ORIGINAL ARTICLE

Using quantum chemical computations of NMR chemical shifts to assign relative configurations of terpenes from an engineered *Streptomyces* host

Q Nhu N Nguyen and Dean J Tantillo

Density functional theory calculations of ¹H and ¹³C chemical shifts are used to assign the relative configurations of isohirsut-1-ene, isohirsut-4-ene and tsukubadiene, three terpenes previously isolated from an engineered *Streptomyces*.

INTRODUCTION

The diversity among known terpenes and terpene derivatives (terpenoids) is a testament to the power of terpene synthases in transforming simple biosynthetic building blocks into complex, stereodense natural products.[1–7] Despite decades of work, scientists continue to isolate and identify terpenes and terpenoids with new carbon frameworks from plants, fungi and bacteria.[2–10] Moreover, the capacity of wild-type terpene synthases can be further magnified through genetic engineering. For instance, Cane, Ikeda and co-workers recently described the frameworks of a variety of new terpenes isolated from an engineered *Streptomyces* host, although relative configurations for many were not fully assigned.[11] Here we show how NMR chemical shifts computed with density functional theory (DFT) methods can be used to assign the relative configurations of such structures, using the three structures in Scheme 1 as representative examples.[12–14] This work builds on our past work on confirming, assigning and reassigning the structures of terpenes[15–17] and terpenoids,[18,19] and provides a missing piece of the structural puzzle defined by Cane and Ikeda. This work represents only the second case of assigning the relative stereochemical configuration of complex hydrocarbon natural products, that is, those lacking polar functional groups, using DFT.[15]

RESULTS AND DISCUSSION

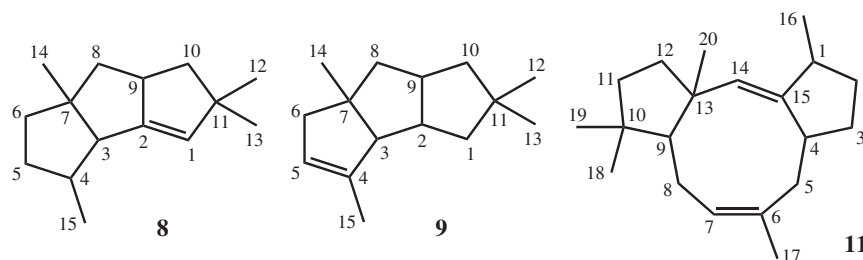### Sesquiterpene 8 (isohirsut-1-ene)

As structure **8** has four stereogenic centers, there are eight possible diastereomers (Scheme 2; each has a corresponding enantiomer, but absolute configuration is not addressed here). First, we constructed and optimized the structures of the eight diastereomers with B3LYP/6-31+G(d,p) in the gas phase (see Methods section). The resulting geometries were then subjected to NMR calculations using SMD(CHCl₃)–mPW1PW91/6-311+G(2d,p). Only structure **8F** was found to have more than one conformer close enough in energy to contribute to the overall chemical shifts (here there were three); the shifts reported in Table 1 for **8F** therefore represent a

Boltzmann-weighted average. The mean absolute deviations (MADs) between the experimental and calculated values were determined for each set of carbon and proton chemical shifts for each diastereomer versus the reported experimental values and are shown in Table 1.[11] Structure **8D** is the best match to the experimental data, based on MAD values for both ¹H (0.10 p.p.m.) and ¹³C (1.60 p.p.m.). Structure **8D** also does not contain any large outliers when comparing to the experimental values. To confirm our assignment, DP4 analysis was performed. This statistical analysis, which makes comparisons on an atom-by-atom basis, indicated that **8F** is actually the best match (87.4% probability) in terms of ¹H shifts (note that **8F** also has a small ¹H MAD: 0.12 p.p.m.), but **8D** is the best match in terms of ¹³C shifts (99.6%) and when both ¹H and ¹³C shifts are combined in the same analysis (98.8%). Although there is no consensus as to which shifts are generally more diagnostic—¹H or ¹³C—it is clear that ¹³C shifts are for compound **8** (based on both MAD and DP4 analysis).

Sesquiterpene 9 (isohirsut-4-ene)

Compound **9** is a constitutional isomer of compound **8**. As a result, in addition to assigning the relative configuration as we did for **8**, we can also confirm that our analysis, simply based on chemical shifts (even in the absence of coupling constants, DEPT data and so on), would differentiate between the hydrocarbon frameworks of **8** and **9**. Compound **9** also has four stereogenic centers, again giving rise to eight diastereomers (Scheme 3). Conformational searching yielded three distinct conformers for **9A** and **9D** after optimization, but for each case, two of these were >3 kcal mol⁻¹ higher in free energy than the lowest-energy conformer, meaning that they would not contribute significantly to the conformational population. Diastereomer **9D** is clearly the best match, in terms of ¹H or ¹³C MADs and DP4 probabilities, for the experimental data (Table 2).

Also, as reflected in the data in Table 3, our computed chemical shifts clearly distinguish between the two constitutional isomers **8** and **9**. We compared the conformers that were predicted to be the best

Department of Chemistry, University of California—Davis, Davis, CA, USA
Correspondence: Professor DJ Tantillo, Department of Chemistry, University of California—Davis, 1 Shields Avenue, Davis, CA 95616, USA.
E-mail: djtantillo@ucdavis.edu

Scheme 1 Two sesquiterpenes and a diterpene from reference 11; structure numbers correspond to the numbers used in this reference.
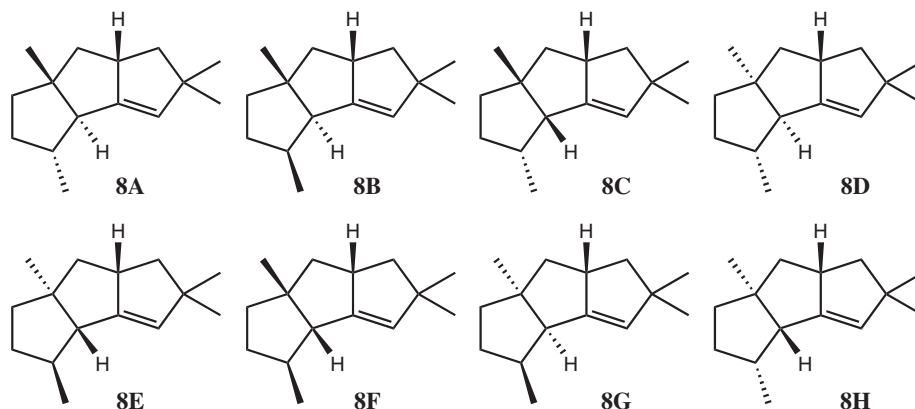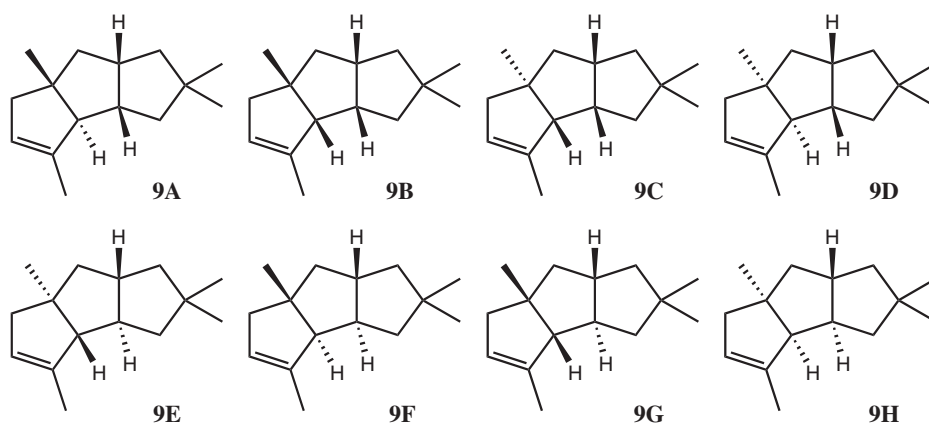

Scheme 2 Eight diastereomers of compound **8**. Note that, although we do not address herein absolute configuration, the configuration of C9 is shown as identical for all diastereomers to facilitate comparisons.

Table 1 Differences between computed and experimental (center column) chemical shifts for each atom in each of the diastereomers of compound 8 (computed minus experimental)

	8A	8B	8C	8D	Exp. δ (p.p.m.)	8E	8F	8G	8H
C1	−0.49	−2.60	0.02	−0.46	127.4	−2.31	−1.78	2.96	−2.50
C10	−0.69	−0.83	−2.98	−1.54	47.7	0.32	−2.02	−1.40	0.24
C11	2.46	2.42	3.74	2.86	50.8	0.98	3.08	2.90	1.76
C12 or C13	−4.74	−4.63	−3.29	−1.99	28.10	−0.22	−2.58	−1.99	−3.15
C12 or C13	−2.34	−2.34	−1.74	−2.10	30.10	−5.00	−2.08	−2.09	−2.07
C14	−11.07	−9.71	−3.03	−2.58	30.1	−10.09	−2.20	−3.55	−8.79
C15	−1.24	−2.80	−5.85	−1.79	20.1	−1.46	−1.30	−4.53	−2.03
C2	−2.77	−3.31	−0.36	2.03	154.1	−3.58	4.27	0.55	−5.21
C3	4.01	2.32	−4.55	0.04	57.7	3.69	0.42	−3.60	−0.36
C4	−11.46	−13.26	−3.37	1.75	43.7	−13.49	0.80	−4.00	−12.91
C5	1.08	1.83	−1.04	0.19	35.1	1.57	2.43	−1.16	3.02
C6	−6.56	−4.95	−0.75	−0.60	41.2	−6.41	0.30	−3.40	−4.46
C7	4.49	3.32	3.96	3.32	55.3	2.61	4.34	4.52	2.47
C8	−5.99	−5.30	−2.84	−1.04	47.8	−9.38	0.15	−2.50	−8.96
C9	9.50	9.92	3.62	1.71	47.9	7.45	6.28	3.97	7.36
MAD	4.59	4.64	2.74	1.60		4.57	2.27	2.87	4.35
H1	0.02	−0.08	−0.04	0.03	5.01	0.05	0.04	0.01	0.08
H10	0.54	0.54	0.36	0.44	0.84	0.57	0.44	0.44	0.58
H10	−0.06	−0.06	−0.06	0.03	1.76	0.05	−0.06	0.00	0.03
H12 or H13	0.03	0.03	0.01	0.01	1.02	0.00	0.02	0.02	0.03
H12 or H13	−0.03	0.00	−0.02	−0.03	1.08	−0.01	−0.02	−0.01	0.01
H14	−0.14	−0.05	0.08	−0.04	1.06	−0.32	0.04	−0.01	−0.15
H15	0.04	−0.06	0.05	−0.01	1.03	−0.03	0.01	−0.28	0.00
H3	0.61	1.10	0.68	0.09	1.69	0.61	0.23	0.77	1.01
H4					not reported				
H5	0.22	0.60	0.12	0.08	1.2	0.30	0.25	0.12	0.57
H5	0.44	0.59	−0.23	−0.05	1.67	0.56	0.12	0.23	0.61
H6	−0.21	−0.25	−0.02	0.00	1.5	−0.14	0.02	−0.05	−0.01
H6	−0.16	−0.07	−0.22	0.09	1.5	−0.20	0.00	0.08	−0.19
H8	−0.16	−0.17	−0.27	−0.34	1.22	−0.25	−0.19	−0.41	−0.24
H8	−0.24	−0.32	−0.21	−0.07	1.82	−0.24	−0.10	−0.07	−0.25
H9	0.57	0.63	0.01	0.12	3.16	0.29	0.21	−0.07	0.24
MAD	0.23	0.30	0.16	0.10		0.24	0.12	0.17	0.27
DP4 (combined, %)				98.8			1.2		
DP4 (13C only, %)				99.6			0.2	0.2	
DP4 (1H only, %)				12.6			87.4		

Abbreviation: MAD, mean absolute deviation.
MAD values and DP4 probabilities are also included.

Scheme 3 Eight diastereomers of compound **9**. Note that, although we do not address herein absolute configuration, the configuration of C9 is shown as identical for all diastereomers to facilitate comparisons.

Table 2 Differences between computed and experimental (center column) chemical shifts for each atom in each of the diastereomers of compound 9 (computed minus experimental)

	9A	9B	9C	9D	Exp. δ (p.p.m.)	9E	9F	9G	9H
C1	−3.45	−8.58	−6.59	−1.80	48.8	−7.99	−7.71	−5.86	−8.83
C10	−0.73	−4.58	0.15	−2.94	47.4	−5.17	−6.72	−5.63	−5.88
C11	6.36	2.12	3.50	2.85	41.6	5.77	4.92	5.44	4.46
C12 or C13	−4.25	−0.89	−3.29	−2.33	29.2	0.22	−0.53	0.33	0.18
C12 or C13	−3.12	−1.37	−3.40	−2.15	30.5	−1.08	−1.31	−0.93	−1.03
C14	−10.62	−1.76	−9.29	−2.04	29.4	−6.55	−9.43	−0.73	−1.52
C15	0.06	0.91	0.20	0.31	15.9	0.29	0.65	0.26	1.14
C2	−8.71	−1.80	−7.13	2.21	50.1	−3.42	−3.02	10.53	5.53
C3	2.79	−0.86	−2.84	−0.37	66.7	−4.68	−6.77	−1.47	−6.48
C4	2.66	1.68	2.06	3.22	142.8	2.96	3.73	2.78	0.11
C5	6.39	1.95	6.87	0.80	121.8	6.50	4.34	0.70	3.88
C6	−6.39	1.00	−3.86	−0.89	47.5	−5.03	−5.26	1.41	1.75
C7	8.81	0.83	7.15	3.07	52.9	10.60	10.66	6.40	6.27
C8	−6.96	0.69	−9.04	0.03	48.9	−16.03	−10.42	−4.54	−5.11
C9	7.35	2.51	8.43	0.74	44.3	13.65	10.29	7.40	4.69
MAD	5.24	2.10	4.92	1.72		6.00	5.72	3.63	3.79
H1	0.10	−0.22	−0.10	−0.13	1.58	−0.14	−0.10	−0.07	−0.25
H1	−0.14	0.01	0.05	0.04	1.3	−0.33	−0.23	−0.15	−0.18
H10	−0.06	−0.07	0.22	−0.03	1.27	−0.15	−0.22	−0.29	−0.35
H10	−0.06	−0.13	0.00	−0.06	1.64	−0.27	−0.34	−0.31	−0.33
H12 or H13	−0.02	−0.02	−0.01	−0.02	0.92	0.07	0.08	0.07	0.05
H12 or H13	−0.04	−0.06	−0.05	−0.04	1.06	−0.05	−0.08	−0.04	−0.09
H14	−0.30	−0.17	−0.13	−0.11	1.17	−0.25	−0.14	−0.14	−0.04
H15	0.05	0.03	0.12	0.02	1.68	0.04	0.16	−0.01	−0.03
H2	0.06	0.36	0.28	−0.03	2.34	−0.33	−0.44	−0.78	−0.31
H3	0.53	0.56	0.74	0.07	2.11	0.38	0.91	−0.14	0.17
H5	0.29	0.10	0.41	0.09	5.06	0.40	0.28	0.05	0.32
H6	−0.16	0.07	−0.23	−0.03	2.01	−0.11	−0.14	0.13	0.15
H6	−0.17	0.06	−0.37	0.08	2.19	−0.13	−0.32	0.14	0.03
H8	−0.10	−0.11	−0.22	−0.06	1.81	−0.20	−0.30	−0.19	−0.31
H8	−0.06	−0.01	−0.07	−0.08	1.32	−0.23	0.10	−0.12	−0.32
H9	0.59	−0.01	0.55	−0.04	2.52	−0.41	0.08	−0.53	−0.53
MAD	0.17	0.12	0.22	0.06		0.22	0.25	0.20	0.22
DP4 (combined, %)				100					
DP4 (13C only, %)		1		99					
DP4 (1H only, %)		0.2		99.8					

Abbreviation: MAD, mean absolute deviation.
MAD values and DP4 probabilities are also included.

match for each isomer with the experimental data for the other, that is, computed shifts for **8D** were compared with the experimental shifts for **9** and computed shifts for **9D** were compared with the experimental shifts for **8**. The MADs for both ^{1}H and ^{13}C were much larger for the incorrect isomers (compare values in Table 3 with those in Tables 1 and 2).

Diterpene 11 (tsukubadiene)

We examined compound **11** because it is larger and more flexible than compounds **8** and **9**, and it contains an uncommon nine-membered ring, thereby providing a more challenging test of our methods. Compound **11** also contains four stereogenic centers, leading again to eight diastereomers (Scheme 4). Due to its somewhat flexible

Table 3 Differences between computed chemical shifts for 8D and experimental shifts for 9 (left) and between computed chemical shifts for 9D and experimental shifts for 8 (right)

	Compound 9	8D			Compound 8	9D	
	Exp. δ (p.p.m.)	Calc. δ (p.p.m.)	abs(Δ)		Exp. δ (p.p.m.)	Calc. δ (p.p.m.)	abs(Δ)
C4	142.8	156.13	13.3	C2	154.1	146.02	8.08
C5	121.8	126.94	5.1	C1	127.4	122.60	4.80
C3	66.7	57.74	9.0	C3	57.7	66.33	8.63
C7	52.9	58.62	5.7	C7	55.3	55.97	0.67
C2	50.1	53.66	3.6	C11	50.8	52.31	1.51
C8	48.9	49.61	0.7	C9	47.9	48.93	1.03
C1	48.8	46.76	2.0	C8	47.8	47.00	0.80
C6	47.5	46.16	1.3	C10	47.7	46.61	1.09
C10	47.4	45.45	2.0	C4	43.7	44.46	0.76
C9	44.3	40.60	3.7	C6	41.2	45.04	3.84
C11	41.6	35.29	6.3	C5	35.1	44.45	9.35
C12 or C13	30.5	28.00	2.5	C12 or C13	30.10	28.35	1.75
C14	29.4	27.52	1.9	C14	30.1	27.36	2.74
C12 or C13	29.2	26.11	3.1	C12 or C13	28.10	26.87	1.23
C15	15.9	18.31	2.4	C15	20.1	16.21	3.89
		MAD	4.2			MAD	3.35
H5	5.06	5.04	0.02	H1	5.01	5.15	0.14
H9	2.52	3.28	0.76	H9	3.16	2.48	0.68
H2	2.34	1.75	0.59	H8	1.82	2.31	0.49
H6	2.19	1.79	0.40	H10	1.76	2.27	0.51
H3	2.11	1.78	0.33	H3	1.69	2.18	0.49
H6	2.01	1.62	0.39	H5	1.67	1.98	0.31
H8	1.81	1.50	0.31	H6	1.5	1.75	0.25
H15	1.68	1.59	0.09	H6	1.5	1.70	0.20
H10	1.64	0.88	0.76	H8	1.22	1.58	0.36
H1	1.58	1.28	0.30	H5	1.2	1.45	0.25
H8	1.32	1.05	0.27	H12 or H13	1.08	1.24	0.16
H1	1.3	1.02	0.28	H14	1.06	1.34	0.28
H10	1.27	1.02	0.25	H15	1.03	1.24	0.21
H14	1.17	1.03	0.14	H12 or H13	1.02	1.06	0.04
H12 or H13	1.06	1.28	0.22	H10	0.84	1.02	0.18
H12 or H13	0.92	1.59		H4	N/A	0.90	
		MAD	0.34			MAD	0.30

Abbreviation: MAD, mean absolute deviation.
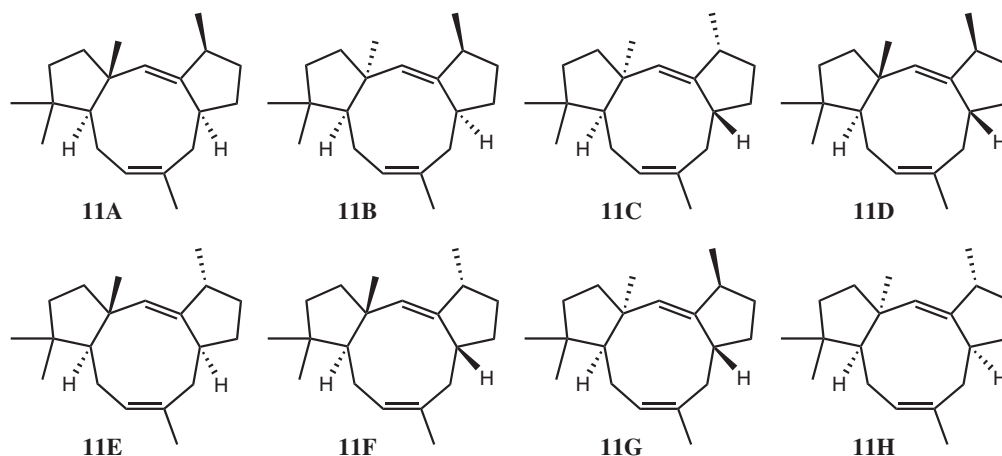
Scheme 4 Eight diastereomers of compound **11**. Note that, although we do not address herein absolute configuration, the configuration of C9 is shown as identical for all diastereomers to facilitate comparisons.

Table 4 Differences between computed and experimental (center column) chemical shifts for each atom in each of the diastereomers of compound 11 (computed minus experimental)

	11A	11B	11C	11D	Exp. δ (p.p.m.)	11E	11F	11G	11H
C1	1.71	1.76	1.32	2.85	41.7	3.14	3.09	3.31	0.72
C10	2.48	4.39	5.26	2.92	40.8	4.58	3.07	4.20	3.02
C11	−0.85	−0.78	−0.81	−2.06	39.9	−0.57	−1.73	−1.47	−1.09
C12	−1.72	2.34	0.41	1.97	40.7	−0.09	2.04	3.47	1.40
C13	4.24	2.44	2.68	1.98	48.1	2.79	2.02	2.77	4.25
C14	−0.49	−1.72	3.98	0.66	132.4	0.77	2.24	0.42	−6.00
C15	5.55	10.24	−4.24	6.10	143.7	−0.80	7.17	6.91	10.54
C16	−1.86	−1.69	−0.52	−4.51	21.7	−4.24	−1.95	−3.23	−3.76
C17	−0.67	−0.42	1.37	0.48	25.3	−0.43	0.38	1.36	−1.00
C18 or C19	−0.49	−2.72	−5.76	−0.59	25.7	−2.21	−0.53	−1.87	0.05
C18 or C19	−1.86	−4.21	−5.71	−2.79	32.1	−3.30	−2.69	−3.71	−3.05
C2	0.43	0.58	−0.34	2.96	31.9	2.66	0.57	1.23	1.72
C20	−2.91	0.63	9.00	−8.82	25.4	−0.64	−8.36	4.86	−0.93
C3	0.06	0.00	−3.64	−2.38	37	−0.16	−1.10	−0.65	−1.96
C4	3.26	1.42	3.17	2.10	40.9	1.05	1.71	4.10	−0.29
C5	1.26	−0.25	−2.63	−3.73	41.3	-3.93	−2.68	0.25	0.29
C6	5.90	−0.45	0.02	0.96	137.5	2.99	0.56	1.07	0.44
C7	3.18	4.11	1.26	3.55	124.7	−0.52	3.88	1.15	3.97
C8	0.74	4.39	4.18	1.17	24	1.07	1.27	2.58	4.77
C9	2.66	12.43	2.94	3.94	53.3	-2.39	5.69	4.91	9.64
MAD	2.12	2.85	2.96	2.83		1.92	2.64	2.68	2.94
H1	0.19	0.25	0.21	0.06	2.32	0.18	0.14	0.21	0.23
H11	0.01	0.02	−0.07	0.08	1.48	−0.09	0.08	−0.05	0.09
H11	0.01	0.08	−0.04	0.03	1.37	−0.03	0.07	−0.01	0.10
H12	−0.03	0.08	0.11	0.06	1.47	−0.02	0.02	−0.03	0.00
H14	0.08	0.23	0.16	0.09	5.11	−0.06	0.31	0.25	0.06
H16	0.02	0.11	−0.05	−0.14	1.05	−0.10	−0.08	0.00	−0.06
H17	0.13	0.12	0.00	0.04	1.64	−0.09	0.06	−0.01	0.02
H18 or H19	0.03	−0.18	−0.06	−0.07	0.93	−0.03	−0.03	−0.14	−0.16
H18 or H19	−0.02	−0.04	−0.05	−0.07	1.00	0.00	−0.06	−0.07	−0.04
H2	−0.02	0.05	0.02	0.09	1.7	0.01	−0.05	0.02	0.10
H2	0.10	0.21	0.21	−0.11	1.13	−0.08	0.21	−0.02	−0.07
H20	0.01	0.23	0.13	0.21	0.96	0.01	0.21	0.12	0.19
H3	0.02	−0.02	−0.18	0.01	1.58	−0.07	0.10	−0.32	−0.23
H3	−0.10	−0.06	−0.10	0.24	1.78	0.29	0.10	0.07	0.19
H4	−0.23	−0.27	0.38	0.09	2.96	0.08	−0.01	0.37	−0.25
H5	−0.31	−0.57	−0.41	−0.13	2.01	−0.30	0.09	0.11	−0.44
H5	0.09	0.31	0.24	0.61	2.25	0.68	0.20	0.06	0.40
H7	0.12	−0.14	−0.11	−0.12	5.49	−0.06	−0.11	−0.10	−0.11
H8	−0.02	−0.07	0.39	0.00	1.9	0.03	0.05	0.30	0.00
H9	−0.26	−0.93	−0.65	−1.05	2.34	0.13	−0.92	−0.98	−0.89
MAD	0.09	0.20	0.18	0.16		0.12	0.15	0.16	0.18
DP4 (combined, %)	100								
DP4 (13C only, %)	98.4					1.5			
DP4 (1H only, %)	97					0.1	2.9		

Abbreviation: MAD, mean absolute deviation.
MAD values and DP4 probabilities are also included.

nine-membered ring, conformational searching gave many possible conformations for each diastereomer (see Supplementary Information for energies of conformers). After optimization with B3LYP/6-31+G (d,p), five conformers were found for **11A** that are within 2.1 kcal mol^{-1} (free energy) of the lowest-energy conformer found (a conformer that is 2 kcal mol^{-1} higher in energy accounts for ~2% of the population). For **11B**, two conformers were found to be within 0.015 kcal mol^{-1} of the lowest, whereas the third conformer was 6.5 kcal mol^{-1} higher in energy. Diastereomer **11C** had two relevant

conformers (spanning, in this case, a 1.4 kcal mol^{-1} range), **11D** had five relevant conformers (spanning, in this case, a 2.3 kcal mol^{-1} range), **11E** had five relevant conformers (spanning, in this case, a 2.4 kcal mol^{-1} range), **11F** had 10 relevant conformers (spanning, in this case, a 1.9 kcal mol^{-1} range), **11G** had four relevant conformers (spanning, in this case, a 2.0 kcal mol^{-1} range) and **11H** had two relevant conformers (spanning, in this case, a 0.5 kcal mol^{-1} range). Due to the presence of more than one contributing conformer, a Boltzmann-weighted average was used to represent the computed
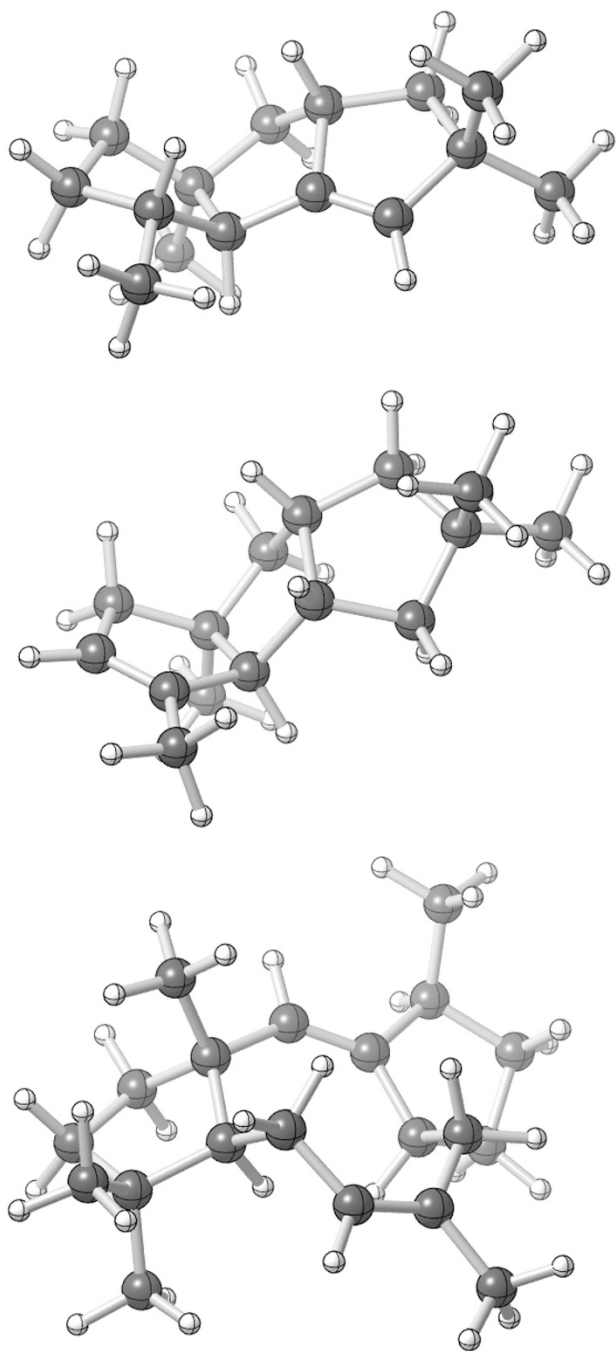
Figure 1 Computed three-dimensional structures of the lowest-energy conformers of (top to bottom) isohirsut-1-ene, isohirsut-4-ene and tsukubadiene.

chemical shifts for each diastereomer of compound **11**. Listed in Table 4 are the experimental chemical shifts for compound **11** and the differences between computed shifts for each diastereomer and the experimental values. These results indicate that **11A** is the best match to the experimental data, with MADs of 2.12 p.p.m. for ^{13}C and 0.09 p.p.m. for ^{1}H and very high DP4 probabilities. Note that diastereomer **11E** actually has a lower ^{13}C MAD, but a higher ^{1}H MAD than **11A**, but when comparisons are made on a shift-by-shift basis (DP4 analysis), ^{13}C shifts for **11E** are a much poorer match for the experimental data than are those for **11A**.

CONCLUSION

We have shown in this report that ^{1}H and ^{13}C chemical shifts computed using quantum chemistry can be used to assign relative configurations of hydrocarbon natural products with high confidence. Specifically, we disclose here the relative configurations of isohirsut-1-ene, isohirsut-4-ene and tsukubadiene, three terpenes previously isolated by Cane, Ikeda and co-workers from an engineered *Streptomyces* (three-dimensional structures shown in Figure 1).[11] Note that if the experimental chiroptical data were available, quantum chemistry could also be used to assign absolute configurations.[20] This work builds on our previous assignment of the structure of rhizathalene,[15] providing further evidence that this approach is generally useful for terpenes. It is remarkable that such simple methods can be used to distinguish one exceedingly nonpolar molecule from many slightly different nonpolar molecules with high confidence.

METHODS

Quantum chemical calculations were performed with Gaussian09.[21] All structural optimizations and frequency calculations (to verify that structures were minima) were performed with B3LYP/6-31+G(d,p) (gas phase).[22,23] ^{1}H and ^{13}C chemical shift calculations were performed using mPW1PW91/6-311+G(2d,p)[24] and the SMD implicit continuum solvation model[25] for chloroform (the solvent used experimentally). Chemical shifts were scaled as described previously[13,26] using scaling factors (slope = − 1.0936 for ^{1}H and − 1.0533 for ^{13}C; intercept = 31.8018 for ^{1}H and 186.5242 for ^{13}C) available at the cheshirenmr.info website.[27] For methyl groups, averages of the computed values of the three hydrogens were used to compare with the experimental data. DP4 statistical analysis[28] was performed to assess the probability that our assignments are correct (given the possible alternatives examined, *vide infra*). For each diastereomer of each compound, a systematic conformational search was first performed using Spartan10.[29] For those searches that resulted in only a few conformers, optimizations and NMR calculations were carried out for all of them. For those searches that resulted in many structures, we first calculated the electronic energy of each using B3LYP/6-31G(d) and then, only for the conformers that were within 15 kcal mol^{-1} of the lowest-energy conformer, we performed optimization with B3LYP/6-31+G(d,p) and NMR calculations (see Supplementary Information for details). To facilitate comparisons with the work of Cane, Ikeda and co-workers, structure numbers correspond to the numbers used in reference 11. Also, note that the chemical shifts for geminal dimethyl groups and diastereotopic CH$_2$ protons are computed separately, but these were averaged to compare with reported values (see Supplementary Information for additional details).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Tantillo, D. J. Biosynthesis via carbocations: theoretical studies on terpene formation (review). *Nat. Prod. Rep.* **28**, 1035–1053 (2011).
2 Cane, D. E. Enzymatic formation of sesquiterpenes. *Chem. Rev.* **90**, 1089–1103 (1990).
3 Yamada, Y., Cane, D. E. & Ikeda, H. Diversity and analysis of bacterial terpene synthases (review). *Methods Enzymol.* **515**, 123–162 (2012).
4 Hong, Y. J. & Tantillo, D. J. How cyclobutanes are assembled in nature – insights from quantum chemistry (review). *Chem. Soc. Rev.* **43**, 5042–5050 (2014).
5 Zerbe, P. & Bohlmann, J. Enzymes for synthetic biology of ambroxide-related diterpenoid fragrance compounds (review). *Adv. Biochem. Eng. Biotechnol.* **148**, 427–447 (2015).

6 Chen, F., Tholl, D., Bohlmann, J. & Pichersky, E. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom (review). *Plant J.* **66**, 212–229 (2011).

7 Cane, D. E. Isoprenoids including carotenoids and steroids. *Compr. Nat. Prod. Chem.* **2**, 155–200 (1999).

8 Blunt, J. W., Copp, B. R., Keyzers, R. A., Munro, M. H. G. & Prinsep, M. R. Marine natural products (review). *Nat. Prod. Rep.* **33**, 382–431 (2016).

9 Christianson, D. W. Unearthing the roots of terpenome. *Curr. Opin. Chem. Biol.* **12**, 141–150 (2008).

10 Davis, E. M. & Croteau, R. Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. *Top. Cur. Chem.* **209**, 53–95 (2000).

11 Yamada, Y. *et al.* Novel terepenes generated by heterologous expression of bacterial terpene synthase genes in an engineered *Streptomyces* host. *J. Antibiot.* **68**, 385–394 (2015).

12 Nguyen, Q. N. N. & Tantillo, D. J. The many roles of quantum chemical predictions in synthetic organic chemistry (a focus review). *Chem. Asian J.* **9**, 674–680 (2014).

13 Lodewyk, M. W., Siebert, M. R. & Tantillo, D. J. Computational prediction of 1H and 13C chemical shifts: a useful tool for natural product, mechanistic and synthetic organic chemistry. *Chem. Rev.* **112**, 1839–1862 (2012).

14 Willoughby, P. H., Jansma, M. J. & Hoye, T. R. A Guide to Small-Molecule Structure Assignment through Computation of (1H and 13C) NMR Chemical Shifts. *Nat. Protoc.* **9**, 643–660 (2014).

15 Vaughan, M. M. *et al.* Formation of the unusual semivolatile diterpene rhizathalene by the *Arabidopsis* Class I terpene synthase TPS08 in the root stele is involved in defense against belowground herbivory. *Plant Cell* **25**, 1108–1125 (2013).

16 Hong, Y. J. & Tantillo, D. J. How many secondary carbocations are involved in the biosynthesis of avermitilol? *Org. Lett.* **13**, 1294–1297 (2011).

17 Nguyen, Q. N. N. & Tantillo, D. J. Caryolene-forming carbocation rearrangements. *Beilstein J. Org. Chem.* **9**, 323–331 (2013).

18 Lodewyk, M. W. *et al.* The correct structure of aquatolide – experimental validation of a theoretically-predicted structural revision. *J. Am. Chem. Soc.* **134**, 18550–18553 (2012).

19 Quasdorf, K. W., Huters, A. D., Lodewyk, M. W., Tantillo, D. J. & Garg, N. K. Total synthesis of oxidized Welwitindolinones and (-)-N-methylwelwitindolinone C isonitrile. *J. Am, Chem. Soc.* **134**, 1396–1399 (2012).

20 Stephens, P. J., McCann, D. M., Devlin, F. J. & Smith, A. B. Determination of the absolute configurations of natural products via density functional theory calculations of optical rotation, electronic circular dichroism, and vibrational circular dichroism: the cytotoxic suberosanone, and suberosenol A acetate. *J. Nat. Prod.* **69**, 1055–1064 (2006).

21 Gaussian 09, Revision D. 01 (Gaussian, Inc.: Wallingford, CT, USA, 2004).

22 Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).

23 Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).

24 Adamo, C. & Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: the mPW and mPW1PW models. *J. Chem. Phys.* **108**, 664–675 (1998).

25 Marenich, A. V., Cramer, C. J. & Truhlar, D. G. Universal solvation model based on solute electron density and a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **113**, 6378–6396 (2009).

26 Jain, R., Bally, T. & Rablen, P. R. Calculating accurate proton chemical shifts of organic molecules with density functional methods and modest basis sets. *J. Org. Chem.* **74**, 4017–4023 (2009).

27 CHESHIRE. CHEmical SHift REpository with Coupling Constants Added Too. Available at http://cheshirenmr.info. Accessed on 23 March 2016.

28 Smith, S. G. & Goodman, J. M. Assigning stereochemistry to single diastereoisomers by GIAO NMR calculation: the DP4 probability. *J. Am. Chem. Soc.* **132**, 12946–12959 (2010).

29 Spartan'10 (Wavefunction, Ind.: Irvine, CA, USA 2010).

Supplementary Information accompanies the paper on The Journal of Antibiotics website (http://www.nature.com/ja)