) of **d6** were also consistent with the proposed structure. Further fragmentation of the species at m/z 130 showed the same pattern (m/z 78, 85, and 97) as that produced by fragmentation of the m/z 130 arising from authentic desulfo-(2*S*)-2-hydroxy-3-butenyl glucosinolate.

The seeds of *A. thaliana* are rich in benzoic acid glucosinolate esters, and the present investigation has added four new structures to the four reported previously (Hogge et al., 1988b). These substances are less polar than their non-benzoylated counterparts, and may represent adaptations for storage in the low water environment of the seed. Three of the four new substances described here are esterified at C-6 of the glucose moiety rather than on the terminal hydroxyl group of the side chain. Other glucosinolates have been reported that are also esterified to aromatic acids through C-6 of the glucose moiety, including 6′-sinapoyl-4-methylsulfinyl-3-butenyl glucosinolate from *Rhaphanus sativus* (Linscheid et al., 1980) and various isoferulic and sinapic acid esters of 4-methylsulfinyl-3-butenyl glucosino-

late (Bjerg and Sørensen, 1987). In considering the physiological significance of these substituents, we were intrigued by the possibility that esterification on the glucose moiety might impede hydrolysis by myrosinase (a prerequisite to formation of the biologically-active hydrolysis products). However, in a preliminary experiment the 6′-benzoyl glucosinolates from the seeds of ecotype Columbia were hydrolyzed as readily as all of the other glucosinolates in this tissue by a commercial preparation of myrosinase from *Sinapis alba* (results not shown).

2.2. Other new glucosinolates from A. thaliana

Nearly all of the glucosinolates derived from methionine are modified by the addition of at least one methylene group to the side chain via a cyclic, three-step pathway (Chapple et al., 1988; Graser et al., 2000). The few such compounds without any side chain elongation include 2-methylthioethyl glucosinolate (Grob and Matile, 1980) and 2-hydroxyethyl glucosinolate (Hu et al., 1989). The TU1 and TU5 mutant lines of *A. thaliana* ecotype Col-0 exhibit a block in the chain elongation process such that 3-methylsulfinylpropyl glucosinolate (the product of one chain elongation cycle) is a principal metabolite instead of 4-methylsulfinylbutyl glucosinolate (the product of two elongation cycles) (Haughn et al., 1991; Kroymann et al., 2001). In the process of characterizing these mutants for our biosynthetic studies, we observed an early-eluting peak in the HPLC chromatogram of the TU1 leaf extract that was not present in the extract from Col-0 wildtype leaves. The UV spectrum was similar to that of the desulfated methylsulfinylalkyl glucosinolates and the chemical ionization mass spectrum of the isolated substance gave a $[M+H]^+$ parent ion at m/z 330 ($C_{10}H_{20}NO_7S_2$), in accordance with that expected for the heretofore unreported desulfo-2-methylsulfinylethyl glucosinolate (**d7**). CID-MS/MS resulted in fragments at m/z 168 and 104, in analogy with the fragmentation pattern known for desulfo-3-methylsulfinylpropyl and desufo-4-methylsulfinylbutyl glucosinolate (Hogge et al., 1988a; Kiddle et al., 2001). The ^1H NMR spectrum, measured in D_2O, (Table 2) showed signals for a glucose unit, two methylene groups (δ 3.04 and 3.08/3.25), and a methyl signal at δ 2.68, with the downfield chemical shift of the latter signal indicative of the adjacent sulfoxide function. Further assignment of the proton signals by means of ^1H–^1H COSY, and determination of ^{13}C NMR shifts (Table 3) from the HMQC spectrum of **d7** helped confirm the structure of **7** to be 2-methylsulfinylethyl glucosinolate, the first non-chain elongated methionine-derived glucosinolate to be isolated from *A. thaliana*. The configuration at the sulfinyl group was not determined.

During our study of the various *A. thaliana* ecotypes (Kliebenstein et al., 2001a), we identified seven other glucosinolates that have been previously described in the literature, but not yet reported from this species, including allyl (2-propenyl), 4-pentenyl, (2*R*)-2-hydroxy-3-butenyl, (2*S*)-2-hydroxy-3-butenyl, benzyl (see also Kliebenstein et al., 2001b), 3-methylbutyl and 4-methylpentyl glucosinolate (Table 1). Most of these compounds are probably methionine-derived, but leucine is the likely precursor of the last two. Curiously, (2*R*)-2-hydroxy-3-butenyl glucosinolate is always accompanied by its (2*S*)-epimer, with the two compounds always occurring in a ratio (2*R*:2*S*) of approximately 1:3 (listed incorrectly in Kliebenstein et al., 2001a). All of these substances were identified by comparison of UV and mass spectra with those reported in the literature (Hogge et al., 1988a; Griffiths et al., 2000; Tolra et al., 2000; Kiddle et al., 2001), and by comparison of HPLC retention times and spectral data with those of standards isolated from other species. Mass spectra were obtained by atmospheric pressure chemical ionization-MS and by electron impact-MS of the TMSi-derivatives (Brown et al., 1991). In addition, the structures of the last two compounds listed were confirmed by GC–MS analysis of their myrosinase hydrolysis products using methods previously described (Brown et al., 1994).

3. Experimental

3.1. Plant material

Arabidopsis thaliana ecotype Columbia (Col-0) seeds were purchased from Lehle Seeds (Round Rock, Texas, USA). Seeds of other ecotypes and the TU1 and TU5 mutant lines were obtained from the Nottingham Arabidopsis Stock Centre, Nottingham, UK (http://nasc.nott.ac.uk/) and the Arabidopsis Biological Resource Center, Columbus, Ohio, USA (http://aims.cps.msu.edu/aims/index.html/). Plants were grown under a mixture of 60 W cool white and broad spectrum bulbs for a 16 h light/8 h dark cycle in a standard soil/vermiculite mixture. [See Kliebenstein et al. (2001a) for complete details.]

3.2. Glucosinolate extraction and desulfation on an anion exchange matrix

Extracts of desulfated glucosinolates were prepared using methods adapted from previously published protocols (Thies, 1979; Graser et al., 2000). For the bulk extract of ecotype Columbia seeds, 16 g seeds were immersed in 300 ml boiling water containing 8 ml 0.3 M $Pb(OAc)_2$–0.3 M $Ba(OAc)_2$ solution for 10 min. After grinding with an Ultraturrax homogenizer, the homogenate was centrifuged at 4000 *g* for 10 min. The pellet was re-extracted with 150 ml water and centrifuged again, and the combined supernatants loaded onto a column of DEAE Sephadex A25 (10 g, Sigma, swelled

in water). The column was rinsed exhaustively with 67% (aq.) MeOH and subsequently with 20 mM MES buffer (pH 5.2). The column was then treated with 4 ml sulfatase solution (prepared as described by Graser et al., 2001) and left to stand overnight. The resulting desulfoglucosinolates were eluted from the column with 40 ml 60% (aq.) MeOH and 40 ml deionized water, and the combined eluent evaporated to dryness with a rotary evaporator. For HPLC separation, samples were redissolved in 8 ml water. Extracts of leaf samples were prepared in the same manner except that, in the initial step, 100 mg of freeze-dried tissue were immersed in 4 ml boiling water (Reintanz et al., 2001).

3.3. Desulfoglucosinolate fractionation and analysis

The desulfoglucosinolate fraction was separated by HPLC on a Hewlett Packard HP 1100 Series system with autosampler, diode-array detector and fraction collector. The procedure employed a C-18, fully-end-capped, reversed phase column (LiChrospher RP-18, 250×4.6 mm i.d., 5 μm particle size, Chrompack) operated at 1 ml min^{-1} and 25 °C. Elution was accomplished with a gradient (solvent A: H_2O, solvent B: MeCN) of 1.5–5% (v/v) B (6 min), 5–7% B (2 min), 7–21% B (10 min), 21–29%B (5 min), and 29–57% B (14 min), followed by a cleaning cycle (57–93% B in 2 min, 5 min hold, 93 to 1.5% B in 3 min, 6 min hold). Eluent was monitored at 229 nm and desulfoglucosinolates were collected with a fraction collector. Fractions were evaporated by use of a rotary evaporator and redissolved in 1 ml water. Purity of the fractions was checked by HPLC-UV, and if necessary fractions were purified further by HPLC using smaller fraction sizes.

3.4. Hydrolysis of benzoate esters

Putative benzoylated desulfoglucosinolates were treated with porcine liver esterase (EC 3.1.1.1, 3.2 M $(NH_4)_2SO_4$ suspension, Sigma). Desulfoglucosinolates (ca. 5–10 μg in 20 μl from purified HPLC fractions) were mixed with 78 μl 50 mM Tris (pH 8.0) and 2 μl esterase suspension (ca. 10 units of activity) in a glass vial and incubated for one hour at room temperature. The total incubation mixture was loaded onto a DEAE Sephadex column and both the flow-through and a 1 ml 60% MeOH eluent were collected and evaporated to dryness. The sample was then resuspended in water and analyzed by HPLC.

3.5. Hydrolysis of intact glucosinolates

Extracts of intact glucosinolates from *A. thaliana* ecotype Columbia seeds were prepared according to Thies (1988). Briefly, 2 g of seed were immersed in 40 ml of boiling water. After homogenization and centrifugation (4000 *g*, 10 min), the supernatant was loaded onto a column of DEAE Sephadex. After washing with water and formic acid/2-propanol/water solution, the intact glucosinolate fraction was eluted from the column with 0.5 M K_2SO_4 directly into ethanol. This fraction was centrifuged, the supernatant evaporated to dryness and the residue redissolved in 1 ml of water. A portion of the extract (100 μl) was treated with myrosinase (thioglucosidase from *Sinapis alba*, EC 3.2.3.1, 859 units g^{-1}, Sigma) by mixing with 700 μl 50 mM Tris (pH 7.0) and 100 μl aqueous myrosinase solution (20 mg ml^{-1}). After 5 min at room temperature, the mixture was boiled for 5 min to denature the enzyme and then applied to a DEAE Sephadex column. Elution of desulfoglucosinolates followed the protocol described above. The resulting sample was analyzed by HPLC-UV and compared with the HPLC trace of a non-hydrolyzed control sample (subjected to the identical procedure but without myrosinase).

3.6. NMR and MS analyses

NMR measurements were carried out on a Bruker DRX 500 spectrometer at 500.13 MHz (^1H), and 125.76 MHz (^{13}C). DMSO-d_6 was used as a solvent and TMS as internal standard, except for compound **7**, which was measured in D_2O (int. std. TMSP-2,2,3,3-d_4). ^1H NMR, ^1H–^1H COSY, and HMQC experiments were recorded in a 2.5 mm inverse detection microprobe head.

Mass spectrometry was performed on a Quattro II (Micromass Ltd, Altrincham, UK) tandem quadrupole mass spectrometer (quadrupole-hexapole-quadrupole geometry) equipped with an atmospheric pressure chemical ionization (APCI) source. The corona pin was operated at 3.5 kV and the sample cone at 16 V. A mixture of acetonitrile and water (1:1) was used as the solvent, at a flow rate of 0.5 ml min^{-1}. Vaporization was achieved with a nitrogen sheath gas (300 l h^{-1}) and drying gas (150 l h^{-1}) at 400 °C. Standard mass spectra were measured using the first quadrupole only. Product ion mass spectra were recorded by setting the first quadrupole to transmit the parent ion of interest and scanning the second quadrupole from m/z 50 to 10 amu above the $[M+H]^+$ of the chosen precursor ion. Argon was used as the collision gas at 2×10^{-3} mbar, and a collision energy of 10–20 eV was employed to achieve fragmentation. For LC/MS analysis, a Hewlett Packard HP 1100 LC was connected to the mass spectrometer. A flow rate of 1 ml min^{-1} was used with 1:1 post-column splitting. The solvent program was identical to that used in the HPLC analysis described above.

Acknowledgements

We thank John Allen for technical assistance, an anonymous reviewer for valuable comments on the

manuscript, and the Deutsche Forschungsgemeinschaft (grant FOR 383 to J.G.) and the Max-Planck Gesellschaft for financial support.

References

Agerbirk, N., Petersen, B.L., Olsen, C.E., Halkier, B.A., Nielsen, J.K., 2001. 1,4-Dimethoxyglucobrassicin in *Barbarea* and 4-hydroxyglucobrassicin in *Arabidopsis* and *Brassica*. Journal of Agricultural and Food Chemistry 49, 1502–1507.

Bjerg, B., Sørensen, H., 1987. Isolation of intact glucosinolates by column chromatography and determination of their purity. In: Wathelet, J.-P. (Ed.), World Crops: Production, Utilization, Description, Vol. 13: Glucosinolates in Rapeseed: Analytical Aspects. Martinus Nijhoff Publishers, Dordrecht/Boston/Lancaster, pp. 59–75.

Brown, P.D., Morra, M.J., Borek, V., 1994. Gas chromatography of allelochemicals produced during glucosinolate degradation in soil. Journal of Agricultural and Food Chemistry 42, 2029–2034.

Brown, P.D., Morra, M.J., McCaffrey, J.P., Auld, D.L., Williams, L., 1991. Allelochemicals produced during glucosinolate degradation in soil. Journal of Chemical Ecology 17, 2021–2034.

Chapple, C.C.S., Decicco, C., Ellis, B.E., 1988. Biosynthesis of 2–2′ methylthioethylmalate in *Brassica carinata*. Phytochemistry 27, 3461–3464.

Chew, F. S., 1988. Biological effects of glucosinolates. In: Cutler, H. G. (Ed.), Biologically Active Natural Products. ACS Symposium Series 380, American Chemical Society, pp. 155–181.

Fahey, J.W., Zalcmann, A.T., Talalay, P., 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56, 5–51.

Graser, G., Oldham, N.J., Brown, P.D., Temp, U., Gershenzon, J., 2001. The biosynthesis of benzoic acid glucosinolate esters in *Arabidopsis thaliana*. Phytochemistry 57, 23–32.

Graser, G., Schneider, B., Oldham, N.J., Gershenzon, J., 2000. The methionine chain elongation pathway in the biosynthesis of glucosinolates in *Eruca sativa* (Brassicaceae). Archives of Biochemistry and Biophysics 378, 411–419.

Griffiths, D.W., Bain, H., Deighton, N., Botting, N.P., Robertson, A.A.B., 2000. Evaluation of liquid chromatography-atmospheric pressure chemical ionisation-mass spectrometry for the identification and quantification of desulphoglucosinolates. Phytochemical Analysis 11, 216–225.

Grob, K., Matile, P., 1980. Capillary GC of glucosinolate-derived horseradish constituents. Phytochemistry 19, 1789–1793.

Halkier, B. A. 1999. Glucosinolates. In: Ikan, R. (Ed.), Naturally Occurring Glycosides: Chemistry Distribution and Biological Properties. John Wiley & Sons, pp. 193–223.

Haughn, G.W., Davin, L., Giblin, M., Underhill, E.W., 1991. Biochemical genetics of plant secondary metabolites in *Arabidopsis thaliana*: The glucosinolates. Plant Physiology 97, 217–226.

Hecht, S.S., 2000. Inhibition of carcinogenesis by isothiocyanates. Drug Metabolism Reviews 32, 395–411.

Hogge, L.R., Reed, D.W., Underhill, E.W., 1988a. The identification of desulfoglucosinolates using thermospray liquid chromatography-mass spectrometry. Journal of Chromatographic Science 26, 348–351.

Hogge, L.R., Reed, D.W., Underhill, E.W., Haughn, G.W., 1988b. HPLC separation of glucosinolates from leaves and seeds of *Arabidopsis thaliana* and their identification using thermospray liquid chromatography-mass spectrometry. Journal of Chromatographic Science 26, 551–556.

Hu, Z., Lewis, J.A., Hanley, A.B., Fenwick, G.R., 1989. 2-Hydroxyethyl glucosinolate from *Capparis masaikai* of Chinese origin. Phytochemistry 28, 1252–1254.

Kiddle, G., Bennett, R.N., Botting, N.P., Davidson, N.E., Robertson, A.A.B., Wallsgrove, R.M., 2001. High-performance liquid chromatographic separation of natural and synthetic desulphoglucosinolates and their chemical validation by UV, NMR and chemical ionisation–MS methods. Phytochemical Analysis 12, 226–242.

Kliebenstein, D.J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., Mitchell-Olds, T., 2001a. Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. Plant Physiology 126, 811–825.

Kliebenstein, D.A., Gershenzon, J., Mitchell-Olds, T., 2001b. Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. Genetics 159, 359–370.

Kroymann, J., Textor, S., Tokuhisa, J., Falk, K., Bartram, S., Gershenzon, J., Mitchell-Olds, T., 2001. A gene controlling variation in *Arabidopsis thaliana* glucosinolate composition is part of the methionine chain elongation pathway. Plant Physiology 127, 1077–1088.

Linscheid, M., Wendisch, D., Strack, D., 1980. The structures of sinapic acid esters and their metabolism in cotyledons of *Raphanus sativus*. Zeitschrift für Naturforschung C 35, 907–914.

Rask, L., Andreasson, E., Ekbom, B., Eriksson, S., Pontoppidan, B., Meijer, J., 2000. Myrosinase: gene family evolution and herbivore defense in Brassicaceae. Plant Molecular Biology 42, 93–113.

Reintanz, B., Lehnen, M., Reichelt, M., Gershenzon, J., Kowalczyk, M., Sandberg, G., Godde, M., Uhl, R., Palme, K., 2001. *Bus*, a bushy *Arabidopsis* CYP79F1 knockout mutant with abolished synthesis of short-chain aliphatic glucosinolates. Plant Cell 13, 351–367.

Thies, W., 1979. Detection and utilization of a glucosinolate sulfohydrolase in the edible snail, *Helix pomatia*. Naturwissenschaften 66, 364–365.

Thies, W., 1988. Isolation of sinigrin and glucotropaeolin from cruciferous seeds. Fat Science and Technology 90, 311–314.

Tolra, R.P., Alonso, R., Poschenrieder, C., Barcelo, D., Barcelo, J., 2000. Determination of glucosinolates in rapeseed and *Thlaspi caerulescens* plants by liquid chromatography-atmospheric pressure chemical ionization mass spectrometry. Journal of Chromatography 889, 75–81.