Taylor & Francis
Taylor & Francis Group

Natural Product Research
Formerly Natural Product Letters

ISSN: 1478-6419 (Print) 1478-6427 (Online) Journal homepage: http://www.tandfonline.com/loi/gnpl20

Linear regression analysis of silychristin A, silybin A and silybin B contents in *Silybum marianum*

Sameh F. AbouZid, Hayam S. Ahmed, Abd-El Mageed A. Abd El Mageed, Abeer S. Moawad, Asmaa I. Owis, Shao-Nong Chen, Amandine Nachtergael, James B. McAlpine, J. Brent Friesen & Guido F. Pauli

To cite this article: Sameh F. AbouZid, Hayam S. Ahmed, Abd-El Mageed A. Abd El Mageed, Abeer S. Moawad, Asmaa I. Owis, Shao-Nong Chen, Amandine Nachtergael, James B. McAlpine, J. Brent Friesen & Guido F. Pauli (2018): Linear regression analysis of silychristin A, silybin A and silybin B contents in *Silybum marianum*, Natural Product Research

To link to this article: https://doi.org/10.1080/14786419.2018.1527838

NATURAL PRODUCT RESEARCH
https://doi.org/10.1080/14786419.2018.1527838

Taylor & Francis
Taylor & Francis Group

SHORT COMMUNICATION

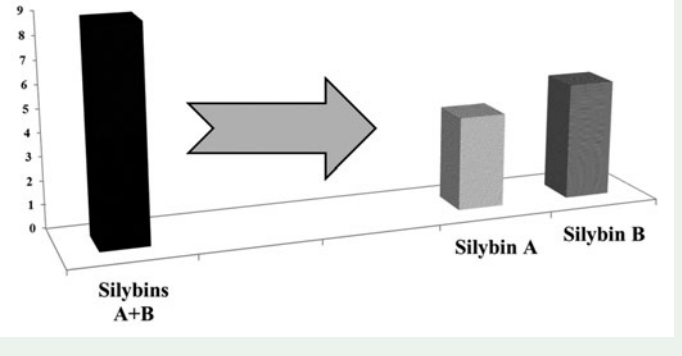
Linear regression analysis of silychristin A, silybin A and silybin B contents in *Silybum marianum*

Sameh F. AbouZid[a], Hayam S. Ahmed[a], Abd-El Mageed A. Abd El Mageed[a], Abeer S. Moawad[a], Asmaa I. Owis[a], Shao-Nong Chen[b,c], Amandine Nachtergael[c], James B. McAlpine[c], J. Brent Friesen[c,d] and Guido F. Pauli[b,c]

[a]Department of Pharmacognosy, Faculty of Pharmacy, Beni-Suef University, Beni-Suef, Egypt; [b]UIC/NIH Center for Botanical Dietary Supplements Research, Chicago, IL, USA; [c]Department of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, University of Illinois at Chicago, Chicago, IL, USA; [d]Physical Sciences Department, Rosary College of Arts and Sciences, Dominican University, River Forest, IL, USA

ABSTRACT

Quantitative correlations between the contents of the flavonolignans silychristin A and silybins A/B provide biosynthetic clues that support a pathway in which one mesomeric form of a taxifolin radical is undergoing an oxidative coupling with a coniferyl alcohol radical. The flavonolignan content and patterns reported in the literature for 53 samples, representing populations of the *Silybum marianum* plant growing in different parts of the world, were subject to a meta-analysis. Linear regression analyses were carried out on these data sets, and a mathematical model was derived that predicts the content of silychristin A relative to the metabolomic pattern of its congeners. The validity of the model was verified by applying it to test samples. This approach could potentially become a tool to enhance the understanding of both the relative composition of the silymarin complex and the biosynthetic pathways that underlie its formation.



ARTICLE HISTORY
Received 23 July 2018
Accepted 20 September 2018

KEYWORDS
Flavonolignan; linear regression; prediction; quantitative biosynthesis modelling

CONTACT Sameh F. AbouZid sameh.zaid@pharm.bsu.edu.eg

Supplemental data for this article can be accessed at https://doi.org/10.1080/14786419.2018.1527838.

1. Introduction

The standardized extract derived from the plant *Silybum marianum* (L.) Gaertn. (Asteraceae) is denoted as silymarin (Kroll et al. 2007). It consists of an isomeric mixture of the flavonolignans silybin A/B, isosilybin A/B, silychristin A, silydianin and isosilychristin. These flavonolignans are biosynthesized via oxidative coupling, between the flavonoid, taxifolin, and the phenylpropanoid, coniferyl alcohol. Based on chromatographic analysis of the silymarin content in *S. marianum* populations growing in Egypt, quantitative correlations between individual flavonolignans were found to exist (AbouZid et al. 2016). It was shown that a near unity positive correlation exists between silychristin A, silybin A, and silybin B. This can be explained using a biosynthetic rationale, conceptualizing that silychristin A is formed from a mesomeric form of the taxifolin radical that is the key intermediate for the formation of the silybins A & B (AbouZid et al. 2017). Based on *S. marianum* samples collected from different parts of Europe, similar observations were reported (Martinelli et al. 2016; Martinelli et al. 2017). In the referred studies, the correlations between flavonolignans were interpreted in terms of their biosynthesis. Such studies can potentially be useful in assembling models that can predict the abundance of biosynthetically related compounds.

The present study aimed at investigating the variation of silychristin A content in samples from various sources relative to the content of the silybins A and B by means of using multiple linear regression analysis. Moreover, the linear correlation between silychristin A and the silybins A & B was hypothesized to exist due to analogous biosynthetic patterns for these three flavonolignans.

2. Results and discussion

Table 1 shows the contents of three correlated flavonolignans silybin A, silybin B and silychristin A analyzed in forty three *S. marianum* samples collected from Europe, Egypt and Iran. Pairwise correlation of the contents of the three compounds (Table S1) revealed that these three flavonolignans are significantly correlated. This can be at least partially explained by the fact that they share close biosynthetic pathways. It was proposed that silybins A and B are formed from the same flavonoid radical precursor and silychristin A is formed from a mesomer of the same radical (AbouZid et al. 2017). The result of the analysis of variance using the Fisher's F test is shown in Table S2. The probability corresponding to the *F* value is lower than 0.0001. This means that the risk (lower than 0.01%) in assuming the null hypothesis is also negligible. The type I and type III sum of squares analysis (Table S2) show the significance of the explanatory variable, silybin A content, on the dependent variable, silychristin A content. Figure S1 shows the standardized regression coefficients and allows a direct comparison of the relative influence of the contents of silybins A & B on the content of silychristin A. The higher influence for silybin A, compared to silybin B, on the content of silychristin A may be interpreted in terms of the biosynthesis of the three compounds. Biosynthetic hypothesis for flavonolignans in *S. marianum* was previously reported (AbouZid et al. 2017). The site of nucleophile addition on the quinone ring involved in the biosynthesis of these compounds results in the *R* configuration of C-1″ in silychristin A and silybin A, and the *S* configuration of the same carbon in silybin B giving an

Table 1. The content of silybin A, silybin B, and silychristin (mg/g DW plant material) in 53 *Silybum marianum* samples.

Reference	Sample	Silychristin A	Silybin A	Silybin B	Silydianin	Isosilybin B
AbouZid et al. 2016	P263	0.56	0.20	0.53	4.40	0.53
	P26394	4.66	4.08	6.36	0.44	0.29
	1P + 293W	4.55	4.02	6.24	0.76	0.46
	1P + 293Br	3.17	2.79	4.41	0.75	0.39
	1P + 293B	2.92	2.43	3.97	0.60	0.34
	2P293	0.33	0.14	0.38	2.87	0.47
	3P293	2.77	2.47	3.91	1.35	0.38
	1W+	0.46	0.17	0.52	4.14	0.67
	1P64	0.55	0.27	0.70	4.36	0.77
	2P64	5.40	4.62	7.08	0.52	0.34
	3P64W	*5.11*	*4.46*	*6.81*	*0.58*	*0.32*
	3P64Br	*5.19*	*4.41*	*6.85*	*0.54*	*0.32*
	3P64B	*5.08*	*4.41*	*6.75*	*0.54*	*0.33*
	3W	*3.86*	*3.42*	*5.35*	*0.39*	*0.24*
	P154	*4.13*	*3.59*	*5.69*	*1.29*	*0.48*
Cheilari et al. 2016	*PL01*	*3.82*	*3.26*	*5.37*	*0.31*	*0.70*
	PL02	1.14	0.82	1.35	2.28	0.78
	PL03	1.19	0.98	1.39	1.62	0.84
	PL04	1.58	1.08	1.91	2.53	0.89
	PL05	1.46	1.03	1.91	2.44	0.90
	PL06	1.21	0.73	1.29	2.90	0.96
	PL07	1.58	1.07	1.85	3.01	0.88
	PL08	1.02	1.16	1.37	2.11	0.77
	PL09	1.41	1.06	1.70	2.64	0.86
	PL10	1.43	1.03	1.70	3.16	0.84
Martinelli et al. 2016	G1	1.94	1.20	1.75	27.55	2.65
	G2	2.07	1.19	1.76	28.11	2.78
	G3	13.89	12.94	13.74	0.23	1.23
	G4	2.49	1.51	2.11	34.55	3.12
	G5	1.96	1.19	1.66	27.20	2.46
	G6	2.38	1.37	2.09	36.98	3.04
	G7	2.08	1.11	1.64	28.68	2.78
	G8	2.07	1.12	1.65	29.31	2.81
	G9	1.60	0.84	1.28	22.87	2.05
	G10	5.63	4.84	5.42	17.85	2.05
	G11	10.70	9.69	10.62	12.15	2.12
	G12	5.07	4.31	4.96	20.61	2.09
	G13	2.20	1.23	1.83	30.67	2.87
	G14	8.57	7.83	8.40	7.82	1.44
	G15	10.95	10.54	11.32	0.23	1.01
	G16	1.99	1.33	1.86	27.28	2.53
	G17	1.84	1.05	1.50	24.74	2.28
	G18	1.88	1.03	1.59	25.19	2.35
	G19	10.41	10.60	11.37	0.21	1.04
	G20	1.94	0.99	1.54	25.49	2.40
	G21	7.03	5.95	6.72	23.42	2.71
	G22	*11.88*	*10.81*	*11.79*	*9.68*	*1.86*
	G23	*14.08*	*12.63*	*14.03*	*2.34*	*1.39*
	G24	*2.57*	*1.39*	*2.16*	*35.73*	*3.35*
	G25	*2.41*	*1.33*	*2.13*	*35.00*	*3.08*
	G26	5.82	5.00	5.83	21.58	2.40
Shokrpour et al. 2008	North	2.75	1.35	2.97	2.81	1.59
	West	3.39	2.00	3.20	2.30	1.00

Sample names represent those reported in the literature. Training samples are shown in normal font and test samples in italics.

evidence to this hypothesis. This may emphasize the role of enzymes, such as peroxidases, that might be involved in the formation of the flavonolignans. Participating enzyme may differ for each isomer, or single enzyme with loose binding pockets catalyzes the formation of congeneric molecules at different rates. Ascorbate peroxidase

was reported to catalyze this reaction (Lv et al. 2017). An alternative pathway for silymarin biosynthesis based on flavonolignans correlation patterns was reported (Martinelli et al. 2017).

An equation describing *silychristin A* content starting from silybin A and silybin B contents can be formulated as: *Silychristin A content* = *0.56169* + *0.909663* * *Silybin A content* + *0.092599* * *Silybin B content*

In order to validate the correctness of this model, a test data set of 10 samples reported in literature was used (AbouZid et al. 2016; Cheilari et al. 2016; Martinelli et al., 2016). The values of silybins A and B were used, and the predicted values of silychristin A content were derived from the fitted model. Table S3 compares the predicted values of silychristin A content with that of reported experimental data. Figure S2 shows the validation of the prediction. The fitted model was able to predict with good confidence ($R^2 = 0.9925$) the reported content of silychristin A in *S. marianum* plant material. The multiple linear regression model developed here for the flavanolignans shows multicolinearity. This reflects the fact that the explanatory variables, silybin A and silybin B contents, are correlated not just to the content of silychristin A, but also to each other. This unwanted effect of multicolinearity can be solved by removing one variable from the model without drastically reducing the value of R^2. This is equivalent to constructing a simple linear regression model between silychristin A and either silybin A or silybin B.

The simple linear regression model between silychristin A and either silybin A or silybin B can be useful in predicting the individual content of these two constitutional isomers. Notably, the commonly observed lack of resolution in chromatographic analysis leads to reporting the sum of the contents of these two diastereomers (Andrzejewska et al. 2011; Elwekeel et al. 2012; Rahal et al. 2015). A similar challenge exists when integral-based quantitative nuclear magnetic resonance (qNMR) methods are used for the analysis of these diastereomers (Cheilari et al., 2016). Figure S3 shows the ^{1}H NMR spectrum of silymarin extracted from the fruits of *S. marianum*. The proton signal of silychristin A at 5.55 ppm is located in an uncrowded region of the spectrum. On the other hand, proton signals of the diastereomers silybin A and silybin B overlapped at 7.10 ppm and 7.08 ppm, respectively. Quantum mechanical driven ^{1}H iterative Full Spin Analysis (HiFSA) approach was reported to quantitatively distinguish silybin A from silybin B as well as isosilybin A from isosilybin B (Napolitano et al. 2013). These regio- and diastereo-isomers exhibit extremely similar coupling patterns and chemical shift differences (Kim et al. 2003). To demonstrate the usefulness of the simple linear regression models shown in Figure S4, these models were used to calculate the contents of the diastereomers silybin A and silybin B. The sum of the content of these two compounds was reported from qNMR data (Cheilari et al. 2016). The correctness of this method was then validated using chromatographic data in the referred study. Data is shown in Table S4. It is interesting to find the error (predicted – experimental) is of negative value for the all of the samples in the data set.

3. Experimental

Experimental details are available online as a supplementary material.

4. Conclusion

An equation for the calculation of silychristin A content starting from silybin A and silybin B contents was provided and its accuracy was tested on 10 samples. The demonstrated relatively strict correlation of certain congeners contained in the otherwise complex silymarin mixture indicate that the variability of a true natural product can be limited and, ultimately, encoded by the biosynthetic process that occurs in the plant potentially regardless of the degree to which it is controlled enzymatically.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Sameh F. AbouZid http://orcid.org/0000-0002-0789-767X
Hayam S. Ahmed http://orcid.org/0000-0002-7379-3078
Abd-El Mageed A. Abd El Mageed http://orcid.org/0000-0002-4582-0606
Abeer S. Moawad http://orcid.org/0000-0002-0228-6757
Asmaa I. Owis http://orcid.org/0000-0003-1860-9124
Shao-Nong Chen http://orcid.org/0000-0003-0748-0863
Amandine Nachtergael http://orcid.org/0000-0002-8697-2809
James B. McAlpine http://orcid.org/0000-0003-1696-8828
J. Brent Friesen http://orcid.org/0000-0002-0739-9224
Guido F. Pauli http://orcid.org/0000-0003-1022-4326

References

AbouZid S, Chen SN, Pauli GF. 2016. Silymarin content in *Silybum marianum* populations growing in Egypt. Ind Crops Prod. 83:729–737.

AbouZid S, Ahmed HS, Moawad AS, Owis AI, Chen CN, Nachtergael A, McAlpine JB, Friesen JB, Pauli GF. 2017. Chemotaxonomic and biosynthetic relationships between flavonolignans produced by *Silybum marianum* populations. Fitoterapia. 119:175–184.

Andrzejewska J, Sadowska K, Mielcarek S. 2011. Effect of sowing date and rate on the yield and flavonolignan content of the fruits of milk thistle (*Silybum marianum* L. Gaertn.) grown on light soil in a moderate climate. Ind Crop Prod. 33(2):462–468.

Cheilari A, Sturm S, Intelmann D, Seger C, Stuppner H. 2016. Head-to-head comparison of ultra-high-performance liquid chromatography with diode array detection versus quantitative nuclear magnetic resonance for the quantitative analysis of the silymarin complex in *Silybum marianum* fruit extracts. J Agric Food Chem. 64(7):1618–1626.

Elwekeel A, AbouZid S, Sokkar N, Elfishway A. 2012. Studies on flavanolignans from cultured cells of *Silybum marianum*. Acta Physiol Plant. 34(4):1445–1449.

Kim NC, Graf TN, Sparacino CM, Wani MC, Wall ME. 2003. Complete isolation and characterization of silybins and isosilybins from milk thistle (*Silybum marianum*). Org Biomol Chem. 1(10):1684–1689.

Kroll DJ, Shaw HS, Oberlies NH. 2007. Milk thistle nomenclature: why it matters in cancer research and pharmacokinetic studies. Integr Cancer Ther. 6(2):110–119.

Lv Y, Gao S, Xu S, Du G, Zhou J, Chen J. 2017. Spatial organization of silybin biosynthesis in milk thistle [*Silybum marianum* (L.) Gaertn]. Plant J. 92(6):995–1004.

Martinelli T, Potenza E, Moschella A, Zaccheria F, Benedettelli S, Andrzejewska J. 2016. Phenotypic evaluation of a milk thistle germplasm collection: fruit morphology and chemical composition. Crop Sci. 56(6):3160–3172.

Martinelli T, Whittaker A, Benedettelli S, Carboni A, Andrzejewska J. 2017. The study of flavonolignan association patterns in fruits of diverging *Silybum marianum* (L.) Gaertn. chemotypes provides new insights into the silymarin biosynthetic pathway. Phytochemistry. 144:9–18.

Napolitano JG, Lankin DC, Graf TN, Friesen JB, Chen JB, McAlpine JB, Oberlies NH, Pauli GF. 2013. HiFSA fingerprinting applied to isomers with near-identical NMR spectra: the silybin/isosilybin case. J Org Chem. 78(7):2827–2839.

Rahal NB, Barba FJ, Barth D, Chevalot I. 2015. Supercritical CO_2 extraction of oil, fatty acids and flavonolignans from milk thistle seeds: Evaluation of their antioxidant and cytotoxic activities in Caco-2 cells. Food Chem Toxicol. 83:275–282.

Shokrpour M, Mohammadi SA, Moghaddam M, Ziai SA, Javanshir A. 2008. Variation in flavonolignan concentration of milk thistle (*Silybum marianum*) fruits grown in Iran. J Herbs Spices Med Plant. 13(4):55–69.