Vol. 13(1), pp. 9-18, January-June 2021
DOI: 10.5897/IJPPB2020.0304
Article Number: 7BBA7C866821
ISSN 2006-9758

http://www.academicjournals.org/IJPPB

ACADEMIC JOURNALS
expand your knowledge

International Journal of Plant Physiology and Biochemistry

Full Length Research Paper

The costunolide biosynthesis enzymes of *Artemisia glabella* Kar. et Kir.: Determination of the nucleotide sequences of the mRNA

Adekenova Gulimzhan S.[1], Shaikenova Zh. S.[1], Chervyakova O. V.[2], Zakarya K. D.[2] and Adekenov S. M.[1*]

[1]JSC "International Research and Production Holding Phytochemistry, 100009, Republic of Kazakhstan.
[2]Research Institute for Biological Safety Problems, Ministry of Education and Science, Republic of Kazakhstan.

Received 8 December, 2020; Accepted 28 April, 2021

Artemisia glabella **Kar. et Kir. is a source of sesquiterpene lactone arglabin, which has antitumor, radiosensitizing and immunomodulatory activity. Studying the biosynthesis of arglabin and its derivatives will allow us to develop the biotechnological basis for its production, thereby increasing its availability. The precursor to most sesquiterpene lactones is costunolide. The purpose of these studies was to detect and determine the nucleotide sequences of mRNA enzymes germacrene A synthase (GAS, EC 4.2.3.23), germacrene A oxidase (GAO, EC 1.14.14.95) and costunolide synthase (COS, EC 1.14.13.120), involved in the biosynthesis of costunolide in *A. glabella*. As a result of studies, mRNA was isolated from various forms (intact plant, regenerant plant, callus) of *A. glabella*. Using specific primers, mRNA fragments of genes encoding sesquiterpene lactones biosynthesis enzymes were amplified and their nucleotide sequences were determined. A comparative analysis of the obtained sequences showed their high (>90%) identity with the genes of GAS, GAO and COS of other representatives of the family Asteraceae. It was revealed that enzymes are expressed both in an intact plant and in calluses and regenerant plants obtained *in vitro*.**

Key words: *Artemisia glabella* Kar. et Kir., mRNA, cDNA, sequencing, germacrene A synthase, germacrene A oxidase, costunolide synthase.

INTRODUCTION

Sesquiterpenoids of the guaiane series belong to fairly widespread substances of the plant kingdom. A particularly large number of guaiane compounds are included in the group of sesquiterpene lactones called guaianolides. α-Methylene guaianolides have a bitter taste, are toxic, and exhibit cytotoxic and antitumor properties. Their molecules, as a rule, are polyfunctional and many of them are characterized by the presence of exocyclic unsaturation (Fischer et al., 1979; Frederick, 1982).

Sesquiterpene lactones are derivatives of isopentenyl diphosphate (IPP), which can be synthesized in two

*Corresponding author. E-mail: info@phyto.kz.

different ways: the mevalonate (MVA) pathway and the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway. In the MEP pathway, glyceraldehyde-3-phosphate and pyruvic acid are converted to IPP and DMAPP by seven enzymes. In the MVA pathway, acetyl-CoA is converted to IPP through six steps followed by IPP isomerase (IPPI). Two IPPs and one DMAPP are then converted to farnesyl diphosphate (1) (FPP), the precursor of all sesquiterpenes, by farnesyl diphosphate synthase (FPS). FPP (1) can be converted to sesquiterpenes using sesquiterpene synthases FPP (1).

Sesquiterpene lactones are formed in the plant organism by oxidation of the methyl part of the isopropyl group attached to the main carbon skeleton. Cyclization of farnesyldiphosphate (1) gives (+)- germacrene A (2). Oxidation of the isopropenyl side chain by (+)-germacrene A-hydroxylase to primary alcohol and further oxidation by NAD (P)+- dependent dehydrogenases gives germacrene acid (3). This is followed by hydroxylation at the C-6 position and subsequent lactonization leads to (+)- costunolide (4). It is assumed that the second stage of the cyclization of germacranolides to the guaianolide skeleton proceeds through epoxidation or hydroxylation of the costunolide skeleton (Ma et al., 2019).

As a result of research on the homology of nucleotide sequences with known sesquiterpene monooxygenases, six promising cytochrome P450 contigs (actin, GAS, GAO, COS, parthenolide synthase and costunolide 3β-hydroxylase) were identified and selected for functional characterization. A new cytochrome P450, cauniolide synthase, which catalyzes the formation of guaianolide cauniolide (5) from the germacranolide substrate costunolide (4), has been characterized. Unlike most cytochromes P450s, cauniolide synthase has a unique mechanism of action, combining stereoselective hydroxylation of costunolide (4) at the C-3 position with elimination, cyclization, and regioselective deprotonation (Liu et al., 2018).

One of the promising guaianolides is arglabin (6), isolated for the first time from *Artemisia glabella* Kar. et Kir., an endemic species of wormwood (*Artemisia* L.), which grows exclusively in Central Kazakhstan. *A. glabella* is a source of a number of biologically active compounds such as sesquiterpene lactones and essential oils (Adekenov et al., 1982; Adekenov et al., 1995). Arglabin (6) is the active substance of the anticancer drug “Arglabin”, which was developed at the International Research and Production Holding “Phytochemistry” and is produced on an industrial scale by the Karaganda Pharmaceutical Plant (Karaganda, Republic of Kazakhstan) (Adekenov, 2001, 2015).

Arglabin (6) is found in all organs of *A. glabella* and throughout the growing season (Adekenov et al., 1995). Its quantitative accumulation is observed during the budding period in leaves (1.90%) and buds (1.56%) (Mantler et al., 2020). However, the biosynthesis stages of the arglabin molecule in a plant organism at the genetic level are studied insufficiently.

Promising approach to obtain arglabin (6) can be the reconstruction of the biosynthetic pathway in a heterologous system. However, despite the widespread medical use of arglabin (6), the biosynthetic pathway in plants remains unexplored. Adekenov and Bouwmeester (2015) described a possible pathway for arglabin biosynthesis in *A. glabella*: farnesyldiphosphate (1), germacrene A (2), germacrene A-ol (7), germacrene A-on (8), germacrene acid A (3), costunolide (4), cauniolide (5), and arglabin (6) (Figure 1).

The formation of germacrene A (2) as a result of the cyclization of farnesyl diphosphate (1) catalyzed by germacrene A synthase (GAS) is a generally recognized step in the biosynthesis of sesquiterpenes (Xu and Dickschat, 2020). GAS was isolated and characterized from a number of plants of the family Asteraceae (Bouwmeester et al., 2002; Majdi et al., 2011; Menin et al., 2012; Pazouki et al., 2015; Nguyen et al., 2016). The next step in biosynthesis is the sequential oxidation of germacrene A (2) by germacrene A oxidase (GAO) and costunolide synthase (COS). As a result, germacrene acid (3) and costunolide (4) are formed, respectively (Nguyen et al., 2019; de Kraker et al., 2002; Ikezawa et al., 2011).

In previous studies, a number of genes encoding sesquiterpene enzymes that control key steps in secondary metabolic pathways have been characterized from a number of herbal plant species. However, the genes responsible for enzymatic processes in guaianolide biosynthesis have not been identified. Therefore, the isolation and characterization of genes involved in the biosynthesis of arglabin (6) in *A. glabella* is a promising direction of research for the reconstruction of the process in a heterologous system. The main objectives of this study were the isolation of genes encoding the GAS, GAO and COS in *A.glabella* and the determination of the nucleotide sequences of mRNA, as well as the assessment of the expression of these genes at the level of transcription in regenerated plants and callus tissues.

MATERIALS AND METHODS

Plant

The intact plant of *A. glabella* Kar. et Kir. in the budding phase under *ex situ* conditions, the regenerant plant and callus tissues obtained *in vitro* culture in the biotechnology laboratory of JSC “International Research and Production Holding “Phytochemistry” were used. Leaves and flower buds were collected and immediately frozen in liquid nitrogen and stored at -80°C (three biological replicates for each tissue were used).

RNA isolation: cDNA synthesis

For total RNA extraction, fresh or freezed plant material (about 100 mg) immediately transferred to 200 µl Tryzol reagent in 1.5 ml

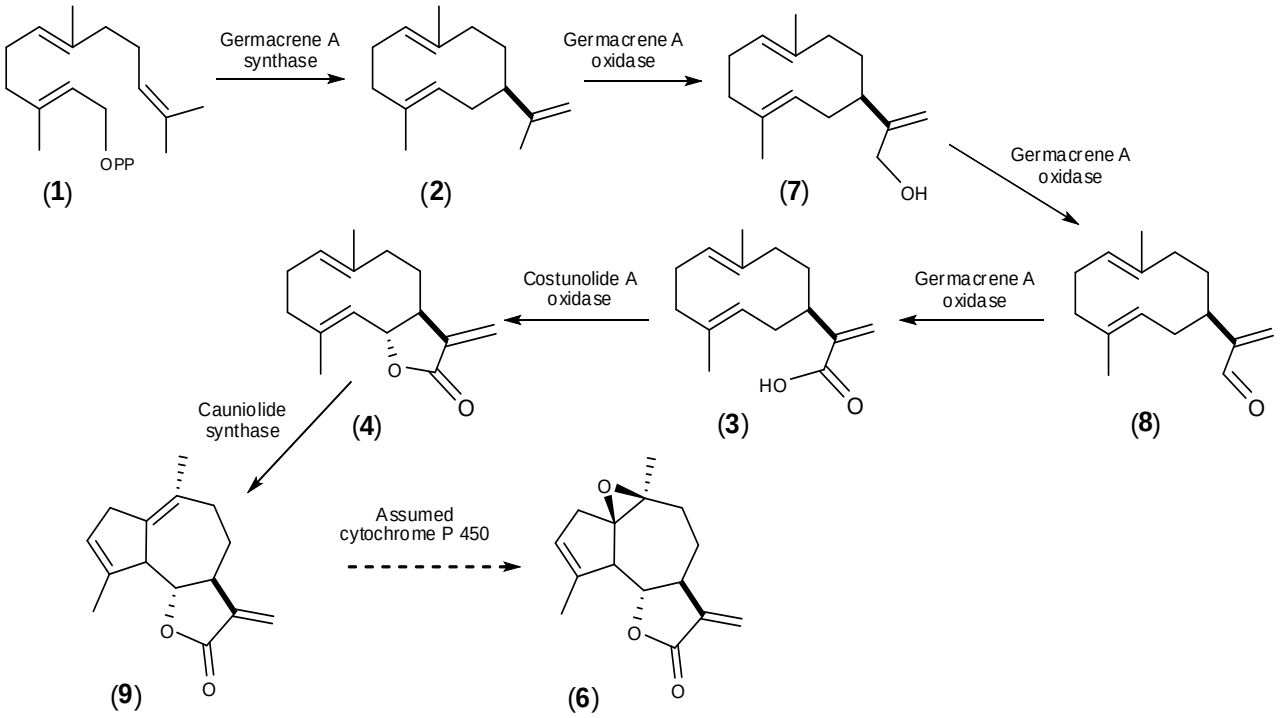


Figure 1. Estimated pathway of arglabin biosynthesis.

microcentrifuge tube. Cells were disrupted mechanically with a pestle in the presence of carborundum. After cell disruption, an additional 800 µl Tryzol reagent was added. All subsequent steps were carried out as described in the manual.

mRNA was isolated from total RNA using the Dynabeads® mRNA DIRECT™ Purification Kit (Invitrogen, USA) according to the manufacturer's protocol.

cDNA synthesis was performed using the Maxima H Minus Double-Stranded cDNA Synthesis Kit (Thermo Scientific, USA), according to the manufacturer's protocol, with oligo (dT) primer.

Amplification and sequencing of fragments of targeted genes

Amplification of cDNA fragments of targeted genes was carried out using specific primers designed as part of these studies. A 50 µl reaction mixture contained: 5 µl 10×PCR buffer (Silex, Russia), 1 µl 10 mM dNTPs (NEB), 0.1 µl cDNA (100 ng/µl), 1 µl of each primer (20 pM/µl), 0.25 µl Taq DNA polymerase (1.25 units, Sileks). Amplification was carried out under the following conditions: at 94°C 5 min; then 30 cycles at 94°C 1 min; at 50°C 1 min; at 68°C 2 min; final elongation at 68°C 7 min.

PCR products were cloned in pJet1.2 vector using the Clone JET PCR Cloning Kit (Thermo Scientific, USA) according to the manufacturer's instructions. Competent cells of *Escherichia coli* TOP10 were transformed with ligation mix. Presence and size of insertions in plasmid vector was determined by PCR analysis of *E. coli* TOP10 colonies according to the manufacturer's instructions for CloneJET PCR Cloning Kit (Thermo Scientific). Plasmid DNA was isolated from *E. coli* TOP10 using the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich, USA) according to the manufacturer's protocol.

Nucleotide sequences of targeted genes were determined using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Austin, TX), on the DNA sequencer Genetic Analyser 3130xl, Applied Biosystems, Hitachi, Japan. Nucleotide and their relevant amino acid sequences were analyzed with the help of the software Vector NTI 10.0.1, Invitrogen. The GAS/GAO/COS nucleotides from the GenBank databases were used for comparison.

Phylogenetic analyses

The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016).

RESULTS AND DISCUSSION

Design of primers

The design of primers for amplification of target genes was carried out based on the analysis of the nucleotide and amino acid sequences of GAS, GAO and COS in members of the Asteraceae family presented in the GenBank of The National Center for Biotechnology Information, U.S. National Library of Medicine.

Amino acid sequences of GAS, GAO, COS are presented in GenBank for 17, 11 and 8 species of the Asteraceae family, respectively. Two forms of GAS were found: short ~ 560 amino acids (a.a.) and long ~ 580 a.a. The average the length of the amino acid sequences of

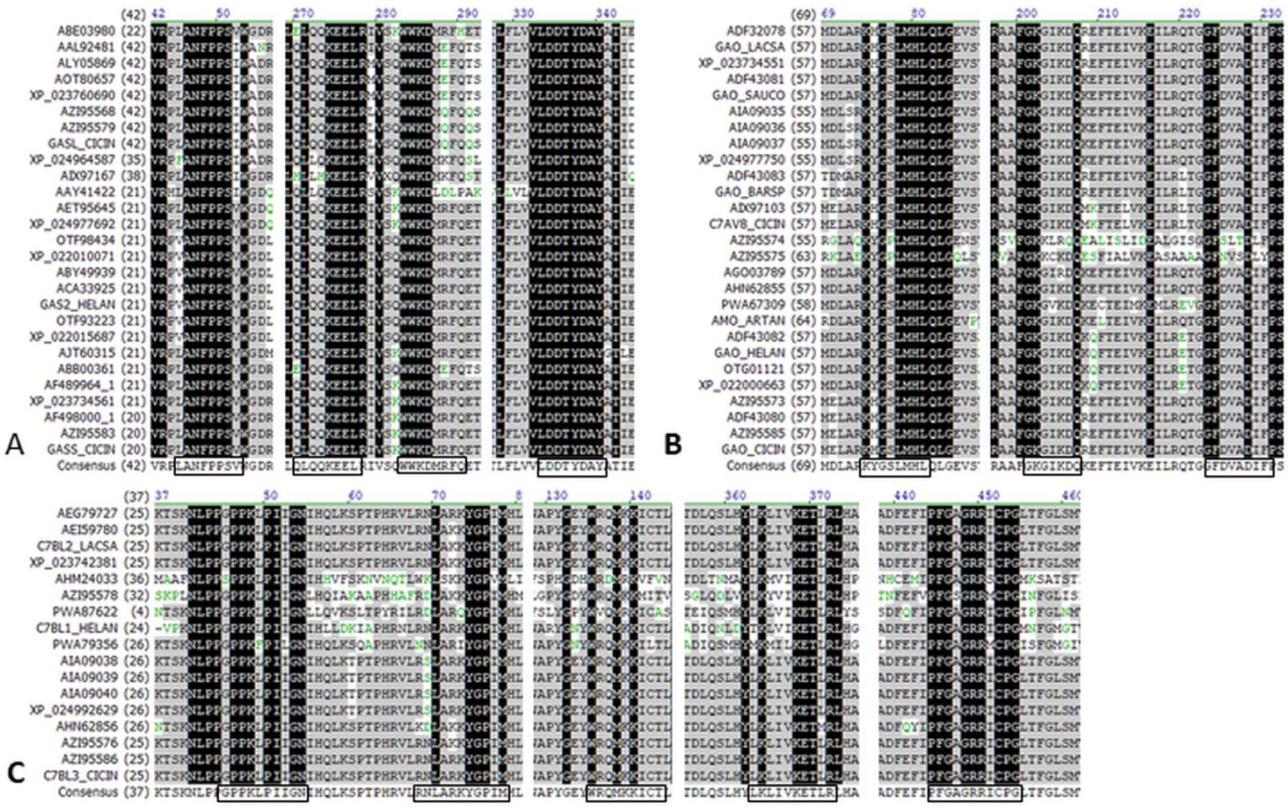


Figure 2. Alignment of amino acid sequences of GAS (A), GAO (B), and COS (C) (conservative sites highlighted).

Table 1. Primers used in this study.

Name	Target peptide	Primer sequence 5'-3'	PCR-product, bp
GAS-F1	LQLQQKEEL	tggagctacgacaaaaagaagaactg	216-885
GAS-F2	LANFPPSVW	ctggccaactttcctccttcagtatgg	
GAS-R1	LDDTYDAY	ataagcatcatatgtgtcatcta	
GAS-R2	WWKDMRF	aaacctcatgtccttccacc	
GAO-F1	KYGSLM	gccagaaagtatggatctttaatgca	405-469
GAO-R1	GKGVKDQ	ctggtccttgactccttttcc	
GAO-R2	GFDVADIF	gaaagatatctgccacatcaaagcc	
COS-F1	GPPKLPIIGN	gggccgccaaaactacccataatcggaaac	300-1200
COS-F2	LKLIVKETLR	ctgaaattaatagtaaaagaaactctgagg	
COS-F3	RNLAKKYGPIM	gaaacttagccaagaaatatggccccatcatg	
COS-R1	PFGAGRRICPG	acaaattctccggccggcaccgaacgg	
COS-R2	WRQMKKICTLE	ctccaaggtgcaaatcttcttcatctgcctcca	

GAO is 488 a.a. and COS is 494 a.a.

Conserved short peptides were identified based on the multiple alignments of the amino acid sequences of GAS, GAO, COS (Figure 2). These peptides became targets for the design of specific primers (Table 1).

Amplification and sequencing of target genes

Fragments of target genes of *A. glabella* (Ag) were amplified using specific primers (Figure 3A) and sequenced. The obtained sequences of mRNA fragments

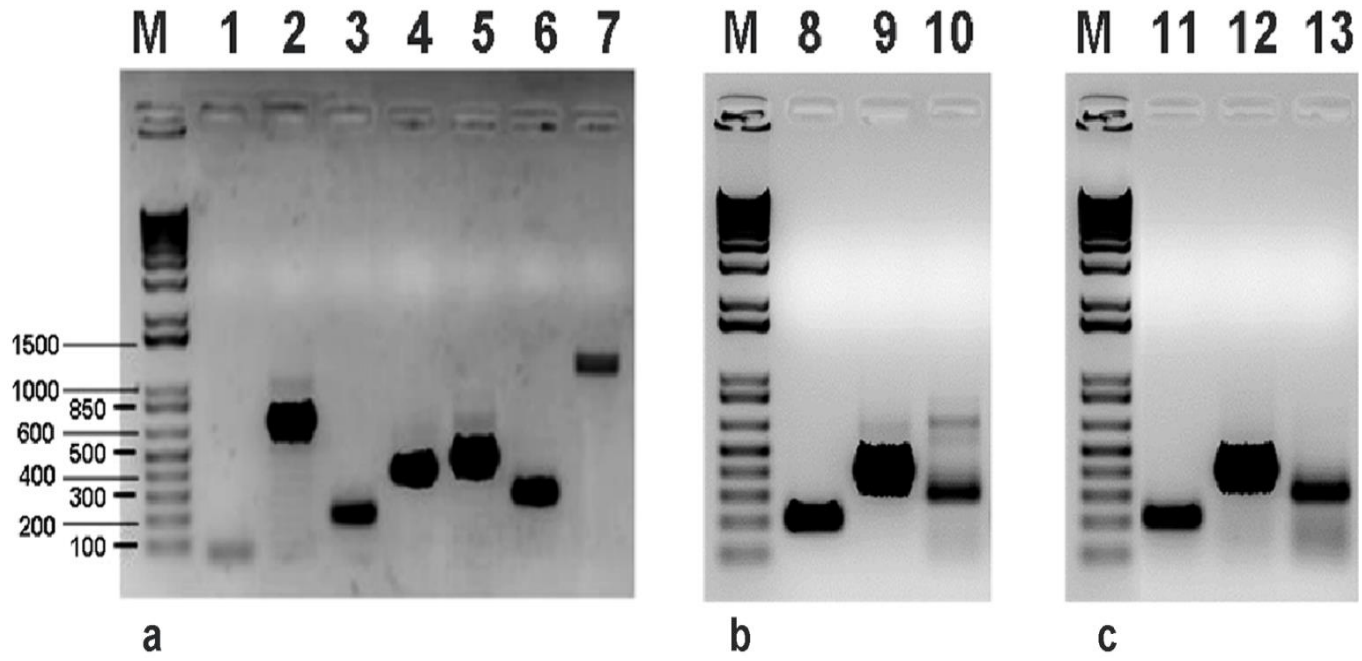


Figure 3. Results of target genes fragments amplification using cDNA obtained from various tissues of *A. glabella* Kar. et Kir. as a template. a - cDNA of an intact plant, b - cDNA of a regeneranted plant, c - cDNA of callus tissues; M - marker of the length of DNA fragments (1 Kb Plus DNA Ladder, Invitrogen); 1 - primers GAS-F2/GAS-R1 (885 bp); 2 - primers GAS-F1/GAS-R2 (216 bp); 3,8,11 - primers GAS-F2/GAS-R2 (737 bp), 4,9,12 - primers GAO-F1/GAO-R1 (405 bp); 5 - primers GAO-F1/GAO-R2 (469 bp); 6 - primers COS-F1/COS-R1 (1200 bp); 7,13 - primers COS-F3/COS-R2 (300 bp).

of AgGAS, AgGAO, and AgCOS were deposited in GenBank with IDs MT276314, MT276315, and MT276313, respectively.

Phylogenetic analysis

Pairwise protein sequences comparison of putative AgGAS, AgGAO and AgCOS with all known enzymes shows wide range of homology.

In silico analysis revealed that putative AgGAS has about 40 to 90% homology at protein level with annotated germacrene A synthases. It has highest identity (90%) with *Tanacetum parthenium* germacrene synthase (TpGAS) and lowest identity (40%) with *Solidago canadensis* germacrene synthases (ScGAS) (Figure 4A). Neighbour joining algorithm generated phylogenetic tree has combined AgGAS, AaGAS and TpGAS into one evolutionary lineage (Figure 4B). *Tanacetum* and *Artemisia* are members of the taxonomic tribe Anthemideae.

Phylogenetic analysis of AgGAO and AgCOS showed high homology with the corresponding enzymes of *T. parthenium*, 97 and 95%, respectively (Figures 5A and 6A). Neighbour joining algorithm generated phylogenetic tree showed separate evolutionary lineage of AgGAO and AaGAO (Figure 5B). However, AgCOS was combined in one lineage with TpCOS (Figure 6B).

Expression of target genes in plant tissues

Expression of AgGAS, AgGAO and AgCOS was evaluated at the transcription stage in regenerant plant and callus tissues obtained *in vitro.* As a result, we found that mRNA of costunolide (4) biosynthesis enzymes is present both in regenerant plants and in callus tissues (Figure 3B and C). Regenerated plants and callus tissue can also provide an alternative source for arglabin (6) production.

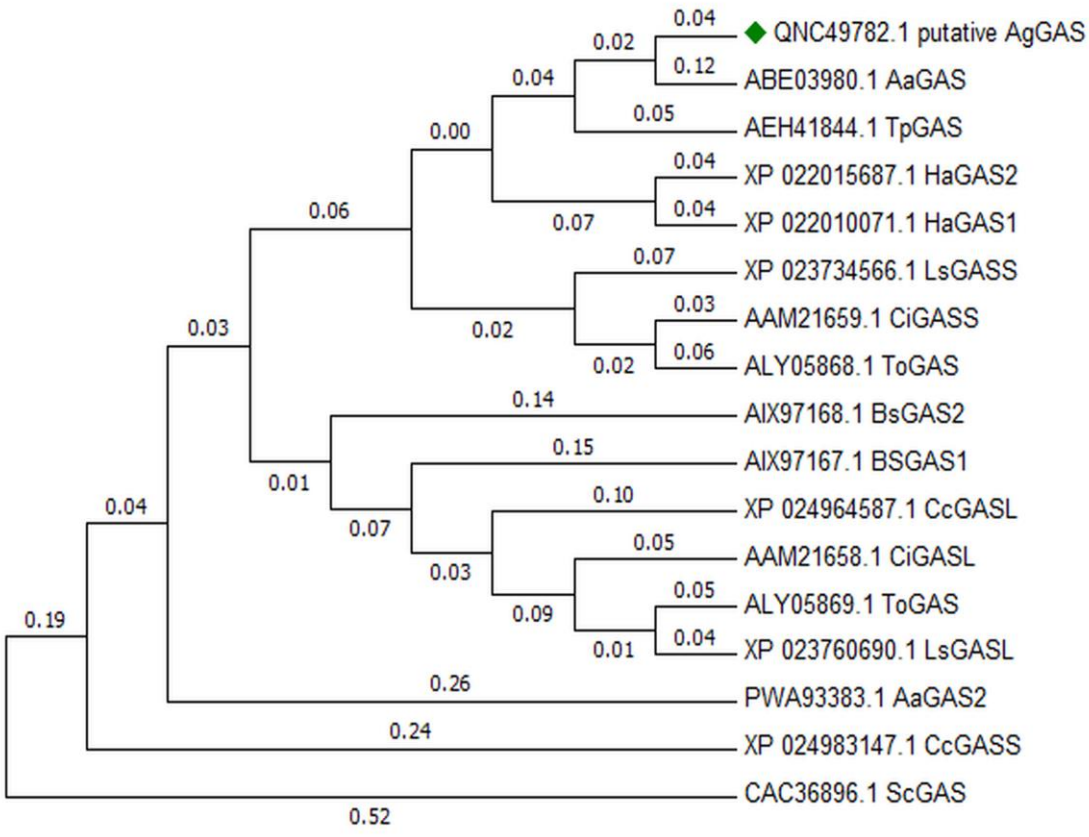
Conclusion

Thus, for the first time the nucleotide sequences of mRNA fragments encoding AgGAS, AgGAO and AgCOS were determined. A comparative analysis of the obtained sequences showed their high (> 90%) identity with the annotated GAS, GAO and COS of other members of the family Asteraceae. It was found that AgGAS, AgGAO and AgCOS are expressed both in the intact plant, as well as in calluses and regenerated plants obtained *in vitro*.

The results obtained provide a basis for further study of the biosynthetic pathway of arglabin (6) in plants. Isolation and characterization of enzymes will make it possible to reconstruct the biosynthesis of arglabin (6) in a heterologous system and to simplify the methods of its

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
QNC49782.1_putative_AgGAS	1		87	70	68	86	84	67	84	69	62	68	84	83	90	40	77	57
ABE03980.1_AaGAS	2	0.15		65	64	79	77	63	79	64	60	61	76	76	83	35	71	51
AIX97167.1_BSGAS1	3	0.40	0.49		80	73	72	79	73	80	63	78	73	72	71	39	76	51
AAM21658.1_CiGASL	4	0.42	0.48	0.34		72	72	93	71	94	62	79	71	71	70	38	74	47
AAM21659.1_CiGASS	5	0.18	0.25	0.35	0.37		94	71	90	72	67	69	88	89	87	38	73	56
ALY05868.1_ToGAS	6	0.20	0.28	0.39	0.38	0.08		71	88	72	66	69	86	87	85	37	80	54
ALY05869.1_ToGAS	7	0.44	0.50	0.35	0.11	0.39	0.40		70	92	61	79	70	70	68	38	73	48
XP_023734566.1_LsGASS	8	0.19	0.26	0.40	0.42	0.12	0.15	0.42		71	66	69	85	86	86	38	79	55
XP_023760690.1_LsGASL	9	0.42	0.48	0.32	0.10	0.38	0.39	0.10	0.41		62	79	71	71	70	38	74	48
XP_024983147.1_CcGASS	10	0.49	0.58	0.50	0.55	0.43	0.45	0.55	0.43	0.54		60	66	66	66	37	68	50
XP_024964587.1_CcGASL	11	0.38	0.47	0.26	0.24	0.32	0.32	0.24	0.37	0.25	0.53		69	68	67	38	72	50
XP_022015687.1_HaGAS2	12	0.20	0.30	0.38	0.40	0.18	0.21	0.42	0.22	0.40	0.46	0.35		95	86	38	78	53
XP_022010071.1_HaGAS1	13	0.22	0.30	0.40	0.41	0.17	0.19	0.43	0.20	0.42	0.48	0.37	0.08		86	38	78	53
AEH41844.1_TpGAS	14	0.11	0.18	0.42	0.41	0.16	0.19	0.44	0.18	0.42	0.46	0.38	0.19	0.19		37	78	55
CAC36896.1_ScGAS	15	0.94	1.02	0.93	1.02	0.88	0.93	0.98	0.93	0.98	0.94	0.98	0.91	0.94	0.98		38	32
AIX97168.1_BsGAS2	16	0.30	0.40	0.35	0.40	0.28	0.30	0.41	0.29	0.39	0.46	0.32	0.33	0.34	0.31	0.93		56
PWA93383.1_AaGAS2	17	0.42	0.53	0.54	0.58	0.42	0.46	0.58	0.45	0.58	0.53	0.53	0.48	0.46	0.44	1.01	0.42	

A

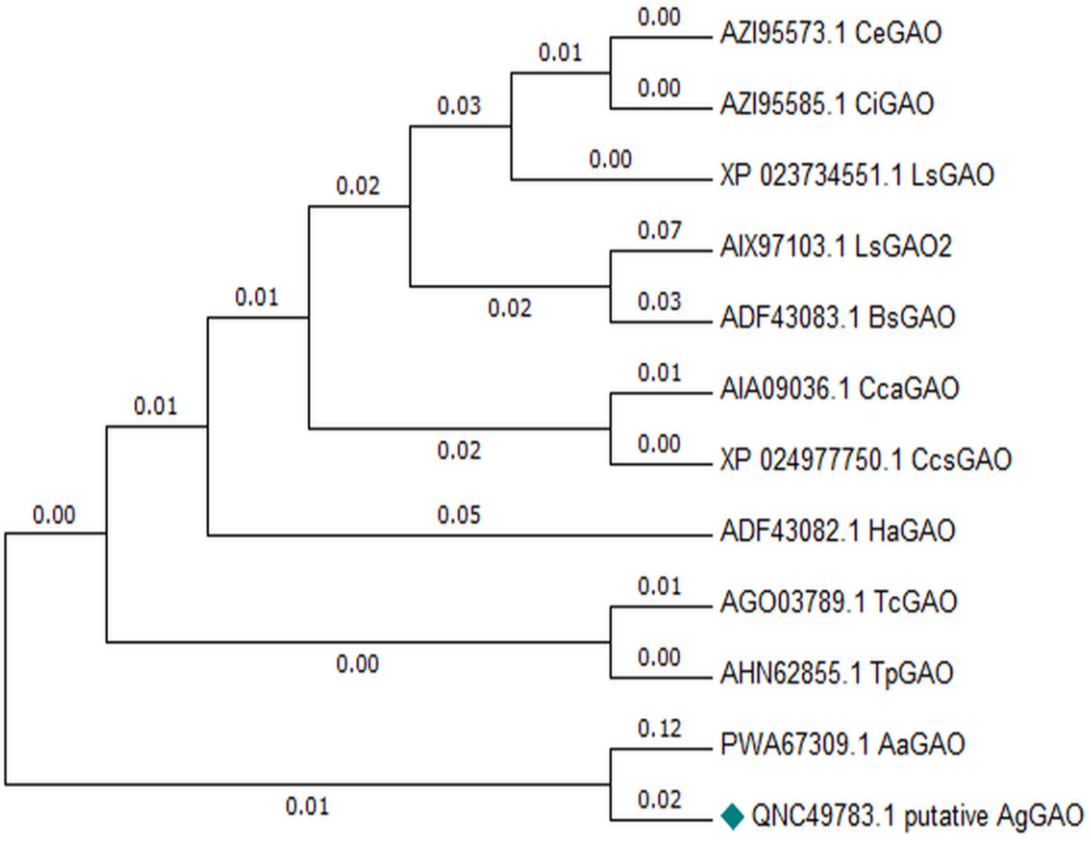


B

Figure 4. Pairwise protein sequence comparison (A) and phylogenetic tree (B) of putative *Artemisia glabella* germacrene A synthase (◆ AgGAS) and annotated germacrene A synthases. AaGAS; AaGAS2: *Artemisia annua*, TpGAS: *Tanacetum parthenium*, HaGAS1; HaGAS2: *Helianthus annuus*, LsGASS; LsGASL: *Lactuca sativa*, CiGASS; CiGASL: *Cichorium intybus*, ToGAS: *Taraxacum officinale*, BsGAS1; BsGAS2: *Barnadesia spinose*, CcGASS; CcGASL: *Cynara cardunculus var. scolymus*, ScGAS: *Solidago canadensis.* A: Lower left triangle shows pairwise evolutionary distance and upper right triangle shows percentage identity.

		1	2	3	4	5	6	7	8	9	10	11	12
AIX97103.1_LsGAO2	1		83	84	83	83	84	83	83	74	82	87	85
AGO03789.1_TcGAO	2	0.14		100	89	90	90	88	89	90	89	84	96
AHN62855.1_TpGAO	3	0.13	0.01		89	90	90	89	89	91	89	84	97
AIA09036.1_CcaGAO	4	0.15	0.05	0.05		89	99	89	89	82	87	83	92
XP_023734551.1_LsGAO	5	0.11	0.08	0.07	0.08		90	98	98	83	89	84	90
XP_024977750.1_CcsGAO	6	0.14	0.04	0.03	0.01	0.07		90	90	83	87	84	94
AZI95573.1_CeGAO	7	0.12	0.09	0.08	0.08	0.01	0.07		99	81	89	84	89
AZI95585.1_CiGAO	8	0.12	0.09	0.08	0.08	0.01	0.07	0.00		81	89	84	89
PWA67309.1_AaGAO	9	0.27	0.14	0.14	0.19	0.21	0.17	0.22	0.22		83	75	87
ADF43082.1_HaGAO	10	0.15	0.07	0.07	0.09	0.10	0.08	0.11	0.11	0.18		81	92
ADF43083.1_BsGAO	11	0.11	0.09	0.09	0.11	0.09	0.09	0.09	0.09	0.22	0.14		89
QNC49783.1_putative_AgGAO	12	0.16	0.04	0.03	0.08	0.11	0.07	0.11	0.11	0.14	0.09	0.11	

A

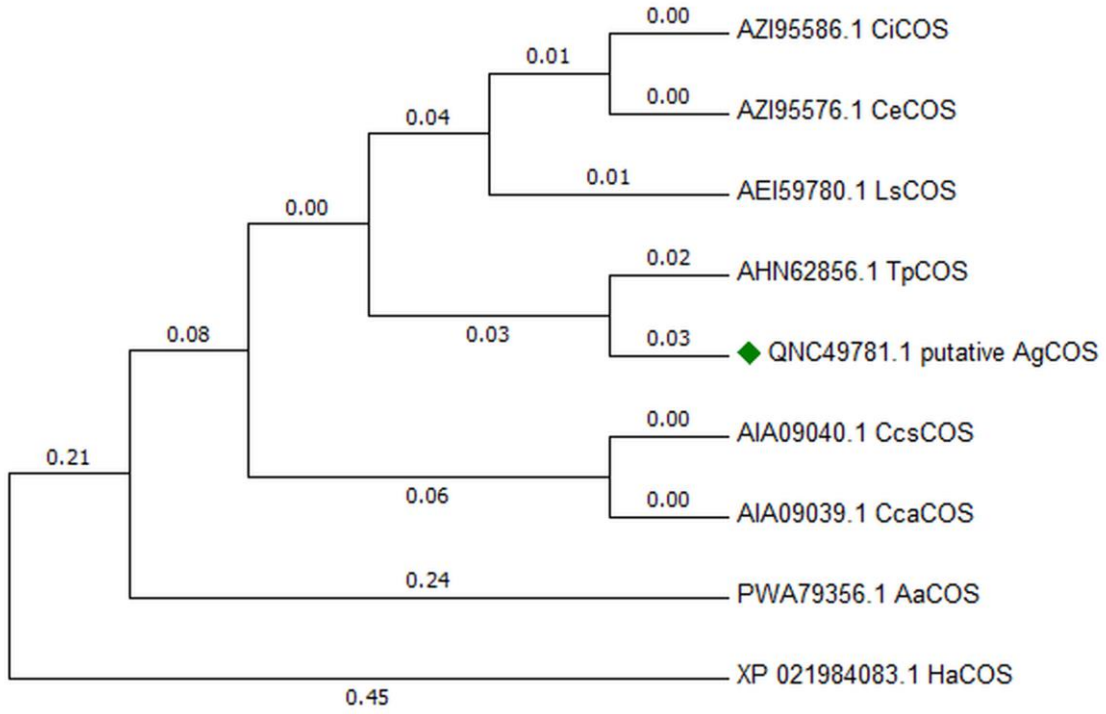


B

Figure 5. Pairwise protein sequence comparison (A) and phylogenetic tree (B) of putative *Artemisia glabella* germacrene A oxidase (◆ AgGAO) and annotated germacrene A oxidases. CeGAO: *Cichorium endivia*, CiGAO: *Cichorium intybus*, LsGAO, LsGAO2: *Lactuca sativa*, BsGAO: *Barnadesia spinose,* CcaGAO: *Cynara cardunculus var. altilis*, CcsGAO: *Cynara cardunculus var. scolymus*, HaGAO: *Helianthus annuus*, TcGAO: *Tanacetum cinerariifolium*, TpGAO: *Tanacetum parthenium*, AaGAO: *Artemisia annua*. A: Lower left triangle shows pairwise evolutionary distance and upper right triangle shows percentage identity.

		1	2	3	4	5	6	7	8	9
AEI59780.1_LsCOS	1		67	44	96	97	89	89	88	89
PWA79356.1_AaCOS	2	0.38		39	67	67	67	67	65	67
XP_021984083.1_HaCOS	3	0.81	0.91		43	43	43	43	42	45
AZI95586.1_CiCOS	4	0.02	0.37	0.80		100	88	88	88	89
AZI95576.1_CeCOS	5	0.02	0.37	0.80	0.00		88	88	88	89
AIA09040.1_CcsCOS	6	0.11	0.38	0.80	0.11	0.11		100	86	89
AIA09039.1_CcaCOS	7	0.11	0.38	0.80	0.11	0.11	0.00		86	89
AHN62856.1_TpCOS	8	0.11	0.39	0.79	0.11	0.11	0.12	0.12		95
QNC49781.1_putative_AgCOS	9	0.12	0.40	0.80	0.12	0.11	0.12	0.12	0.05	

A



B

Figure 6. Pairwise protein sequence comparison (A) and phylogenetic tree (B) of putative *Artemisia glabella* costunolide synthase (◆AgCOS) and annotated costunolide synthases. CiCOS: *Cichorium intybus*, CeCOS: *Cichorium endivia*, LsCOS: *Lactuca sativa*, TpCOS: *Tanacetum parthenium*, CcsCOS: *Cynara cardunculus var. scolymus*, CcaCOS: *Cynara cardunculus var. altilis*, AaCOS: *Artemisia annua,* HaCOS: *Helianthus annuus*. A: Lower left triangle shows pairwise evolutionary distance and upper right triangle shows percentage identity.

preparation for pharmaceutical production.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

This work was carried out according to grant project No. AP05134198 “The study of the biosynthesis of terpenoids in plants and the search for new pharmacologically active bimolecular compounds”, funded by the Science

Committee of the Ministry of Education and Science of the Republic of Kazakhstan.

REFERENCES

Adekenov SM (2001). Method and device for production of lyophilized hydrochloride-1β, 10β-epoxy-13-dimethylaminoguaia-3(4)-en-6,12-olide. US6242617B1. https://patents.google.com/patent/US6242617B1/en.

Adekenov SM (2015). Method for production of hydrochloride 1(10) beta-epoxy-13-dimethylamino-5,7-alpha,6,11-beta (h)-guaia-3(4)-en-6,12-olide, the lyophilized antitumor preparation "arglabin". EP2069357B1. https://patents.google.com/patent/EP2069357B1/en.

Adekenov SM, Bouwmeester HJ (2015). About the Biosynthesis of Sesquiterpene Lactone Arglabin. In: Achievements and Prospects for the Development of Phytochemistry, 27-28. Karaganda, Kazakhstan.

Adekenov SM, Mukhametzhanov MN, Kagarlitsky AD, Kupriyanov AN (1982). Arglabin – A new sesquiterpene lactone from *Artemisia glabella.* Chemistry of Natural Compounds 18(5):623-624.

Adekenov SM, Rakhimova BB, Dzhazin KA, Alikov VB, Shaushekov ZK, Toregozhina ZhR, Turdybekov KM (1995). Sesquiterpene lactones from *Artemisia glabella.* Fitoterapia 66(2):142-146.

Bouwmeester HJ, Kodde J, Verstappen FWA, Altug IG, de Kraker J-W, Wallaart TE (2002). Isolation and Characterization of Two Germacrene A Synthase cDNA Clones from Chicory. Plant Physiology 129(1):134-144.

de Kraker J-W, Franssen MCR, Joerink M, de Groot A, Bouwmeester HJ (2002). Biosynthesis of Costunolide, Dihydrocostunolide, and Leucodin. Demonstration of Cytochrome P450-Catalyzed Formation of the Lactone Ring Present in Sesquiterpene Lactones of Chicory. Plant Physiology 129(1):257-268.

Fischer NH, Czerson H, Bohlmann F, Stuessy TF (1979). Sesquiterpenoid and acetylenic constituents of seven Clibadium species. Phytochemistry 18(2):257-260.

Frederick CS (1982). Sesquiterpene lactones as taxonomic characters in the asteraceae. The Botanical Review 48(2):121–594.

Ikezawa N, Göpfert JC, Nguyen DT, Kim SU, O'Maille PE, Spring O, Ro DK (2011). Lettuce costunolide synthase (CYP71BL2) and its homolog (CYP71BL1) from sunflower catalyze distinct regio- and stereoselective hydroxylations in sesquiterpene lactone metabolism. The Journal of biological chemistry 286(24):21601-21611.

Kumar S, Stecher G, Tamura K (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33(7):1870-1874.

Liu Q, Kashkooli AB, Manzano D, Pateraki I, Richard L, Kolkman P, Lucas MF, Guallar V, Ric CH de Vos, Franssen MCR., Van der Krol A, Bouwmeester H (2018). Kauniolide synthase is a P450 with unusual hydroxylation and cyclization-elimination activity. Nature Communications 9(1):1-13.

Ma GH, Chen KX, Zhang LQ (2019). Advance in biological activities of natural guaiane-type sesquiterpenes. Medicinal Chemistry Research 28((9):1339-1358.

Majdi M, Liu Q, Karimzadeh G, Malboobi M, Beekwilder J, Cankar K, Todorovic S, Simonovic A, Bouwmeester HJ (2011). Biosynthesis and localization of parthenolide in glandular trichomes off ever few (*Tanacetum parthenium* L. Schulz Bip.). Phytochemistry 72(14-15):1739-1750.

Mantler SN, Zhakanov MM, Adekenov SM (2020). Biosynthesis and dynamics of accumulation of sesquiterpene lactones in *Artemisia glabella* Kar. et Kir. News of the national academy of sciences of the republic of Kazakhstan - Series chemistry and technology 4(442):22-29.

Menin B, Comino C, Portis E, Moglia A, Cankar K, Bouwmeester HJ, Lanteri S, Beekwilder J (2012). Genetic mapping and characterization of the globe artichoke (+)-germacrene A synthase gene, encoding the first dedicated enzyme for biosynthesis of the bitter sesquiterpene lactone cynaropicrin. Plant science: an international journal of experimental plant biology 190:1-8.

Nguyen TD, Faraldos JA, Vardakou M, Salmon M, O'Maille PE, Ro DK (2016). Discovery of germacrene A synthases in Barnadesia spinosa: The first committed step in sesquiterpene lactone biosynthesis in the basal member of the Asteraceae. Biochemical and biophysical research communications 479(4):622-627.

Nguyen TD, Kwon M, Kim SU, Fischer C, Ro DK (2019). Catalytic Plasticity of Germacrene A Oxidase Underlies Sesquiterpene Lactone Diversification. Plant physiology 181(3):945-960.

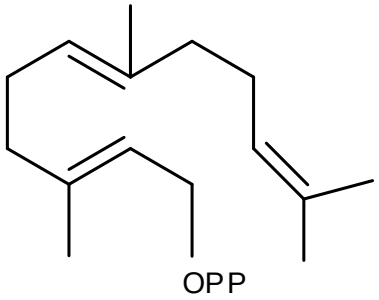
Pazouki L, Memari HR, Kännaste A, Bichele R, Niinemets Ü (2015). Germacrene A synthase in yarrow (Achillea millefolium) is an enzyme with mixed substrate specificity: gene cloning, functional characterization and expression analysis. Frontiers in plant science 6:111.

Saitou N, Nei M (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4(4):406-425.

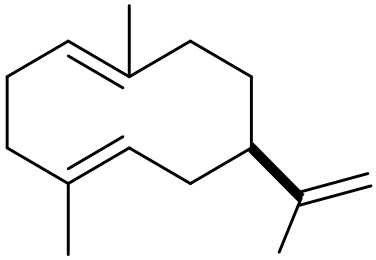
Xu H, Dickschat JS (2020). Germacrene A – A Central Intermediate in Sesquiterpene Biosynthesis. Chemistry (Weinheim an der Bergstrasse, Germany) 26(72):17318-17341.

Zuckerkandl E, Pauling L (1965). Evolutionary divergence and convergence in proteins. In: Evolving Genes and Proteins. Ed. by Bryson V and Vogel HJ, Academic Press pp. 97-166, New York. https://doi.org/10.1016/C2013-0-11981-2

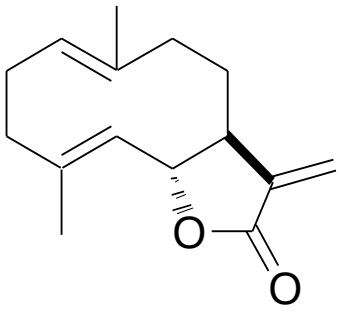
SUPPLEMENTARY



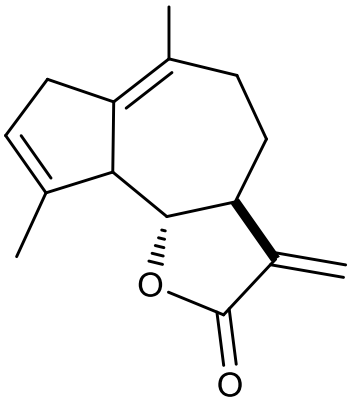
(1)



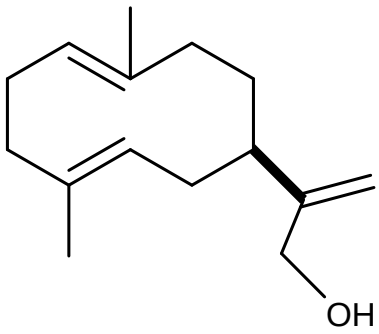
(2)



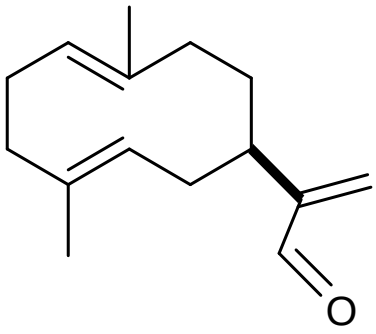
(**4**)



(**5**)



(**7**)



(**8**)