Identification of Brassicadiene, a Diterpene Hydrocarbon Attractive to the Invasive Stink Bug *Bagrada hilaris,* from Volatiles of Cauliflower Seedlings, *Brassica oleracea* var. *botrytis*.

Kyle Arriola,[†] Salvatore Guarino,[‡] Mokhtar Abdulsattar Arif,[§‡] Stefano Colazza,[§‡] Ezio Peri,[§‡] Jocelyn G. Millar[†*]

[†] Departments of Chemistry and Entomology, University of California, Riverside, California 92506, USA

[‡] Institute of Biosciences and Bioresources (IBBR), National Research Council of Italy (CNR), Corso Calatafimi 414, 90129 Palermo, Italy

[§] Università degli Studi di Palermo, Dipartimento di Scienze Agrarie, Alimentari e Forestali, Viale delle Scienze edificio 5, 90128, Palermo, Italy

Supporting Information Placeholder

ABSTRACT: Brassicadiene, a novel tricyclic diterpene hydrocarbon, was identified by a combination of mass spectrometry, microchemical tests, and analysis of NMR spectra. The compound constitutes >90% of the volatile organic compounds (VOCs) produced by cauliflower seedlings, *Brassica oleracea* var. *botrytis*. The invasive stink bug *Bagrada hilaris* is strongly attracted to brassicadiene, providing a mechanism for this herbivore, which specializes on cruciferous plants, to locate its hosts at a nutrient-rich and vulnerable stage.

The stink bug *Bagrada hilaris* Burmeister (Heteroptera: Pentatomidae), endemic to parts of Asia and Africa, attacks a variety of plant species in the genus *Brassica*, including a number of important crops such as broccoli, cauliflower, and kale.[1,2] This pest invaded North America approximately 10 years ago, spreading rapidly through the southern United States, and then further south through Central America and into South America.[2] Wherever it has invaded, it has caused substantial damage to crops. A recent study showed that this bug is particularly attracted to seedlings of cruciferous plants, which are frequently killed by bug feeding activity[3]. In preliminary work, we showed that crude extracts of the volatile organic compounds (VOCs) released by undamaged seedlings of *Brassica oleracea* var. *botrytis* and *B. napus* were highly attractive to *B. hilaris* adults. Bioassay-guided fractionation of the crude extracts revealed that the active compound(s) were contained in the nonpolar hydrocarbons fraction eluted from silica gel with hexane.[3] The major component (>90%) of this fraction was tentatively identified as a diterpene on the basis of several pieces of information. First, its molecular weight of 272 amu suggested a likely molecular formula of $C_{20}H_{32}$, which in turn would require five rings or sites of unsaturation. Second, hydrogenation with palladium on carbon catalyst produced a derivative with a molecular weight of 276 amu, indicating that there were two sites of unsaturation and hence three rings. Third, the electron impact ionization (EI) mass spectra of both the parent compound (Figure S1) and the derivative (Figure S2) showed fragmentation patterns typical of terpenoids, and both were dominated by a base peak from loss of 43 mass units, suggestive of an isopropyl group that was readily lost from both structures. These structural features and further fragmentary information obtained from the ^{1}H NMR spectrum did not match any of the ~1,000 known diterpene hydrocarbon structures listed in SciFinder™, but at the time, we were unable to make further progress on its identification.

A larger sample of the compound was isolated from a composite crude extract, obtained by combining multiple VOC collections from *B. oleracea* var. *botrytis* seedlings, by liquid chromatography on silica gel, eluting with hexane. Assignment of protons to carbon centers with a ^{1}H-^{13}C-HSQC spectrum showed the presence of 5 methyl groups, one of which was probably allylic from its chemical shift (1.75 ppm) and broadening, two of which were the diastereotopic methyls of an isopropyl group, one of which exhibited as a doublet indicating that it was attached to a methine, and one of which was a sharp singlet, indicative of a methyl attached to a quaternary center. There were six methylenes, each consisting of two diastereotopic protons with different chemical shifts. The five methine protons included two alkenyl protons, the proton of the isopropyl group, and two additional protons on sp^3 carbons. The remaining four quaternary carbons included two sp^2 and two sp^3 carbons. The chemical shifts of the carbons and their attached protons are given in Table 1, numbered in

Formattato: Inglese (Stati Uniti)

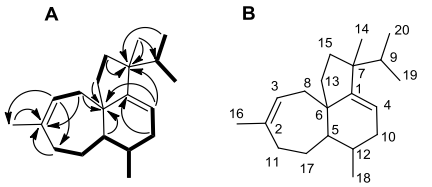

Figure 1. A) Structure elucidation by 2D-^{1}H-^{1}H-DQFCOSY (bold) and some 2D-^{13}C-^{1}H-HMBC correlations (arrows); B) the basic structure assembled with carbons numbered in descending magnitude of chemical shift (see Table 1).

~~order~~ Order of decreasing chemical shift of the carbons, and the placement of the carbons is shown in Figure 1.

The basic proton spectrum was comprised of four spin systems terminating at quaternary carbons (Figure 1A). Elucidation began with consideration of the two alkenes. The two alkenyl protons were not coupled to each other, indicating that there were two trisubstituted alkenes. The first of these was substituted with the allylic methyl group CH_3-16 with chemical shift 1.75 ppm, and displayed a broadened singlet. Data from the ^{1}H-^{13}C HMBC spectrum showed strong correlations between C-2 to H_3-16, C-2 to H_2-11, and C-3 to H_3-16, establishing the C-2 to C-3 alkene fragment.

Table 1. ^{13}C and ^{1}H data of brassicadiene.

δ (^{13}C)/ ppm	Carbon type	Position	δ (^{1}H) ppm, Mult	J(H,H) /Hz
156.14	>C=	1	-	
136.56	>C=	2	-	
122.19	-CH=	3	5.30 m	
118.32	-CH=	4	5.50 dd	J=8.2, 2.5
55.29	>CH-	5	1.43 td	J~11.5, ~11.5, 2.1
51.81	>C<	6	-	
45.40	>C<	7	-	
43.23	-CH_2-	8α	2.59 dd	J=16.4, 6.9
		8β	1.77 br d	J~16.4
39.06	>CH-	9	1.70 m	
36.97	-CH_2-	10α	1.99 ddd	J=15.8, 8.2, 3.3
		10β	1.73 ddd	J=15.8, 10.9, 2.5
35.36	-CH_2-	11α	2.28 ddd	J=17.0, ~5.9, ~5.2
		11β	2.10 br ddd	J=17.0, ~7, ~7
33.88	>CH-	12	1.16 m	
32.22	-CH_2-	13α	1.79 dd	J=11.6, 6.9
		13β	1.12 m	
30.80	-CH_3	14	1.15 s	
29.95	-CH_2-	15α	1.66 dd	J=13.6, 8.0
		15β	1.29 m	
28.82	-CH_3	16	1.75 br	
25.70	-CH_2-	17α	1.86 dddd	J=13.7, 7.0, 5.2, 2.4
		17β	1.32 m	
20.03	-CH_3	18	0.92 d	J=6.5
19.21	-CH_3	19	0.73 d	J=6.8
18.70	-CH_3	20	0.89 d	J=6.7

There were additional HMBC correlations between C-2 and allylic H_2-8, supported by $^{3}J_{H\text{-}H}$-couplings between H-3 and H-8α in the ^{1}H spectrum, establishing the bond C_3─C_8. The fact that there were no obvious additional $^{3}J_{H\text{-}H}$-couplings to H-8β or H-8α suggested that C-8 terminated at a quaternary carbon on the other side, providing one end of the spin system.

The connection of C-2 to allylic CH_2-11 was verified by the HMBCs of C-11 to H_3-16 and C-3 to H_2-11. $^{3}J_{H\text{-}H}$-couplings in the ^{1}H-NMR spectrum then established the successive connectivity between C_{11}─C_{17}─C_5─C_{12}─C_{10}─C_4. The alkenyl unit CH-4, by default, had to be connected to the remaining quaternary alkenyl carbon C-1. In addition, the ^{1}H and ^{1}H-^{1}H COSY spectra showed H-12 was coupled to H_3-18, and H-5 was only coupled to protons on H_2-17 and H-12, suggesting that the final connection of CH-5 was to a quaternary carbon. The complete spin system is shown in Figure 1A.

The third spin system consisted of an isolated isopropyl group, characterized by the EI mass spectrum and $^{3}J_{H\text{-}H}$-couplings between H-9 with H_3-19 and H_3-20, with no other direct couplings to H-9, establishing that it was attached to a quaternary carbon. To establish further connectivity, strong HMBCs were observed between H_3-19, H_3-20, and C-7 in addition to a correlation between C-7 to H-9, producing the connection of C_9─C_7. The possibility of the sharp methyl singlet H_3-14 being connected to C-7, as opposed to the other quaternary centers C-1 and C-6, was corroborated with HMBCs of H_3-14 to C-15, C-7, and C-9, its relatively low chemical shift (1.15 ppm), and the absence of significant HMBCs to C-4 and C-6. This established the connectivity of C_{14}─C_7─C_9─$C_{19/20}$. Additional HMBCs of C-7 to H_2-15 and H-4, and across the quaternary center C-7, such as the correlations C-1 to H_3-14 and C-15 to H_3-14, allowed the identification of the connections C_1─C_7─C_{15} and C_7-C_1-C_4. The fourth and final spin system consisted of methylenes C_{13}─C_{15} that are attached on both ends to quaternary centers, where one had already been identified above as C_7─C_{15} and the other was to be determined.

With four single-bond connections and a lone sp^3-quaternary carbon (C-6) remaining, HMBCs between C-6 and H_2-8, H-3, H_2-13, H-4, and H-5, gave a strong indication that C-6 was directly connected to carbons 8, 13, 1, and 5. In sum, this established the basic structure as the fused 5-6-7-spirotricyclic skeleton shown in Figure 1B, with the unusual moiety of an isopropyl and a methyl group attached to the same sp^3-quaternary center.

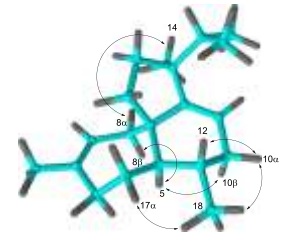

Figure 2. Model of brassicadiene showing key NOE enhancements.

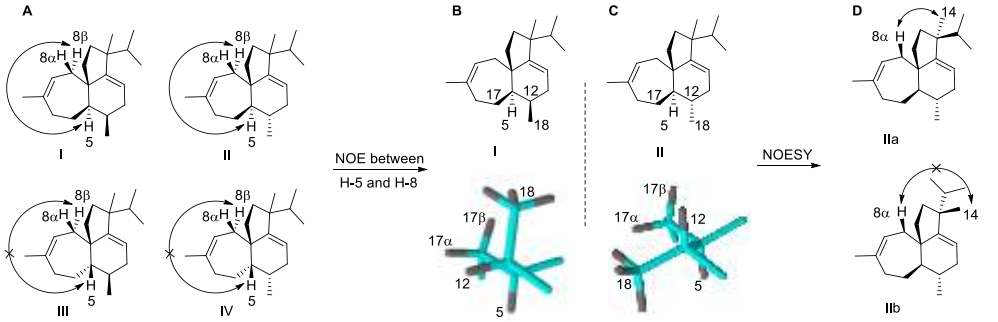

Figure 3. A) The four diastereomers selected to begin elucidation of the relative stereochemistry. B) The simulated *anti*-diastereomer and C) the simulated *syn*-diastereomer in regard to the relative stereochemical relationship between H-5 and C-18, showing key dihedral angles among protons adjacent to H-5. D) Elucidation of the relative stereochemistry of the final stereocenter, showing NOE between H-8α and H_3-14.

Assignment of the relative stereochemistry (Figure 2) was facilitated by using a digital molecular modeling program (ChemSketch®) to simulate the shape and orientation of the molecule and its atoms. From the simulated structures, the relative stereochemistry was elucidated beginning with the four diastereoisomeric forms I-IV which varied in the relative configurations of the three contiguous stereocenters on carbons 6, 5, and 12 (Figure 3A), using data from the $^3J_{H\text{-}H}$-coupling constants and a NOESY spectrum. A strong NOE enhancement was observed between H-5 and H-8β, eliminating possibilities III and IV where the C_6─C_{13} bond is *syn* with respect to the CH-5 bond. Next, the spatial relationships between H-5 and the flanking H-12, H-17β, and H-17α were established from the magnitudes of the respective coupling constants. In particular, the pseudo-triplet of doublets of H-5 corresponded to two large couplings of approximately equal value (J≃11.5 Hz) to H-12 and H-17β, indicating that they were approximately parallel or antiparallel to the CH-5 bond, whereas the remaining small coupling (J≃2.1 Hz) between H-5 and H-17α indicated a roughly orthogonal relationship between the CH-5 and the CH-17α bonds. For simulated structure I with an *anti*-relationship between H-5 and C-18 (Figure 3B), the expected splitting pattern of H-5 would best be described as a ddd, owing to the approximate dihedral angles of 180º, 55º, and 90º relative to the flanking protons H_2-17 and H-12. This would not support the two large coupling constants of ~11.5 Hz. Conversely, structure II with a *syn*-relationship between H-5 and C-18 (Figure 3C) exhibits two pseudo-antiperiplanar relationships, specifically between CH-5 with CH-12, CH-5 with CH-17β, and an approximately 90º dihedral angle between CH-5 and CH-17α, which would accommodate the pseudo-triplet of doublets with observed coupling constants of ~11.5, ~11.5, and 2.1 Hz. Taken together, these data established the relative stereochemistry of the three contiguous stereocenters as that shown in structure II in Figure 3C.

The relative stereochemistry of the fourth and final stereocenter was determined to be as shown in structure IIa in Figure 3D, from the strong NOE enhancement observed between H_3-14 and H-8α, with a separation estimated from the model of 1.8 Å. Such an enhancement would be improbable in the alternative stereoisomer IIb, where the methyl group and H-8α are much farther apart (estimated 4.9 Å). Furthermore, there was no NOE between either H_2-8 and H_3-19 or H_3-20, as would be expected for the alternative stereoisomer IIb. Thus, the data best supported the structure represented by IIa, to which we assign the common name brassicadiene.

In addition to the unusual feature of methyl and isopropyl groups being attached to the same quaternary center in a terpenoid, the connectivity of the 5,6,7-tricyclic ring structure of brassicadiene also appears to be rare. A search of the literature revealed only one other known terpenoid with this particular pattern of fused rings, the sesterterpenoid gascardic acid (Figure 4), produced by the scale insect *Gascardia madagascariensis* Targioni Tozzetti.[4,5] Furthermore, treatment of gascardic acid with acid had been reported to result in, among other rearrangements, a 1-2 shift of the quaternary methyl group to the ternary carbon bearing the side chain, while simultaneously introducing a new double bond, creating a core structure analogous to that in brassicadiene.[4,5] A plausible biosynthetic pathway to brassicadiene from geranylgeranyl diphosphate is shown in Figure S4.

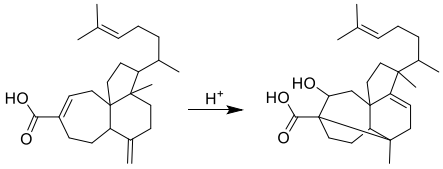

Figure 4. Treatment of gascardic acid with acid resulted in a 1-2 methyl shift and introduction of a double bond, producing a core structure analogous to that in brassicadiene.

Brassicadiene exhibited a specific rotation of approximately $[a]_D^{20}$ +21° (*c* 0.17, CH_2Cl_2), but because of the small amount of material available, and the lack of a conjugated chromophore in the structure to provide strong absorbtions at accessible wavelengths, we have not attempted to determine its absolute configuration by ECD or related methods.

Plants in the family Brassicacae are known for being chemically defended from generalist herbivores by glucosinolates.[6] Tissue damage from herbivore feeding releases myrosinases, which cleave the glucosinolates into smaller, strongly irritating isothiocyanates and related chemicals, the odorous compounds typically associated with mustards, cabbage, cauliflower, and other cruciferous plants. Thus, it was unexpected to find that the headspace odors of undamaged cauliflower seedlings consisted largely of a novel diterpene hydrocarbon, along with trace amounts of several isomers.[3] The stink bug *B. hilaris*, a specialist herbivore which exhibits a strong preference for newly emerged seedlings of crucifers,[1,7] has apparently evolved to use brassicadiene as a host location cue. In previously reported bioassays, the hexane fraction of the plant VOCs consisting of >90% brassicadiene was highly attractive to adult bugs,[3] and we confirmed this attraction with Y-olfactometer bioassays with the purified compound (Figure S3). The attraction behavior of *B. hilaris* toward brassicadiene encourage the use of this molecule as candidate attractant lure for trapping this insect in the field, also in consideration that the monitoring methods based on pheromone traps for this species did not give consisting results so far.[8] The fact that *B. hilaris* exploits this compound for host location is intriguing, given that diterpene hydrocarbons are generally thought to contribute to plant defenses against herbivores,[8-10] phytopathogenic fungi,[11,12,14,15] and nematodes.[15,16,17] The possible functions of brassicadiene within the plants that produce it remain to be determined.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Isolation of brassicadiene, EI mass spectra of brassicadiene and its reduction product, accurate mass measurement of brassicadiene, specific rotation measurement, bioassay methods and results, tables of other significant HMBC and NOE correlations, 1D (^{1}H and ^{13}C) and 2D (1H-1H, DQFCOSY, 1H-1H-HMBC, 1H-13C-HSQC, 1H-1H-NOESY) NMR spectra, GC-FT-IR spectrum of brassicadiene, and a possible biosynthetic scheme (PDF).

AUTHOR INFORMATION

Corresponding Author

* Email: millar@ucr.edu

ORCID

Jocelyn G. Millar: 0000-0001-7639-9001

Author Contributions

SG, MAA, SC, and EP prepared the crude extracts; KA and JGM isolated and identified the compound. The manuscript was written and edited jointly by all authors, and the final manuscript has been read and approved by all authors.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

SG, MAA, SC, and EP acknowledge the Project "High Studies Programme (Master and Ph.D.) for Iraqi officials at Italian Academic Institutions (HiSPro)" co-financed by the Iraqi Government and the Italian Ministry of Foreign Affairs (MAE). JGM acknowledges financial support from Hatch project CA-R*ENT-5181-H. We thank Prof. Paulo H. G. Zarbin, Universidade Federal do Paraná, Brazil, for obtaining the GC-FTIR spectrum, and Prof. Stefan Schulz and Christian Schlawis, University of Braunschweig, for calculating simulated ECD spectra.

REFERENCES

(1) Huang, T. I.; Reed, D. A.; Perring, T. M.; Palumbo, J. C. *Arthropod-Plant Interactions* **2014**, *8*, 89-100.

(2) Bundy, C. S.; Perring, T. M.; Reed, D. A.; Palumbo, J. C.; Grasswitz, T. R.; Jones, W. A. *Bagrada hilaris.* In: McPherson J. E. (ed) Invasive Stink Bugs and Related Species (Pentatomoidea): Biology, Higher Systematics, Semiochemistry, and Management, CRC Press, Boca Raton, FL **2018**; pp. 205-242.

(3) Guarino, S.; Arif, M. A.; Millar, J. G.; Colazza, S.; Peri, E. *PloS one*, **2018**, *13*(12), p.e0209870.

(4) Scartazzini, R. Gascardinsaure, ein Sester-terpen. **1966**. PhD. Dissertation No. 3899, ETH (Zurich).

(5) Boeckman, Jr., R. K.; Blum, D. M.; Arnold, E. V.; Clardy, J. *Tetrahedron Lett.* **1979**, *20*, 4609-4612.

(6) Agerbirk, N.; Olsen, C. E. *Phytochemistry* **2012**, *77*, 16-45.

(7) Ludwig, S.; Kok, L. *Crop Protec*. **2001**, *20*, 247-51.

(8) Arif, M. A.; Guarino, S.; Colazza, S.; Peri, E. *Insects* **2020**, *11*, 109.

(9) Gershenzon, J.; Dudareva, N. *Nat. Chem. Biol.* **2007**, *3*, 408.

(10) Chen, X. Q.; Gao, L. H.; Li, Y. P.; Li, H. M.; Liu, D.; Liao, X. L.; Li, R. T. *J. Agr. Food Chem*. **2017**, *65*, 4456-4463.

(11) Hua, J.; Luo, S. H.; Liu, Y.; Liu, Y. C.; Tan, Y. Y.; Feng, L.; Xiao, C. J.; Zhang, K. Q.; Li, S. H.; Niu, X. M. *Chem. Biodivers*. **2017**, *14*(10):e1700327.

(12) Rhoades, D. F. Evolution of plant chemical defenses against herbivores. In Janzen, D. H.; Rosenthal, G. A. (eds.). Herbivores, Their Interaction with Secondary Plant Metabolites. Academic Press Inc., New York 1979; pp. 3-54.

(13) Koga, J.; Ogawa, N.; Yamauchi, T.; Kikuchi, M.; Ogasawara, N.; Shimura, M. *Phytochemistry* **1997**, *44*, 249-253.

(14) Salah, M. A.; Bedir, E.; Toyang, N. J.; Khan, I. A.; Harries, M. D.; Wedge, D. E. *J. Agr. Food Chem*. **2003**, *51*, 7607-7610.

(15) Bi, Y.; Yu, Z. *J. Agr. Food Chem*. **2016**, *64*, 8525-8529.

(16) Sun, Y. J.; Gao, M. L.; Zhang, Y. L.; Wang, J. M.; Wu, Y.; Wang, Y.; Liu, T. *Molecules* **2016**, *21*, 434.

(17) Liu, T.; Meyer, S. L.; Chitwood, D. J.; Chauhan, K. R;, Dong, D.; Zhang, T.; li, J.; Liu, W. *J. Agr. Food Chem*. **2017**, *65*, 3127-3132.