A galloylated dimeric proanthocyanidin from grape seed exhibits dentin biomodification potential

Rasika S. Phansalkar[a], Joo-Won Nam[a], Shao-Nong Chen[a], James B. McAlpine[a], José G. Napolitano[a], Ariene Leme[b], Christina Vidal[b], Thaiane Aguiar[b], Ana K. Bedran Russo[b], Guido F. Pauli[a,*]

[a] *Department of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, University of Illinois at Chicago, Chicago, IL, USA*

[b] *Department of Restorative Dentistry, College of Dentistry, University of Illinois at Chicago, Chicago, IL, USA*

Supplementary Data

Corresponding author at: UIC College of Pharmacy, 833 S. Wood St., Chicago, IL 60612. Tel.: +1 312 355-1949; fax: +1 312 355-2693.

E-mail address: gfp@uic.edu (G. Pauli)

Table of Contents

S1. The Nomenclature of "Condensed Tannins"

Nomenclature Systems

Nomenclature systems have been initially proposed by Hemingway [1], then extended by Porter [1], and applied subsequently by Ferreira et al. [2, 3]. A summary of this nomenclature is provided below, along with explanations how it is applied to the present study.

Leuco- vs. Pro-

As per the following definitions, the prefix pro- already implies the non-monomeric nature (DP>1) of a compound.

Leucoanthocyanidins	monomeric compounds which produce anthocyanidins by cleavage of a C–O bond on heating with mineral acid
Proanthocyanidins	oligomers/polymers which give anthocyanidins by cleavage of a C–C bond under strongly acidic conditions in the presence of molecular oxygen

However, it should be noted that the prefix pro- does not specify whether a compound is oligomeric or polymeric. This means that the use of "oligo-" can be justified to differentiate oligomers from polymers.

Oligomeric "condensed tannins" can be di-, tri-, tetra-, penta-, hexa-, hepta-, octa-, or nona-mers, etc. To our best knowledge, no clear definition exists for the numerical cutoff of "oligo-" vs. "poly-" representing the degree of polymerization (DP).

In general, oligomers with DP values of 2 or 3 are commonly reported in the literature. Oligomers with DP = 4 are much more uncommon, and those with DP≥4 are very rarely reported. If, in the same analogy to carbohydrate chemistry that led to the initial nomenclatural proposal by Hemingway, carbohydrate conventions would guide the nomenclature of proanthocyanidins, compounds with DP = 2-9 could be considered as "oligo-" vs. DP≥10 as "poly-". However, the IUPAC definitions are vague, stating that oligomers are comprised of "a small plurality of units" (http://goldbook.iupac.org/O04286.html) [4].

Procyanidins vs. Proanthocyanidins

The term "proanthocyanidin" is general and encompasses numerous groups of flavan-3-ol derivatives with different substitution patterns such as the procyanidins (which are the by far most common subgroup), prodelphinidins, propelargonidins, and several other classes (see [3]

for an overview of a total of 16 known classes). Thus, "proanthocyanidin" is the more inclusive term than "procyanidins".

Whenever referring to (groups of) condensed tannins that cover more than just the (most common) "procyanidin" subclass of "proanthocyanidins", such as in an extract of (mostly or partially) unknown constitution, it is necessary to use the term "proanthocyanidins" (PACs) rather than just procyanidins (PCs).

Monomers & Linkages

In the literature, monomers are consistently referred to as flavan-3-ols or flavanyl units.

The linkage between multiple flavan-3-ol units (monomers) can be designated as "interflavanyl linkage", "interflavanyl bond", or "interflavonoid linkage".

The term "C-C linked" refers to B-type linkages.

The term "C-C and C-O-C doubly linked" refers to A-type linkages.

Acronyms

In the literature on condensed tannins, the use of acronyms is largely inconsistent. Frequently, publications including authoritative reviews do not make use acronyms.

A rationale for the acronym in the present study is presented in the summary below.

It shall be noted that acronyms that are pronounceable should generally be preferred. For example, the acronym "OPAC" for an "oligomeric proanthocyanidin" is easier to pronounce than "OPC" for an "oligomeric procyanidin". Considering that the first group actually includes the second, the first is the more practical term and can substitute the second - unless a statement is specifically geared towards the distinction of the various substitution pattern of the flavan-3-ol units.

Summary of Nomenclature Used in the Present Study

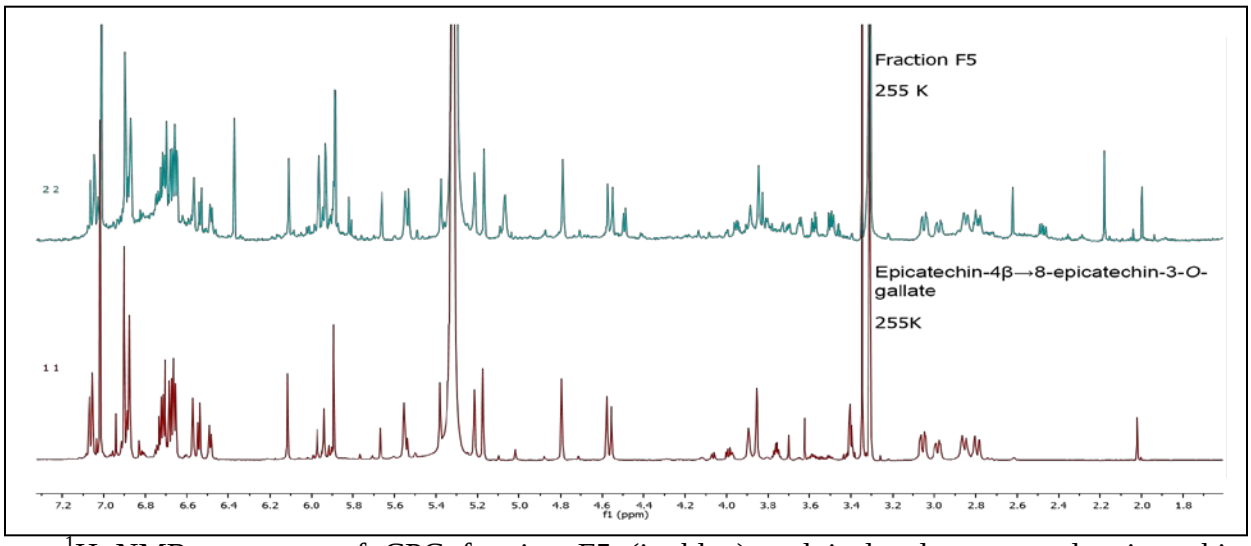
The nomenclature and acronyms used in the present study follow the following rationales:

- Use of the general term "proanthocyanidins" to cover the full breadth of all "condensed tannins" (which is the less desirable synonym due to potential confusion with "hydrolyzable tannins") contained in the study plant.
- Grape seed extracts (GSE) contains both procyanidins and prodelphinidins. Hence, when referring to all condensed tannins in GSE, the proper term and abbreviation to

use is proanthocyanidins (PACs). However, as the structures described in this study are all procyanidins, it would also be correct to refer to them as procyanidins (PCs).

- Differentiate between oligomers and polymers as the two major subgroups of proanthocyanidins. Considering that the structural complexity of oligomers increases exponentially with the degree of polymerization (see calculations in [5]), we consider a cutoff of DP<9 as oligomers vs. DP≥10 for polymers as adequate.
- Build acronyms using the key letters in the respective words:
    - PAC for proanthocyanidin
    - PC for procyanidins
    - OPAC for oligomeric proanthocyanidin
    - OPC for oligomeric procyanidin
- Add "s" for plural use of the acronyms; e.g. OPACs, PACs etc.

*S2. ^{1}H NMR Spectra of Fraction F5 vs. Isolated Compound **1***



^{1}H NMR spectrum of CPC fraction F5 (in blue) and isolated compound epicatechin-(4β→8)-epicatechin-3-*O*-gallate (**1**, in red) stacked using MestReNova NMR processing software. Both spectra were recorded in methanol-d_4 on an 800 MHz NMR instrument at 255K.

*S3. Quantitation of **1** in Fraction F5 by qHNMR*

Calculation of the qHNMR Purity (% w/w) of epicatechin-(4β→8)-epicatechin-3-*O*-gallate in fraction F5 was performed using the 100% method as follows:

QUANTITATIVE ANALYSIS										
100% METHOD										
Fraction F5	**Epicatechin-(4β→8)-epicatechin-3-*O*-gallate (1)**	(integration in MestReNova)								
Target	**signals**	**int**	**# H**	**int 1H**	**average 1H**	**MW**	***F (Molar Mass ratio)***	**average mass**	**δH [ppm]**
		100.00	1	100.00	102.12	730.63	1.00	**102.12**	7.046 (d)
		111.54	1	111.54					2.85 (dd)
		96.25	1	96.25					3.05 (dd)
		88.77	1	88.77					3.846 (brs)
		85.57	1	85.57					4.788 (brs)
		85.67	1	85.67					5.167 (brs)
		99.92	1	99.92					5.216 (brs)
		86.68	1	86.68					5.548 (ddd)
		100.0	1	100.00					5.886 (s)
		200.09	2	100.05					7.012 (s)
		224.41	2	112.21					6.87, 6.88
		143.43	1	143.43					6.67 (d)
		143.36	1	143.36					6.70 (d)
		93.38	1	93.38					5.932 (d)
		85.00	1	85					5.96 (d)
	Impurities								
Structurally related*		**14.11**	1	14.11	32.75	730.63	1.00	32.75	1.39 (dd)
		3.58	1	3.58					1.44 (d)
		44.96	1	44.96					2.48 (dd)
		27.87	1	27.87					2.80 (dd)
		14.48	1	14.48					3.58 (dd)
		67.4	1	67.42					3.50 (dd)
		53.61	1	53.61					2.481 (m)
		25.87	1	25.87					3.826 (s)
		39.45	2	19.73					4.497,4.487
		91.9	1	91.85					5.069 (d)
		45.26	1	45.26					5.53 (d)
		45.48	1	45.48					5.66 (d)
		33.68	1	33.68					5.82 (d)
		69.0	2	34.49					5.89, 5.90
		29.30	1	29.30					5.94
		14.82	1	14.82					5.98
		10.29	1	10.29					6.01
		56.86	2	28.43					6.371 (s)
		33.14	1	33.14					6.74
		39.57	1	39.57					6.75
		19.46	2	9.73					7.028 (s)
Unrelated impurities	Methyl acetate	22.87	3	7.62	8.79	74.08	0.10	0.89	2.00 (s)
		29.89	3	9.96					2.621 (s)
	Ethyl acetate	85.13	3	28.38	28.38	88.11	0.39	11.07	1.24 (s)
	n-hexane	42.00	2	21.00	15.05	86.18	0.12	1.78	0.90
		27.32	3	9.11					0.938 (t)
	Acetone	37.10	3	12.37	4.10	58.08	0.08	0.33	2.15 (s)
								46.8	

	Mass (% w/w)
1H pure cmpd	102.12
sum 1H all imp	46.80
% purity	**68.57**
% impurity	**31.43**

*S4. Determination of the Purity of **1** by qHNMR*

Calculation of qHNMR purity (% w/w) of **1** in its isolated form and determination of the relative ratio of major and minor rotamers. Purity of the isolated compound was found to be 88.5% and the relative ratio of major and minor rotamers was 62:38.

QUANTITATIVE ANALYSIS									
100% METHOD									
Isolated compound	Epicatechin-(4β→8)-epicatechin-3-*O*-gallate (**1**)	(integration in MestReNova)							
Target	**Major rotamer**	**int**	**# H**	**int 1H**	**average 1H**	**MW**	***F (Molar Mass ratio)***	**average mass**	**δH [ppm]**
	signals	100.00	1	100.00	101.86	730.63	1.00	101.86	7.046 (d)
		107.95	1	107.95					2.85 (dd)
		103.90	1	103.90					3.05 (dd)
		96.23	1	96.23					3.85 (brs)
		89.95	1	89.95					4.79 (brs)
		102.21	1	102.21					5.17 (brs)
		109.87	1	109.87					5.22 (brs)
		101.60	1	101.60					5.55 (ddd)
		79.7	1	79.71					5.89 (s)
		197.12	2	98.56					7.01 (s)
		220.53	2	110.27					6.87, 6.88
		148.74	1	148.74					6.65 (d)
		112.11	1	112.11					6.67 (d)
		77.63	1	77.63					5.93 (d)
		75.15	1	75.15					5.96 (d)
		115.88	1	115.88					6.70 (d)
	Minor rotamer								
		69.12	1	69.12	62.38	730.63	1.00	62.38	2.78 (dd)
		73.05	1	73.05					2.97 (dd)
		55.50	1	55.50					3.89 (brs)
		59.20	1	59.20					4.55 (brs)
		66.72	1	66.72					4.57 (s)
		77.01	1	77.01					5.37 (brs)
		50.03	1	50.03					5.53 (d)
		45.3	1	45.30					5.66 (d)
		50.59	1	50.59					6.11 (s)
		65.53	1	65.53					6.48 (dd)
		65.29	1	65.29					6.53 (d)
		63.40	1	63.40					6.56 (d)
		69.69	1	69.69					6.72 (d)
		125.96	2	62.98					6.89 (s)
		62.25	1	62.25					7.06 (d)
	Impurities								
	Structurally related	9.37	1	9.37	15.98	730.63	1.00	15.98	1.39 (dd)
		9.49	1	9.49					1.44 (d)
		29.08	1	29.08					2.48 (dd)
Unrelated impurities	Ethyl acetate	86.28	3	28.76	25.42	88.11	0.12	3.07	1.24 (t)
		48.45	2	24.23					4.09 (q)
		69.85	3	23.28					2.02 (s)
	Acetonitrile	125.56	3	41.85	41.85	41.05	0.06	2.35	2.03 (s)
								21.40	

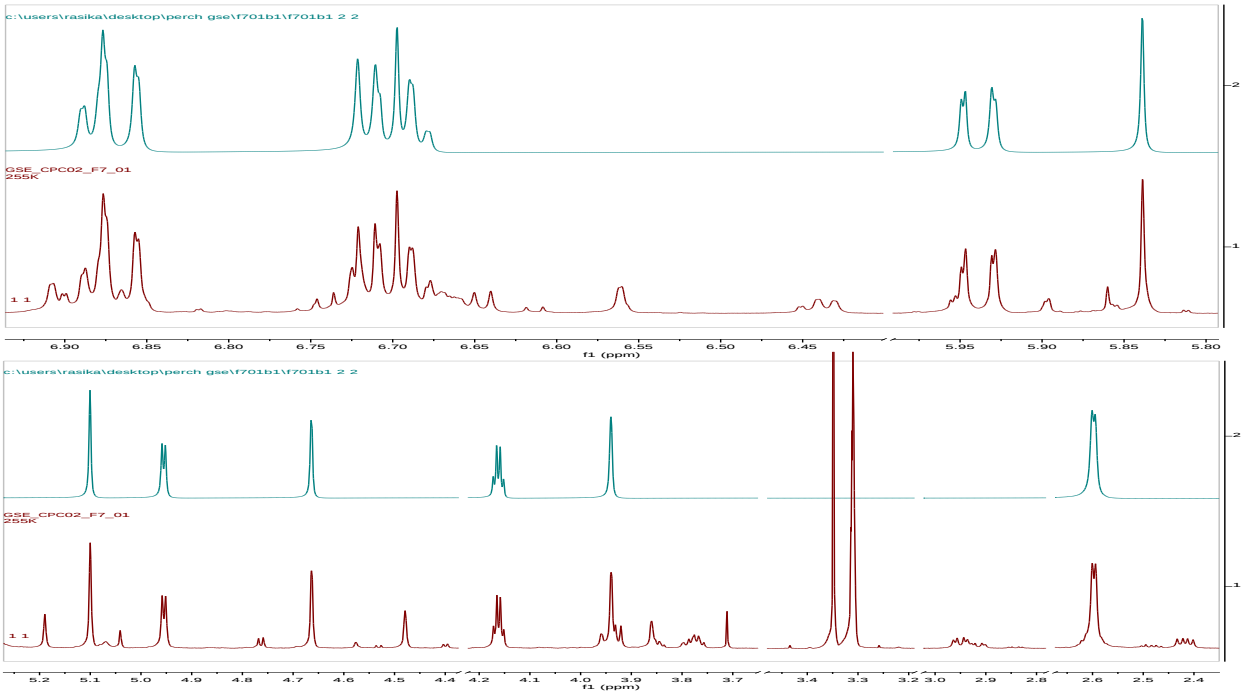
Mass (%w/w)			
1H Major rotamer	101.86		
1H Minor rotamer	62.38		
1H compound	164.24		
sum 1H all imp	21.40		
% Major rotamer	**54.87**	**Relative % of major and minor rotamers**	
% Minor rotamer	**33.60**	**62.019**	
% purity	**88.47**	**37.981**	
% impurity	**17.36**		

*S5. The HiFSA Profile of **1** in PERCH .pms Format*

```
* The couplings can be fixed in the same way

NMR-data: c:\users\rasika\desktop\perch gse\f403\f403betterf701\perch
#$œ  Date 17. 6.2014;  Time 20:32:42     perch.pms

CHEMICAL SHIFTS(PPM):
PROTON    2*SPIN= 1 SPECIES=1H    POPULATION(Y)=   1.00000
 H1    / 1     5.930642  1*1*1  STAT=Y  PRED= 5.909 RANGE= 0.254 WIDTH(Y)= 2.036 RESP(Y)= 0.3855 HSQC= C1
 H5    / 1     5.963556  1*1*1  STAT=Y  PRED= 5.883 RANGE= 0.230 WIDTH(Y)= 1.905 RESP(Y)= 0.3585 HSQC= C5
 H8    / 1     5.160737  1*1*1  STAT=Y  PRED= 4.944 RANGE= 0.375 WIDTH(Y)= 2.322 RESP(Y)= 0.4917 HSQC= C8
 H9    / 1     3.845153  1*1*1  STAT=Y  PRED= 3.968 RANGE= 0.735 WIDTH(Y)= 3.393 RESP(Y)= 0.5510 HSQC= C9
 H10   / 1     4.784641  1*1*1  STAT=Y  PRED= 4.766 RANGE= 0.499 WIDTH(Y)= 3.750 RESP(Y)= 0.4597 HSQC= C10
 H12   / 1     6.652900  1*1*1  STAT=Y  PRED= 6.881 RANGE= 0.110 WIDTH(Y)= 2.263 RESP(Y)= 1.0000 HSQC= C12
 H13   / 1     6.700888  1*1*1  STAT=Y  PRED= 6.689 RANGE= 0.219 WIDTH(Y)= 3.500 RESP(Y)= 0.7114 HSQC= C13
 H16   / 1     6.866147  1*1*1  STAT=Y  PRED= 7.031 RANGE= 0.290 WIDTH(Y)= 1.800 RESP(Y)= 0.5070 HSQC= C16
 H24A  / 1     3.046660  1*1*1  STAT=Y  PRED= 3.099 RANGE= 0.315 WIDTH(Y)= 4.081 RESP(Y)= 0.5698 HSQC= C24
 H24B  / 1     2.845180  1*1*1  STAT=Y  PRED= 2.678 RANGE= 0.380 WIDTH(Y)= 7.247 RESP(Y)= 0.6592 HSQC= C24
 H25   / 1     5.547437  1*1*1  STAT=Y  PRED= 5.408 RANGE= 0.445 WIDTH(Y)= 3.543 RESP(Y)= 0.6041 HSQC= C25
 H26   / 1     5.215290  1*1*1  STAT=Y  PRED= 5.447 RANGE= 0.420 WIDTH(Y)= 5.834 RESP(Y)= 0.7110 HSQC= C26
 H30   / 1     5.882931  1*1*1  STAT=Y  PRED= 5.914 RANGE= 0.575 WIDTH(Y)= 1.923 RESP(Y)= 0.4096 HSQC= C30
 H34   / 1     7.042530  1*1*1  STAT=Y  PRED= 6.762 RANGE= 0.250 WIDTH(Y)= 2.769 RESP(Y)= 0.5084 HSQC= C34
 H37   / 1     6.672773  1*1*1  STAT=Y  PRED= 6.736 RANGE= 0.345 WIDTH(Y)= 2.607 RESP(Y)= 0.5643 HSQC= C37
 H38   / 1     6.875800  1*1*1  STAT=Y  PRED= 7.031 RANGE= 0.455 WIDTH(Y)= 2.700 RESP(Y)= 0.7642 HSQC= C38
 H42_46/ 1     7.009321  1*2*1  STAT=Y  PRED= 7.018 RANGE= 0.150 WIDTH(Y)= 2.195 RESP(Y)= 0.4825 HSQC= C42_46

COUPLING CONSTANTS(HZ):
 J54_55      2.0833   J H1     H5       STAT=Y  PRED= 2.230 RANGE= 0.890
 J56_57      1.4748   J H8     H9       STAT=Y  PRED= 5.320 RANGE= 2.800
 J56_59     -0.7238   J H8     H12      STAT=Y  PRED= -0.730 RANGE= 0.200
 J56_61     -0.8716   J H8     H16      STAT=Y  PRED= -0.740 RANGE= 0.200
 J57_58      1.8927   J H9     H10      STAT=Y  PRED= 10.290 RANGE= 2.000
 J59_60      8.2279   J H12    H13      STAT=Y  PRED= 8.260 RANGE= 0.500
 J59_61      2.1026   J H12    H16      STAT=Y  PRED= 2.030 RANGE= 0.800
 J60_61      0.0914   J H13    H16      STAT=Y  PRED= 0.430 RANGE= 0.320
 J67_68    -17.0104   J H24A   H24B     STAT=Y  PRED= -14.750 RANGE= 1.280
 J67_69      4.6476   J H24A   H25      STAT=Y  PRED= 5.080 RANGE= 2.800
 J67_70      0.2091   J H24A   H26      STAT=Y  PRED= 2.440 RANGE= 0.700
 J68_69      3.7577   J H24B   H25      STAT=Y  PRED= 11.140 RANGE= 2.000
 J69_70      1.0413   J H25    H26      STAT=Y  PRED= 5.890 RANGE= 2.800
 J72_73      0.1809   J H34    H37      STAT=Y  PRED= 0.430 RANGE= 0.320
 J72_74      1.8655   J H34    H38      STAT=Y  PRED= 2.030 RANGE= 0.800
 J73_74      8.3139   J H37    H38      STAT=Y  PRED= 8.260 RANGE= 0.500
 J77_76      1.5300   J H42_46 H42_46   STAT=Y  PRED= 1.530 RANGE= 0.850

CONTROL PARAMETERS:
       Solvent =  none (def. 99% enriched)
         1.000 =  Concentration (vol%, def=1.0%)
    0.00100000 =  Minimum line-intensity
    0.00100000 =  Diagonalization criterium (not in use)
  800.20400000 =  FIELD(1H,MHz), used to transform shifts to ppms
   12.26565676 =  Left frequency (ppm)
   -1.68167678 =  Right frequency (ppm)
        10.000 =  Acquisition time (s, for QMTLS)
         0.000 =  Line-width (for modes D, P & T, 0=use defaults)
   0.095260478 =  Data-point resolution (Hz)
        -6.260 =  GAUSSIAN (%, 0=use default from INF)
         3.269 =  Dispersion contribution (%, 0=use default from INF)
    0.00000000 =  Decoupling frequency (for DORES)

END of FILE
```

*S6. The HiFSA Fingerprint of compound **2***

Simulated and experimental spectra of compound **2**, procyanidin B1 (PA B1); stacked plots in MestReNova. The simulated spectrum is shown in green and the experimental spectrum in red. Iteration was performed using both the D-mode (overall integral fitting) and the T-mode (total line shape fitting) modules in PERCH NMR software. The experimental ^{1}H NMR spectrum was recorded on an 800 MHz Bruker instrument at 255 K. The 1D and 2D NMR data along with NMR simulation and literature review allowed the conclusion that compound **2** is procyanidin B1.



*S7. The ^{1}H NMR Data of **2** and its HiFSA profile in PERCH .pms Format*

The table consists of the proton chemical shifts and coupling constants for procyanidin B1 (compound **2**) as generated by HiFSA using the PERCH software tool. The NMR signals for protons U-2, U-3, and U-4 appear as broad singlets, and the signals for protons L-2', L-5' appear as doublets. The small coupling constants are derived from the calculated spectral parameters that resulted from the HiFSA.

Unit	**Position**	**Proton chemical shifts (δ [ppm]), coupling constants and multiplicities (*J* [Hz]); mult.)**
		Epicatechin-4β→8-catechin
U	2	5.1004 (0.99; d)
	3	3.9400 (0.99, 1.9; dd)
	4	4.6638 (1.90; d)
	6	5.9294 (2.04; d)
	8	5.9479 (2.04; d)
	2'	6.8753 (2.12; d)
	5'	6.7008 (8.22; d)
	6'	6.6849 (2.12, 8.22; dd)
L	2	4.9551 (5.32; d)
	3	4.1620 (5.32, 5.15, 5.92; ddd)
	4a	2.5998 (-14.55, 5.15; dd)
	4b	2.5958 (-14.55, 5.92; dd)
	6	5.8387 (s)
	2'	6.8560 (2.10, 0.30; dd)
	5'	6.7160 (8.60, 0.30; dd)
	6'	6.8834 (2.10, 8.60; dd)

```
* NEW: the lines beginning by * are comment lines !
* To keep all the chemical shifts fixed during iteration
* replace "CHEMICAL SHIFTS(HZ):" by "..SHIFTS(HZ): fixed"
* The couplings can be fixed in the same way

NMR-data: c:\users\rasika\desktop\perch gse\f701b1\f701b1
#$œ  Date  8. 9.2014;  Time 14:34: 7     f701b1.pms

CHEMICAL SHIFTS(PPM):
PROTON    2*SPIN= 1 SPECIES=1H    POPULATION(Y)=   1.00000
 H1  / 1     5.838731  1*1*1  STAT=Y  PRED= 5.963 RANGE= 0.584 WIDTH(Y)= 1.818 RESP(Y)= 0.5239 HSQC= C1
 H8  / 1     4.955129  1*1*1  STAT=Y  PRED= 4.377 RANGE= 0.390 WIDTH(Y)= 3.511 RESP(Y)= 0.6540 HSQC= C8
 H9  / 1     4.162085  1*1*1  STAT=Y  PRED= 3.843 RANGE= 0.290 WIDTH(Y)= 2.645 RESP(Y)= 0.6822 HSQC= C9
 H10A/ 1     2.599756  1*1*1  STAT=Y  PRED= 2.878 RANGE= 0.254 WIDTH(Y)= 9.166 RESP(Y)= 0.8162 HSQC= C10
 H10B/ 1     2.595805  1*1*1  STAT=Y  PRED= 2.525 RANGE= 0.285 WIDTH(Y)= 3.211 RESP(Y)= 0.8195 HSQC= C10
 H14 / 1     6.883379  1*1*1  STAT=N  PRED= 6.685 RANGE= 0.275 WIDTH(N)= 2.700 RESP(Y)= 0.7357 HSQC= C14
 H15 / 1     6.715955  1*1*1  STAT=N  PRED= 6.602 RANGE= 0.325 WIDTH(N)= 2.800 RESP(Y)= 1.0000 HSQC= C15
 H18 / 1     6.856041  1*1*1  STAT=Y  PRED= 6.764 RANGE= 0.275 WIDTH(Y)= 2.470 RESP(Y)= 0.6725 HSQC= C18
 H21 / 1     4.663829  1*1*1  STAT=Y  PRED= 4.535 RANGE= 0.494 WIDTH(N)= 3.000 RESP(Y)= 0.6040 HSQC= C21
 H23 / 1     5.100356  1*1*1  STAT=Y  PRED= 5.196 RANGE= 0.350 WIDTH(Y)= 3.110 RESP(Y)= 0.6904 HSQC= C23
 H24 / 1     3.940097  1*1*1  STAT=Y  PRED= 4.144 RANGE= 0.424 WIDTH(Y)= 3.397 RESP(Y)= 0.6839 HSQC= C24
 H28 / 1     5.929490  1*1*1  STAT=Y  PRED= 5.946 RANGE= 0.205 WIDTH(Y)= 2.240 RESP(Y)= 0.4553 HSQC= C28
 H30 / 1     5.947903  1*1*1  STAT=Y  PRED= 5.790 RANGE= 0.495 WIDTH(Y)= 1.878 RESP(Y)= 0.3829 HSQC= C30
 H34 / 1     6.684962  1*1*1  STAT=Y  PRED= 6.962 RANGE= 0.120 WIDTH(N)= 2.700 RESP(Y)= 0.7385 HSQC= C34
 H35 / 1     6.700775  1*1*1  STAT=N  PRED= 6.690 RANGE= 0.154 WIDTH(Y)= 1.918 RESP(Y)= 0.6613 HSQC= C35
 H38 / 1     6.875268  1*1*1  STAT=Y  PRED= 6.644 RANGE= 0.265 WIDTH(Y)= 2.268 RESP(Y)= 0.6019 HSQC= C38

COUPLING CONSTANTS(HZ):
 J44_45      5.3166   J H8   H9     STAT=Y  PRED= 10.200 RANGE= 2.000
 J45_46      5.1516   J H9   H10A   STAT=Y  PRED= 8.150 RANGE= 3.000
 J45_47      5.9213   J H9   H10B   STAT=Y  PRED= 7.400 RANGE= 3.000
 30          0.0001   J H9   H15    STAT=Y
 32          0.0000   J H9   H34    STAT=Y
 31         -0.0001   J H9   H35    STAT=Y
 33          0.0000   J H9   H38    STAT=Y
 J46_47    -14.5538   J H10A H10B   STAT=Y  PRED= -14.740 RANGE= 1.280
 J50_51      8.6000   J H14  H15    STAT=N  PRED= 8.260 RANGE= 0.500
 J50_52      2.1000   J H14  H18    STAT=N  PRED= 2.030 RANGE= 0.800
 J51_52      0.3000   J H15  H18    STAT=N  PRED= 0.430 RANGE= 0.320
 J55_57      1.9000   J H21  H24    STAT=N  PRED= 1.710 RANGE= 2.200
 J56_57      0.9861   J H23  H24    STAT=N  PRED= 1.200 RANGE= 2.200
 J58_59      2.0420   J H28  H30    STAT=Y  PRED= 2.230 RANGE= 0.890
 J62_63      8.2206   J H34  H35    STAT=N  PRED= 8.260 RANGE= 0.500
 J62_64      2.1163   J H34  H38    STAT=N  PRED= 2.030 RANGE= 0.800
 J63_64      0.0620   J H35  H38    STAT=Y  PRED= 0.430 RANGE= 0.320

CONTROL PARAMETERS:
       Solvent =  none (def. 99% enriched)
         1.000 =  Concentration (vol%, def=1.0%)
    0.00100000 =  Minimum line-intensity
    0.00100000 =  Diagonalization criterium (not in use)
  800.20400000 =  FIELD(1H,MHz), used to transform shifts to ppms
   12.26589229 =  Left frequency (ppm)
   -1.68144383 =  Right frequency (ppm)
        10.000 =  Acquisition time (s, for QMTLS)
         0.000 =  Line-width (for modes D, P & T, 0=use defaults)
   0.090899999 =  Data-point resolution (Hz)
        10.624 =  GAUSSIAN (%, 0=use default from INF)
         7.267 =  Dispersion contribution (%, 0=use default from INF)
    0.00000000 =  Decoupling frequency (for DORES)


END of FILE
```

*S8. Calculation of the Purity of **2** by qHNMR*

Calculation of the qHNMR purity (% w/w) of **2** using the 100 % method. The purity was determined to be 74 %.

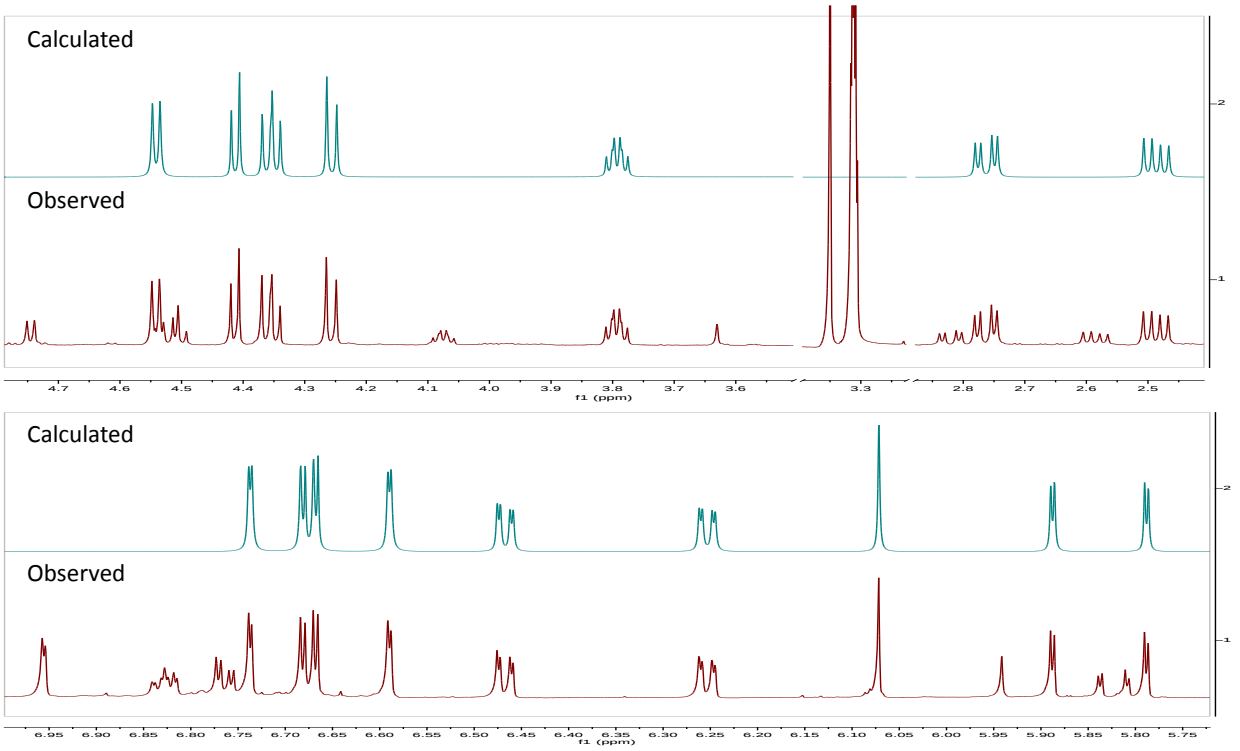
QUANTITATIVE ANALYSIS									
100% METHOD									
Isolated compound	Epicatechin-(4β→8)-catechin (2)	(integration in MestReNova)							
	signals	int	# H	int 1H	average 1H	MW	*F (Molar Mass ratio)*	average mass	δH [ppm]
Target		251.02	2	125.51	98.39	578.52	1.00	98.39	2.62 (d)
		101.17	1	101.17					3.96 (brs)
		107.94	1	107.94					4.18 (ddd)
		95.26	1	95.26					4.68 (brs)
		100.00	1	100.00					4.98 (d)
		108.91	1	108.91					5.12 (brs)
		90.39	1	90.39					5.86 (s)
		71.38	1	71.38					5.95 (ddd)
		61.32	1	61.32					5.97 (s)
		212.65	2	106.33					6.68, 6.70 (dd,d)
		108.68	1	108.68					6.74 (d)
		96.66	1	96.66					6.88 (d)
		211.05	2	105.53					6.88,6.87 (d,d)
	Impurities								
	Structurally related	40.07	1	40.07	33.11	578.82	1.00	33.11	2.44 (dd)
		39.80	1	39.80					2.97 (dd)
		44.18	1	44.18					3.79 (brs)
		31.00	1	31.00					3.88 (brs)
		46.82	1	46.82					4.50 (brs)
		36.60	1	36.60					5.21 (brs)
		28.55	1	28.55					6.10 (s)
		24.12	1	24.12					5.71 (d)
		30.54	1	30.54					3.95 (d)
		21.59	1	21.59					5.43 (d)
		31.34	1	31.34					6.46 (dd)
		33.47	1	33.47					6.58 (d)
		31.57	1	31.57					6.67 (d)
		34.67	1	34.67					6.93 (d)
		22.34	1	22.34					6.76 (d)
Unr	Unknown	19.05	3	6.35	5.64	74.08	0.13	0.72	3.71 (d)
		4.92	1	4.92					8.09 (8)
	Acetonitrile	5.81	3	1.94	1.94	41.05	0.07	0.14	2.03 (s)
								33.97	

Mass (%w/w)	
1H compound	98.39
sum 1H all imp	33.97
% purity	**74.34**
% impurity	**25.66**

NOTE: Multiple rotamers were not observed in case of compound **2**. The PAC like signals observed in parallel are not the signals from the rotamers, but from the structurally related impurities. This can be confirmed from a considerably large difference observed between the coupling constants of the comparable protons.

*S9. The HiFSA Fingerprint of Compound **3***

Simulated and experimental spectra of compound **3**, procyanidin B3 (PA B3), using stacked plots in MestReNova software. The final simulated spectrum is shown in green and the experimental spectrum in red. The iteration was performed using the D-mode (overall integral fitting) and T-mode (total line shape fitting) modules in the PERCH NMR software. The experimental ^{1}H NMR spectrum was recorded on an 800 MHz Bruker instrument at 255 K. The 1D and 2D NMR data along with the NMR HiFSA results and review of the literature allowed the conclusion that compound **3** is procyanidin B3.



*S10. The ^{1}H NMR Data of compound **3** and the HiFSA Profile in PERCH .pms Format*

The table consists of the proton chemical shifts and coupling constants for procyanidin B3 (compound 3) as generated by HiFSA of the experimental spectrum. The NMR signals for protons U-2, U-3, and U-4 appear as broad singlets due to small coupling constants, and the signals for protons L-2’ and L-5’ appear as doublets. The small coupling constants were extracted from the experimental spectra via HiFSA.

Unit	**Position**	**Proton chemical shifts (δ [ppm]), coupling constants and multiplicities (*J* [Hz]); mult.)**
		Catechin-4*α*→8-catechin
U	2	4.25648 (9.84; d)
	3	4.35403 (9.84, 7.98; dd)
	4	4.41154 (7.98; d)
	6	5.78828 (2.39; d)
	8	5.88762 (2.39; d)
	2'	6.73700 (2.05; d)
	5'	6.67182 (8.13; d)
	6'	6.46736 (2.05, 8.13; dd)
L	2	4.54091 (7.39; d)
	3	3.79230 (7.39, 5.50, 8.04; ddd)
	4a	2.48719 (-16.28, 8.04; dd)
	4b	2.76191 (-16.28, 5.50; dd)
	6	6.07157 (s)
	2'	6.58963 (2.05; d)
	5'	6.67681 (8.20; d)
	6'	6.25342 (2.052, 8.20; dd)

```
* NEW: the lines beginning by * are comment lines !
* To keep all the chemical shifts fixed during iteration
* replace "CHEMICAL SHIFTS(HZ):" by "..SHIFTS(HZ): fixed"
* The couplings can be fixed in the same way

NMR-data: c:\users\rasika\desktop\perch gse\f704\f704
#$œ  Date 27. 8.2014;  Time 13:40:13     f704.pms

CHEMICAL SHIFTS(PPM):
PROTON    2*SPIN= 1 SPECIES=1H    POPULATION(Y)=   1.00000
 H2  / 1     5.788289  1*1*1  STAT=Y  PRED= 5.747 RANGE= 0.782 WIDTH(Y)= 1.268 RESP(Y)= 0.4531 HSQC= C2
 H6  / 1     5.887627  1*1*1  STAT=Y  PRED= 6.022 RANGE= 0.157 WIDTH(Y)= 1.342 RESP(Y)= 0.4853 HSQC= C6
 H7  / 1     4.411540  1*1*1  STAT=Y  PRED= 4.571 RANGE= 0.492 WIDTH(Y)= 1.471 RESP(Y)= 0.6368 HSQC= C7
 H8  / 1     4.354036  1*1*1  STAT=Y  PRED= 4.647 RANGE= 0.460 WIDTH(Y)= 1.777 RESP(Y)= 1.0000 HSQC= C8
 H9  / 1     4.256489  1*1*1  STAT=Y  PRED= 4.848 RANGE= 0.375 WIDTH(Y)= 1.647 RESP(Y)= 0.7313 HSQC= C9
 H14 / 1     6.071572  1*1*1  STAT=Y  PRED= 5.967 RANGE= 0.567 WIDTH(Y)= 1.351 RESP(Y)= 0.4925 HSQC= C14
 H18A/ 1     2.487190  1*1*1  STAT=Y  PRED= 2.505 RANGE= 0.287 WIDTH(Y)= 2.187 RESP(Y)= 0.7385 HSQC= C18
 H18B/ 1     2.761918  1*1*1  STAT=Y  PRED= 2.586 RANGE= 0.340 WIDTH(Y)= 2.181 RESP(Y)= 0.7791 HSQC= C18
 H19 / 1     3.792300  1*1*1  STAT=Y  PRED= 3.788 RANGE= 0.365 WIDTH(Y)= 1.964 RESP(Y)= 0.7371 HSQC= C19
 H20 / 1     4.540919  1*1*1  STAT=Y  PRED= 4.399 RANGE= 0.422 WIDTH(Y)= 2.376 RESP(Y)= 0.8511 HSQC= C20
 H26 / 1     6.467369  1*1*1  STAT=Y  PRED= 6.566 RANGE= 0.545 WIDTH(Y)= 1.451 RESP(Y)= 0.6922 HSQC= C26
 H27 / 1     6.671827  1*1*1  STAT=Y  PRED= 6.665 RANGE= 0.327 WIDTH(Y)= 1.156 RESP(Y)= 0.5398 HSQC= C27
 H30 / 1     6.737001  1*1*1  STAT=Y  PRED= 7.126 RANGE= 0.287 WIDTH(Y)= 1.725 RESP(Y)= 0.7282 HSQC= C30
 H32 / 1     6.253421  1*1*1  STAT=Y  PRED= 6.535 RANGE= 0.260 WIDTH(Y)= 1.669 RESP(Y)= 0.6894 HSQC= C32
 H33 / 1     6.676815  1*1*1  STAT=Y  PRED= 6.690 RANGE= 0.190 WIDTH(Y)= 1.743 RESP(Y)= 0.8405 HSQC= C33
 H36 / 1     6.589638  1*1*1  STAT=Y  PRED= 6.684 RANGE= 0.835 WIDTH(Y)= 1.752 RESP(Y)= 0.7009 HSQC= C36

COUPLING CONSTANTS(HZ):
 J43_44      2.3897   J H2   H6     STAT=Y  PRED= 2.230 RANGE= 0.890
 J45_46      7.9841   J H7   H8     STAT=Y  PRED= 1.830 RANGE= 2.200
 J45_47      0.0038   J H7   H9     STAT=Y  PRED= 1.250 RANGE= 0.440
 J46_47      9.8440   J H8   H9     STAT=Y  PRED= 2.050 RANGE= 2.200
 J47_57     -0.3492   J H9   H26    STAT=Y  PRED= -0.770 RANGE= 0.200
 J47_59      0.0000   J H9   H30    STAT=Y  PRED= -0.790 RANGE= 0.200
 J49_50    -16.2756   J H18A H18B   STAT=Y  PRED= -14.830 RANGE= 1.280
 J49_51      8.0358   J H18A H19    STAT=Y  PRED= 2.180 RANGE= 2.200
 J49_52      0.0037   J H18A H20    STAT=Y  PRED= 1.650 RANGE= 0.550
 J50_51      5.5012   J H18B H19    STAT=Y  PRED= 4.230 RANGE= 2.800
 J51_52      7.3873   J H19  H20    STAT=Y  PRED= 1.620 RANGE= 2.200
 17         -0.0114   J H20  H32    STAT=Y
 J57_58      8.1299   J H26  H27    STAT=Y  PRED= 8.260 RANGE= 0.500
 J57_59      2.0478   J H26  H30    STAT=Y  PRED= 2.030 RANGE= 0.800
 J58_59      0.0213   J H27  H30    STAT=Y  PRED= 0.430 RANGE= 0.320
 J60_61      8.1978   J H32  H33    STAT=Y  PRED= 8.260 RANGE= 0.500
 J60_62      2.0542   J H32  H36    STAT=Y  PRED= 2.030 RANGE= 0.800
 J61_62      0.0013   J H33  H36    STAT=Y  PRED= 0.430 RANGE= 0.320

CONTROL PARAMETERS:
       Solvent =  none (def. 99% enriched)
         1.000 =  Concentration (vol%, def=1.0%)
    0.00100000 =  Minimum line-intensity
    0.00100000 =  Diagonalization criterium (not in use)
  600.12700000 =  FIELD(1H,MHz), used to transform shifts to ppms
   13.14654155 =  Left frequency (ppm)
   -0.83260135 =  Right frequency (ppm)
        10.000 =  Acquisition time (s, for QMTLS)
         0.000 =  Line-width (for modes D, P & T, 0=use defaults)
   0.032002612 =  Data-point resolution (Hz)
         0.000 =  GAUSSIAN (%, 0=use default from INF)
         0.000 =  Dispersion contribution (%, 0=use default from INF)
    0.00000000 =  Decoupling frequency (for DORES)


END of FILE
```

*S11. Calculation of the Purity of **3** by qHNMR*

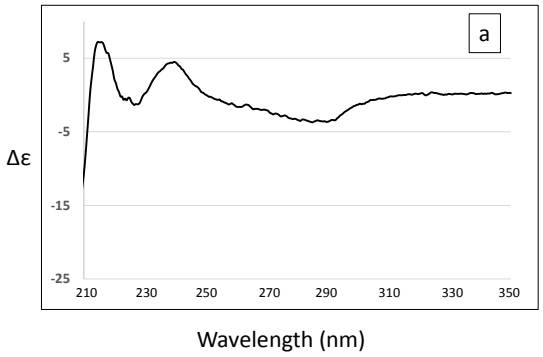
Calculation of the qHNMR purity (% w/w) of **3** using the 100 % method. The purity was determined to be 89 %, and the relative ratio of major and minor rotamers was found to be 63:37.

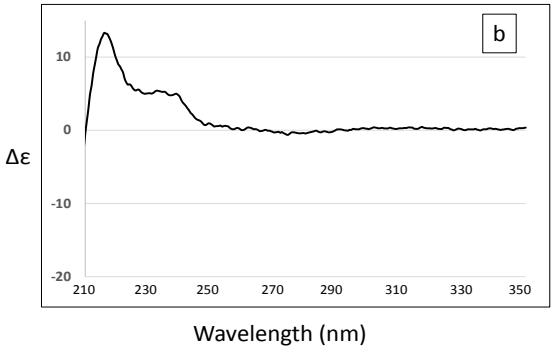
QUANTITATIVE ANALYSIS									
100% METHOD									
Isolated compound	Catechin-(4α→8)-catechin **(3)**	(integration in MestReNova)							
	Major rotamer	**int**	**# H**	**int 1H**	**average 1H**	**MW**	***F (Molar Mass ratio)***	**average mass**	**δH [ppm]**
Target	**signals**	128.11	1	128.11	101.10	578.52	1.00	101.10	2.49 (dd)
		119.71	1	119.71					2.76 (dd)
		122.15	1	122.15					3.79 (ddd)
		108.60	1	108.60					4.26 (d)
		103.02	1	103.02					4.36 (d)
		99.44	1	99.44					4.41 (d)
		100.34	1	100.34					4.54 (d)
		70.87	1	70.87					5.79 (d)
		77.77	1	77.77					5.89 (d)
		87.34	1	87.34					6.07 (s)
		102.58	1	102.58					6.25 (dd)
		103.09	1	103.09					6.47 (dd)
		100.00	1	100.00					6.59 (d)
		191.51	2	95.76					6.67, 6.68 (d)
		97.72	1	97.72					6.74 (d)
	Minor rotamer								
		81.80	1	81.80	58.57	578.52	1.00	58.57	2.58 (dd)
		72.49	1	72.49					2.82 (dd)
		69.37	1	69.37					4.07 (ddd)
		54.65	1	54.65					4.35 (d)
		51.86	1	51.86					4.51 (d)
		63.73	1	63.73					4.54 (d)
		54.02	1	54.02					4.74 (d)
		45.2	1	45.18					5.81 (d)
		38.94	1	38.94					5.84 (d)
		53.18	1	53.18					5.94 (s)
		99.88	2	49.94					6.77, 6.76 (d)
		104.57	2	52.29					6.83, 6.82 (dd)
		73.94	1	73.94					6.96 (d)
	Impurities								
	Structurally relate	18.49	1	18.49	16.73	578.52	1.00	16.73	2.64 (d)
		15.29	1	15.29					2.67 (d)
		15.20	1	15.20					2.71 (d)
		26.01	1	26.01					4.61 (d)
		19.07	1	19.07					4.77 (d)
		12.86	1	12.86					7.05 (d)
		10.18	1	10.18					7.12 (d)
Unrelated	Ethylene glycocl (grease)	55.47	2	27.74	27.74	62.07	0.11	2.98	1.29(brs)
								19.71	

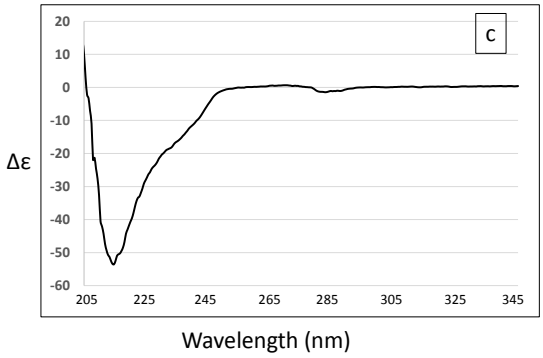
Mass (% w/w)						
1H pure cmpd	159.67					
1H Major rotamer	101.10					
1H Minor rotamer	58.57					
sum 1H all imp	19.71	Relative % of major and minor rotamers				
% Major rotamer	**56.36**	63.32				
% Minor rotamer	**32.65**	36.68				
% purity	**89.01**					
% impurity	**10.99**					

*S12. The CD Spectra of Compounds **1**, **2**, and **3***

The figures a-c show the CD spectra of compounds **1**, **2**, and **3**, respectively. The positive Cotton effect between 210 and 240 nm confirms the C-4β configuration at the inter-flavonoid bond in both compounds **1**and **2**. The negative Cotton effect between 210 and 240 nm confirms the C-4α configuration of compound **3**.







S13. References

1. Hemingway, R.W., L.Y. Foo, and L.J. Porter, *Linkage isomerism in trimeric and polymeric 2,3-cis-procyanidins.* Journal of the Chemical Society, Perkin Transactions, 1982**. 1**: p. 1209-1216.
2. Ferreira, D. and D. Slade, *Oligomeric proanthocyanidins: naturally occurring O-heterocycles.* Natural Product Reports, 2002**. 19**(5): p. 517-541.
3. Ferreira, D., D. Slade, and J.P.J. Marais. *Flavonoids: chemistry, biochemistry, and applications* in *Flavans and proanthocyanidins*. CRC, Taylor & Francis, Boca Raton (FL), 2006.
4. IUPAC, *Compendium of Chemical Terminology. Gold Book. Online Version 2.3.2.; http://goldbook.iupac.org/*. 2005-today.
5. Bedran-Russo, A. K., Pauli, G. F., Chen, S.-N., McAlpine, J. B., Castellan, C., Phansalkar, R., Aguilar, T., Vidal, C., Napolitano, J. G., Nam, J.-W., Lem, A. *Dentin biomodification: strategies, renewable resources and clinical applications.* Journal of Dentistry, 2014**. 30**(1): *p.* 62-76.