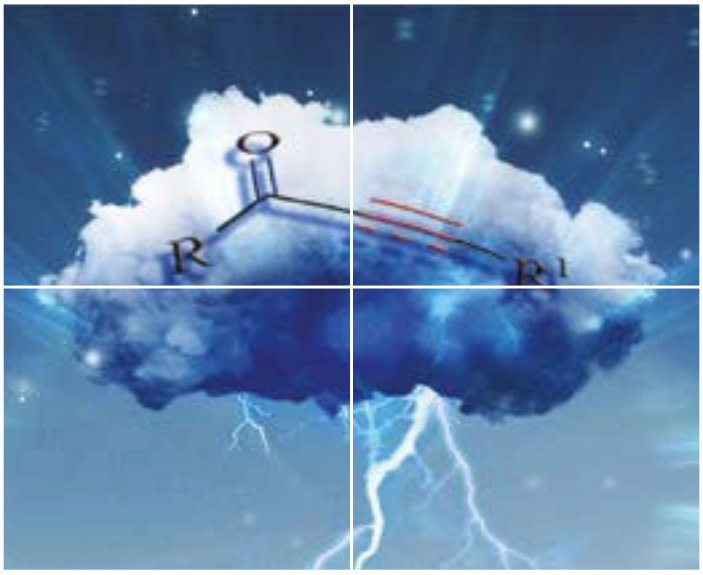
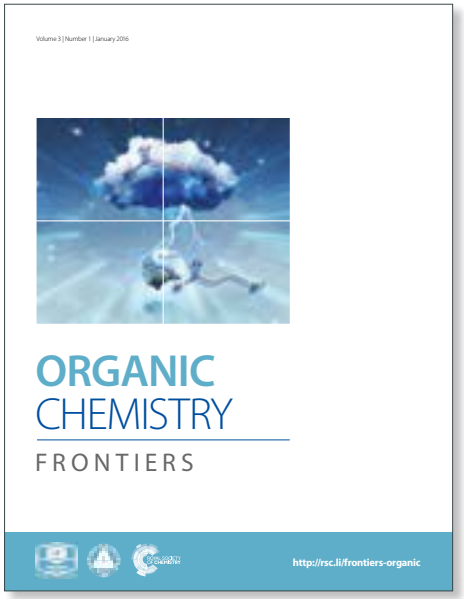
ORGANIC CHEMISTRY
FRONTIERS



Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: H. Xu, K. Hu, X. Shi, J. Tang, X. Li, H. Sun and P. Puno, *Org. Chem. Front.*, 2019, DOI: 10.1039/C9QO00281B.



View Article Online
DOI: 10.1039/C9QO00281B

ARTICLE

Synergistic Use of NMR Computation and Quantitative Interproton Distance Analysis in Structural Determination of Neokadcoccitane A, a Rearranged Triterpenoid Featuring an Aromatic Ring D from *Kadsura coccinea*

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Hou-Chao Xu,‡[a,b] Kun Hu,‡[a] Xiao-Huo Shi,[a] Jian-Wei Tang,[a] Xiao-Nian Li,[a] Han-Dong Sun,[a] and Pema-Tenzin Puno*[a]

Neokadcoccitane A (**1**), an unprecedented 14(13→12):28(14→8)-diabeo-3,4-secolanostane triterpenoid possessing an aromatic ring D and an unusual Me-21 orientation, along with two novel 14(13→12)-abeo-3,4-seco-norlanostane triterpenoids featuring three and one carbon degradation, respectively (**2** and **3**), were isolated from the roots of *Kadsura coccinea*. The absolute configuration of **1** was determined by synergistic use of NMR computation, quantitative interproton distance analysis, and TDDFT ECD calculation, while that of **2** was established by X-ray diffraction. **1** was found to exhibit moderate anti-platelet aggregation activity.

Introduction

The Schisandraceae family, comprising only genera *Kadsura* and *Schisandra*, is an invaluable source of herbal medicines. Extensive research on species within this family has proven them to be versatile manufacturers of diverse and bioactive triterpenoids, which mainly falls into three classes: lanostanes, cycloartanes, as well as schinortriterpenoids (SNTs) exclusive to this family[1]. *K. coccinea*, or *hei lao hu* (meaning "black tiger") in Chinese, is a traditional Chinese herbal medicine often used to treat rheumatic arthritis, postpartum abdominalgia with blood stasis, etc[2]. Previous investigations on the stems of this species by our lab have afforded six kinds of novel triterpene scaffolds[3] (Table S1), which were uniformly biosynthetically derived from lanostanes. In the present study, roots of this plant cultivated in Hunan Province, China, were fortunately obtained and brought into investigation to hunt for more structurally interesting and bioactive triterpenoids. Consequently, three novel lanostane-type triterpenoids, neokadcoccitanes A–C (**1**–**3**) were obtained, and they were proposed to share the common biosynthetic precursor, seco-neokadsuranic acid A (sna A)[4] (Figure 1). Sna A possessed rare 14(13→12)-abeo-3,4-secolanostane scaffold[3c], yet it was found to be a major constituent in the present material. On the basis of sna A, **1** was proposed to form through 28(14→8)-migration and subsequent ring D aromatization, resulting in its unusual 14(13→12):28(14→8)-diabeo-3,4-secolanostane skeleton featuring a tetrasubstituted benzene ring D. To the best of our knowledge, only three lanostanes, daedaleanic acid A, daedaleaside A[5] and 19-nor-lanosta-5(10),6,8,24-tetraene-1α,3β,12β,22*S*-tetraol[6], possess benzene rings within their intrinsic skeletons (all in ring B)[7] (Table S2). Besides, two SNTs with highly substituted aromatic ring D, rubriflordilactones A and B, have also been reported by our group[8], which aroused great interests in their synthetic study[9]. In addition, upon various degrees of carbon degradation occurring on sna A, **2** and **3** were generated and represented the first examples of 25,26,27-trinor-14(13→12)-abeo-3,4-secolanostane and 30-nor-14(13→12)-abeo-3,4-secolanostane triterpenoids, respectively. Notably, to address the challenges in stereochemical determination of **1**, quantum chemical calculation of NMR parameters (qccNMR)[10] and quantitative interproton distance analysis (QIDA)11 were collaboratively

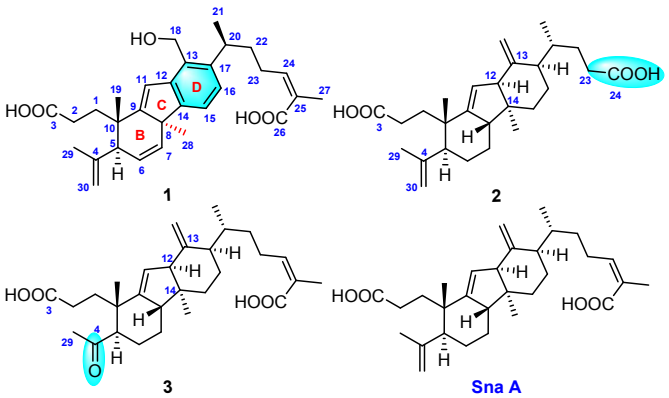


Figure 1. Chemical structures of compounds **1**–**3**, and sna A.

[a] *State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, and Yunnan Key Laboratory of Natural Medicinal Chemistry, Kunming 650201, Yunnan, People's Republic of China. E-mail: punopematenzin@mail.kib.ac.cn*
[b] *University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China.*

† Electronic Supplementary Information (ESI) available: NMR data and 2D NMR correlations of **1**–**3**; General experimental procedures; Methods, relevant data and results for qccNMR and QIDA of **1**; The methods and results for bioactivity screening; NMR, MS, IR, ECD and UV spectra of **1**–**3**. See DOI: 10.1039/x0xx00000x
‡ These authors contributed equally to the work.

View Article Online
DOI: 10.1039/C9QO00281B

employed. The qccNMR method has been frequently used in our research on configurational determination[12], structural validation[13] and revision[14]. Moreover, QIDA significantly aided in our recent investigation into structurally rigid diterpenoids from *Isodon* species[12b], however, its application in flexible molecules as **1** was rarely reported, though it has been affirmed to be practically feasible[15]. Herein, the structure elucidation and bioactivity evaluation of **1–3** were elaborated, and the employment of qccNMR/QIDA method was highlighted.

Results and discussion

Neokadcoccitane A (**1**) possessed a molecular formula of $C_{30}H_{38}O_5$ as determined by the HRESIMS ion at *m/z* 477.2655 ([M–H]$^-$, *calcd* 477.2646). Detailed analysis of ^{1}H, ^{13}C, DEPT, and HSQC spectra enabled all proton resonances of **1** to be assigned to their respective carbons (Table 1), and suggested it to be a triterpenoid with tricyclic skeleton. Careful analysis of the ^{1}H-^{1}H COSY and HMBC spectra of **1** can confirm the demolition of ring A through C-3/C-4 cleavage and the typical C-20–C-27 side chain as in sna A (Figure 2). In addition, the ^{1}H-^{1}H COSY correlations of H-5 (δ_H 2.89)/H-6 (δ_H 5.42)/H-7 (δ_H 6.09), in combination with the HMBC correlations from H_3-19 (δ_H 1.21) to C-5 (δ_C 60.2), C-9 (δ_C 158.2) and C-10 (δ_C 41.9), from H-11 (δ_H 6.68) to C-9, C-12 (δ_C 141.4) and C-14 (δ_C 150.3), and from H_3-28 (δ_H 1.38) to C-7 (δ_C 130.8), C-8 (δ_C 52.8), C-9 and C-14 revealed the 6/5 ring system in ring B and C, as well as an unprecedented migration of Me-28 from C-14 to C-8. Furthermore, the ^{1}H-^{1}H COSY correlation of H-15 (δ_H 7.24)/H-16 (δ_H 7.08), along with the HMBC correlations from H-15 to C-12 and C-17 (δ_C 144.6), from H-16 to C-13 (δ_C 129.7) and C-14, and from H_2-18 (δ_H 4.82) to C-12, C-13 and C-17, demonstrated that ring D was a tetra-substituted benzene ring. Thus, the planar structure of **1** could be determined as depicted in Figure 2.

Table 1. NMR data of **1** in $CDCl_3$ (δ in ppm, *J* in Hz)

NO.	δ_H[a]	δ_C[b]	No.	δ_H[a]	δ_C[b]
1a	2.52, overlap	35.5	17		144.6
1b	1.77, m		18	4.82, q (11.8)	59.2
2a	2.43, overlap		19	1.21, s	22.3
2b	2.22, m	30.3	20	3.19, dd (14.0, 7.1)	34.3
3		178.8	21	1.27, d (6.9)	22.9
4		145.1	22a	1.75, overlap	38.1
5	2.89, d (4.8)	60.2	22b	1.71, m	
6	5.42, dd (9.7, 4.8)	127.7	23a	2.47, overlap	28.6
7	6.09, d (9.7)	130.8	23b	2.42, overlap	
8		52.8	24	6.02, m	146.1
9		158.2	25		126.4
10		41.9	26		172.3
11	6.68, s	125.5	27	1.83, d (0.9)	20.6
12		141.4	28	1.38, s	27.7
13		129.7	29	1.37, s	22.3
14		150.3	30a	4.71, s	114.9
15	7.24, overlap	121.4	30b	4.34, s	
16	7.08, d (7.8)	122.9			

[a]Recorded at 800 MHz, [b]Recorded at 200 MHz.

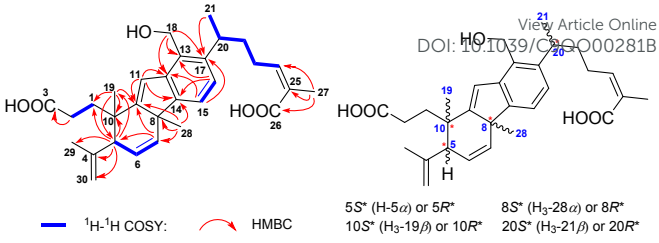


Figure 2. Left: ^{1}H-^{1}H COSY and key HMBC correlations of **1**; Right: four unsettled chiral centers of **1**, and the 8 possible diastereoisomers of **1**: (*S**, *S**, *S**, *R**)-**1a**, (*S**, *S**, *R**, *R**)-**1b**, (*S**, *R**, *R**, *S**)-**1c**, (*S**, *R**, *S**, *R**)-**1d**, (*S**, *S**, *S**, *S**)-**1e**, (*S**, *R**, *R**, *R**)-**1f**, (*S**, *R**, *S**, *S**)-**1g**, and (*S**, *S**, *R**, *S**)-**1h**.

As for the configuration determination of **1**, the H-24/H_3-27 NOESY correlation indicated the *Z*-geometry of the C-24/C-25 double bond. However, owing to the separated chiral center (C-20), weak and confusing NOESY correlations (for example, H-5/H_3-19), as well as the ^{1}H signal overlap (for example, H_3-28 and H_3-29) existing in **1**, the stereochemistry of C-5, C-8, C-10 and C-20 (Figure 2) can't be clearly resolved presently. Additionally, efforts towards the acquisition of suitable crystals of **1** for crystallographic experiment were conducted in various solvents, but finally failed. Thus, qccNMR method was adopted to establish the relative configurations of these four chiral centers. Tentatively assigning the biosynthetically conservative C-5 as *S**, all 8 possible diastereoisomers of **1** (**1a–1e**, Figures 2 and S6) were generated and subjected to a strict conformational screening process to alleviate the weakness of small-molecule force fields in low energy conformer predictions[16]. Subsequently, NMR chemical shifts calculation were run at MPW1PW91-SCRF/6-31+G(d,p)//B3LYP-D3(BJ)/6-31G(d) level of theory with IEFPCM solvent model, the optimal method for DP4+ analysis[17]. Analysis of the results was undertaken on correlation coefficient (R^2), mean absolute error (MAE), corrected mean absolute error (CMAE)[18], and DP4+ probability of each candidate, which can confidently exclude the possibilities of **1b**, **1d**, **1g**, and **1h** with opposite C-8, C-10 configurations. However, the parameters for the other four stereoisomers, though overall supported **1e** most, appeared to be relatively inconsistent, for example, when making DP4+ analysis with only ^{13}C or ^{1}H data (Figure 3, Tables S8 and S9).

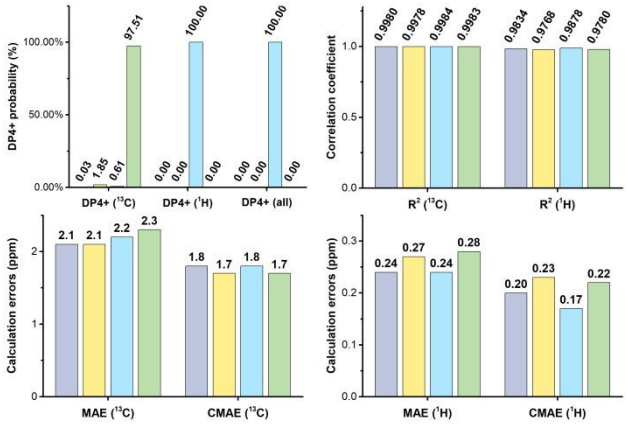


Figure 3. Analysis of the qccNMR results of **1a** (purple), **1c** (yellow), **1e** (cyan), and **1f** (green).

View Article Online
DOI: 10.1039/C9QO00281B

Allowing for the imperfect results of the qccNMR method, as well as the unusual C-20 configuration in **1e**, the QIDA method was chosen to validate the proposed structure. To provide more accurate conformers for QIDA, all B3LYP-D3(BJ)/6-31G(d) geometries were reoptimized at M06-2X-D3/def2-SVP level, and energy of optimized conformers were evaluated at M06-2X-D3/def2-TZVP level. Solvation free energies were calculated at M05-2X/6-31G* level with SMD solvation model[19]. To determine the feasibility of QIDA in current research, ensemble averaged interproton distances between 9 selected proton-pairs were calculated for **1a**, **1c**, and **1f**, respectively, and compared to those of **1e**, fortunately, the differences turned out to be distinguishable enough (Figure S21). Subsequently, NOE intensities between selected proton-pairs of **1** were extracted from 1D NOESY spectra (Figures S23–S25) and converted into interproton distances (Table S60) according to the peak amplitude normalization for improved cross-relaxation (PANIC) method[11b, 20], which were taken into comparison with their theoretically predicted counterparts. As a result, the calculated interproton distances of **1e** were found to match the experimental ones best (Table 2 and Figure S22), which further proved its correctness. Noteworthily, the QIDA of H-20/H_2-18, H-16/H_3-21, H-5/H_3-19, H-5/H-1b distances strongly supported the *S** configurations of C-5 and C-20 (Figure 4), as well as clarified the confusing H-5/H_3-19 NOESY correlation.

Table 2. The QIDA results of **1a**, **1c**, **1e**, and **1f**.

Proton-pairs	Exptl. distances (Å)	Calcd. distances (Å)			
		1a	**1c**	**1e**	**1f**
H-6/H-7	2.43	2.44	2.42	2.43	2.42
H-5/H-6	2.45	2.67	2.47	2.63	2.49
H-5/H_3-19	3.39	3.57	2.47	3.48	2.49
H-5/H-1b	2.44	2.35	3.45	2.45	3.48
H-5/30b	2.50	2.78	2.42	2.63	2.53
H-7/H_3-28	3.22	3.23	3.15	3.16	3.19
H-20/H_2-18	2.32	3.59	4.21	2.34	2.34
H-16/H_3-21	2.90	3.80	3.87	2.85	2.87
H-7/H-15	2.81	2.96	2.94	2.92	2.89
MAD		0.35	0.57	0.07	0.24

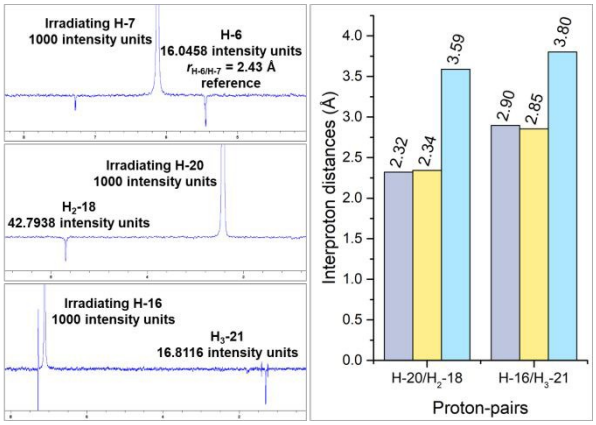


Figure 4. Left: 1D NOESY spectra of **1** when irradiating H-7, H-20, and H-16, respectively; Right: a comparison between the experimental H-20/H_3-18, and H-16/H_3-21 interproton distances (purple) and their calculated counterparts in **1e** (yellow) and **1a** (cyan).

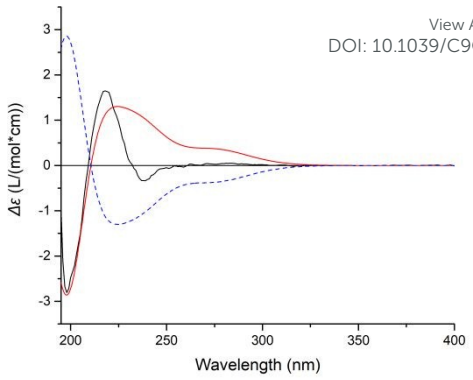


Figure 5. Experimental ECD spectrum of **1** (black); Calculated ECD spectra of (5*S*, 8*S*, 10*S*, 20*S*)-**1e** (shift = 8 nm, red) and (5*R*, 8*R*, 10*R*, 20*R*)-**1e** (shift = 8 nm, blue dash).

Finally, TDDFT ECD calculation[21] on one of the two possible enantiomers of **1e** at CAM-B3LYP/TZVP level in methanol with IEFPCM solvent model succeeded in establishing the absolute configuration of **1** as 5*S*, 8*S*, 10*S*, 20*S*, and 24*Z* (Figure 5). And this conclusion can be further reinforced by considering that C-5 and C-10 configurations are rather conserved in lanostane triterpenoids.

Neokadcoccitane B (**2**) was obtained as colorless acicular crystals with the molecular formula $C_{27}H_{40}O_4$ determined by HRESIMS ([M+Na]$^+$ *m/z* 451.2788, calcd 451.2819). Detailed analysis of the NMR data of **2** uncovered it to be closely analogous to sna A regarding their scaffolds. However, the H_3-21/H-20/H-22/H-23 spin system observed in ^{1}H-^{1}H COSY spectrum of **2**, together with the HMBC correlations from H_2-22 and H-23 to C-24 (δ_C 177.2) (Figure S2) indicated C-25–C-27 on the side chain were degraded. The excessively overlapped ROESY signals severely hampered the determination of the relative configuration of **2**. Fortunately, crystals of **2** were obtained and subjected to X-ray diffraction analysis through Cu Kα radiation (CCDC:1887346) (Flack parameter = 0.19(5)) (Figure 6), together with biosynthetic considerations, which verified the proposed planar structure of **2** and determined its absolute configuration to be 5*S*, 8*S*, 10*S*, 13*R*, 14*S*, 17*R*, and 20*R*.

Neokadcoccitane C (**3**) has the molecular formula of $C_{29}H_{42}O_5$ as supported by HRESIMS ([M−H]$^-$ *m/z* 469.2949, calcd 469.2959). Comprehensive analysis of the 1D and 2D NMR data revealed it to be also the derivative of sna A. As compared with **2**, **3** has an intact side chain but underwent C-30 degradation, which could be deduced from the HMBC correlations from H-5

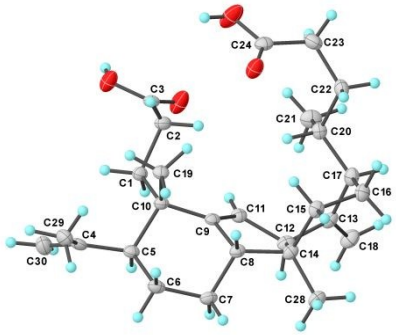


Figure 6. X-ray crystallographic structure of **2**.

View Article Online
DOI: 10.1039/C9QO00281B

Table 3. ^{1}H and ^{13}C NMR spectroscopic data for compounds **2** and **3** in pyridine-d_5 (δ in ppm, J in Hz)

NO.	2		3		NO.	2		3	
	$\delta_H{}^a$	$\delta_C{}^b$	$\delta_H{}^a$	$\delta_C{}^b$		$\delta_H{}^a$	$\delta_C{}^b$	$\delta_H{}^a$	$\delta_C{}^b$
1a	2.25 (overlap)	34.6 t	2.32 (dd, 10.5, 3.9)	34.8 t	16a	1.89 (overlap)	25.3 t	1.79 (overlap)	25.1 t
1b	2.11 (m)		2.19 (overlap)		16b	1.62 (overlap)		1.60 (overlap)	
2a	2.79 (m)	30.7 t	2.81 (overlap)	30.5 t	17	1.96 (dd, 10.6, 4.3)	50.2 d	1.97 (m)	49.8 d
2b	2.66 (overlap)		2.72 (m)		18a	4.94 (d, 2.9)	112.9 t	4.92 (d, 2.9)	112.9 t
3		177.6 s		176.9 s	18b	4.83 (d, 2.9)		4.81 (d, 2.8)	
4		147.5 s		211.3 s	19	1.07 (3H, s)	18.0 q	1.32 (3H, s)	18.3 q
5	1.88 (overlap)	57.6 d	2.38 (dd, 12.7, 3.5)	60.9 d	20	2.21 (overlap)	32.0 d	2.02 (overlap)	32.6 d
6a	1.71 (overlap)	29.1 t	1.79 (overlap)	26.2 t	21	1.10 (3H, d, 6.6)	18.4 q	1.09 (3H, d, 6.5)	18.5 q
6b	1.50 (m)		1.68 (dd, 7.7, 3.1)		22a	2.32 (overlap)	30.2 t	1.81 (overlap)	34.5 t
7a	1.61 (overlap)	29.3 t	1.58 (overlap)	28.7 t	22b	1.37 (br s)		1.19 (br s)	
7b	1.03 (overlap)		1.03 (dd, 12.8, 3.8)		23a	2.67 (overlap)	32.8 t	2.87 (2H, m)	27.7 t
8	2.02 (m)	54.9 d	2.03 (overlap)	54.3 d	23b	2.24 (overlap)			
9		153.3 s		151.4 s	24		176.8 s	6.05 (t, 7.6)	142.4 d
10		42.0 s		41.3 s	25				129.6 s
11	5.30 (br s)	123.0 d	5.36 (s)	123.5 d	26				171.2 s
12	3.04 (s)	59.2 d	3.01 (s)	59.3 d	27			2.18 (3H, s)	22.1 q
13		150.9 s		150.7 s	28	1.01 (3H, s)	21.0 q	0.99 (3H, s)	21.2 q
14		44.4 s		44.2 s	29	1.73 (3H, s)	24.2 q	2.16 (3H, s)	32.6 q
15a	1.88 (overlap)	32.9 t	1.73 (overlap)	33.1 t	30a	4.92 (br s)	114.4 t		
15b	1.10 (overlap)		1.07 (overlap)		30b	4.78 (br s)			

[a]Recorded at 500 MHz, [b]Recorded at 125 MHz.

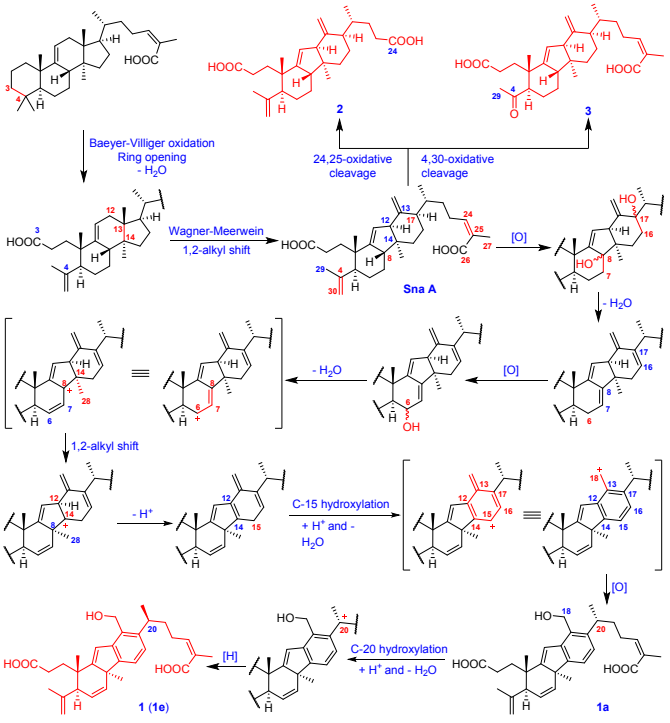
and H_3-29 to C-4 (δ_C 211.3) (Figure S3). Most of the stereochemistry of **3** can be resolved through ROESY spectrum analysis (Figure S3) and NMR data comparison with those of **2** and sna A, while that of C-5 can be determined according to the absence of both H-5/H_3-19 and H-5/H-8 ROESY correlations, as well as through biosynthetic considerations.

From the perspective of biosynthesis, **1–3** could be derived from a common precursor, sna A, which is proposed to generate mainly through Baeyer-Villiger oxidation, Wagner-Meerwein rearrangement, etc., starting from the intact lanostane. The oxidative cleavage of C-24/C-25 and C-4/C-30 double bonds upon sna A is the key process to yield **2** and **3,** respectively. As for **1**, multiple allylic oxidations and dehydrations lead to the transfer of Me-28 and aromatization of ring D (Scheme 1). Furthermore, the unusual configuration of C-20 in **1** may be explained by the tendency for benzylic oxidation (C-20), as well as that the fact that the lowest energy conformer of **1** (**1e-i**) has much lower energy (about 3 kcal/mol) than that of its C-20 epimer (**1a-i**) (Tables S15 and S38).

Compound **1–3** were screened for activities against platelet aggregation induced by colloid. The suppression ratio of **1** was 21.0 ± 9.4% (p<0.001) under the concentration of 100 μM. The anticoagulant activity and cytotoxic activity of **1–3** were also evaluated but no remarkable activities were observed, and deficiency in sample quantity hampered further evaluations.

Conclusions

In conclusion, three lanostane triterpenoids, neokadcoccitanes A–C (**1–3**), were obtained from the roots of *K. coccinea*. The discovery of these novel compounds not only enriches the skeletal diversity of the lanostane triterpenoids, but also manifested the strong ability of roots of *K. coccinea* in producing unique secondary metabolites. Moreover, together with our previous research[12b], the QIDA method was further demonstrated to be a robust tool in structure elucidation of both rigid and flexible natural products, especially when it is used in combination with the qccNMR method.



Scheme 1. Plausible biogenetic pathways for **1–3**.

View Article Online
DOI: 10.1039/C8QO00384B

Experimental Section

Plant Materials

The roots of *K. coccinea* were collected from Jingzhou Miao and Dong Autonomous County in Hunan Province, People's Republic of China, in June 2016 and identified by Prof. Heng Li at Kunming Institute of Botany. A voucher specimen (KIB 2016062101) has been deposited in State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences.

Physicochemical data of 1–3

Neokadcoccitane A (**1**): Amorphous white powder; $[\alpha]_D^{23}$ −9.7 (c 0.138, MeOH); UV (MeOH) λ_{max} (log ε) 272 (3.42) nm, 223 (3.95) nm; IR (KBr) ν_{max} 3437.93, 2927.84, 1701.80, 1637.10, 1383.62, 1262.36, 958.72, 582.62 cm^{-1}; CD (MeOH): 198 nm ($\Delta\varepsilon$ = −2.81), 218 nm ($\Delta\varepsilon$ = +1.65), 239 nm ($\Delta\varepsilon$ = −0.34); ^{1}H and ^{13}C NMR data, see Table 1 and Table S3; negative ESIMS m/z 477 [M−H]$^-$; HRESIMS *m/z* 477.2655 [M−H]$^-$ (calcd 477.2646).

Neokadcoccitane B (**2**): Colorless acicular crystals; Melting point: 208 ℃. $[\alpha]_D^{26}$ −139.0 (c 0.118, MeOH); UV (MeOH) λ_{max} (log ε) 205 (4.18) nm; IR (KBr) ν_{max} 3432.55, 2927.40, 1701.93, 1637.13, 1413.86, 1260.53, 893.39, 590.39 cm^{-1}; CD (MeOH): 211 nm ($\Delta\varepsilon$ = −3.25); ^{1}H and ^{13}C NMR data, see Table 3; positive ESIMS *m/z* 451 [M+Na]$^+$; HRESIMS *m/z* 451.2788 [M+Na]$^+$ (calcd 451.2819).

Crystal data for neokadcoccitane B (**2**): $C_{27}H_{40}O_4$, M = 428.59, a = 8.0610(2) Å, b = 10.8588(2) Å, c = 27.7356(6) Å, α = 90°, β = 90°, γ = 90°, V = 2427.78(9) Å^3, T = 100(2) K, space group $P212121$, Z = 4, μ(CuKα) = 0.605 mm^{-1}, 14256 reflections measured, 4413 independent reflections (R_{int} = 0.0418). The final R_1 values were 0.0358 ($I > 2\sigma(I)$). The final $wR(F^2)$ values were 0.0981 ($I > 2\sigma(I)$). The final R_1 values were 0.0360 (all data). The final $wR(F^2)$ values were 0.0984 (all data). The goodness of fit on F^2 was 1.085. Flack parameter = 0.19(5).

Neokadcoccitane C (**3**): Amorphous white powder, $[\alpha]_D^{26}$ −71.4 (c 0.154, MeOH), UV (MeOH) λ_{max} (log ε) 209 (4.05) nm, 293 (2.91) nm. IR (KBr) ν_{max} 3429.80, 2928.73, 1703.36, 1640.30, 1456.47, 1381.96, 1206.85, 1895.00, 575.12 cm^{-1}. ^{1}H and ^{13}C NMR data, see Table 3; positive ESIMS m/z 493 [M+Na]$^+$; HRESIMS *m/z* 469.2949 [M−H]$^-$ (calcd 469.2959). CD (MeOH): 208 nm ($\Delta\varepsilon$ = −0.96), 293 nm ($\Delta\varepsilon$ = −0.07)

Computational method

Conformational analysis of **1a–1e** was initially performed in Spartan'16 (Wavenfunction, Irvine, CA, USA, 2016) using the Monte Carlo algorithm and Merck molecular force field (MMFF). To avoid losing relevant conformations during the conformational search stage, the "set torsions" function was used to give all rotatable bonds a 6-fold sampling. Maximum 20000 conformers were examined for each diastereoisomer, and those obtained conformers within 20 kcal/mol were kept (1000 ones for each diastereoisomer).

These conformers were subjected to semiempirical geometry optimization at PM6/DH+ level using MOPAC program (Stewart Computational Chemistry, Colorado Springs) in order to obtain conformers better correlating with DFT calculations. Subsequently, PM6 geometries were clustered according to a criterion of 0.25 kcal (relative energy) and 0.25 (distance geometry). Then, clustered PM6 geometries were subjected to a DFT energy calculation at B3LYP/6-311G(d,p) level of theory, and those conformers within an energy window of 5 kcal/mol were kept. This step was completed with the aid of the Molclus program[22] (and its "isostat" module), which can call third-party programs (MOPAC, Gaussian, Openbabel, ORCA, XTB, etc) to run energy calculations or geometry optimizations on a large ensemble of conformers, as well as to undertake clustering analysis on them.

The above screened conformers were subjected to DFT geometry optimization at B3LYP/6-31G(d) level of theory with Grimme's DFT-D3 dispersion correction using Becke-Johnson damping function (EmpiricalDispersion=GD3BJ)[23] using Gaussian 09 program[24]. Frequency analysis of all optimized conformations were undertaken at the same level of theory to ensure they were true local minima on the potential energy surface. Then, energies of all optimized conformations were evaluated at B3LYP-D3(BJ)/6-311G(d,p) level of theory. Gibbs free energies of each conformers were calculated by adding "Thermal correction to Gibbs Free Energy" obtained by frequency analysis to electronic energies obtained at B3LYP-D3(BJ)/6-311G(d,p). Room-temperature (298.15 K) equilibrium populations were calculated according to Boltzmann distribution law:

$$p_i = \frac{n_i}{\sum_j n_i} = \frac{e^{-\Delta G_i/RT}}{\sum_j e^{-\Delta G_j/RT}} = \frac{Q_{i(Relative)}}{Q_{(Relative)}}$$

Where P_i is the population of the i^{th} conformer; n_i the number of molecules in i^{th} conformer; ΔG is the relative Gibbs free energy (kcal/mol); T is the temperature, usually the room temperature (298.15 K); R is the ideal gas constant (0.0019858995); Q is the partition function. Those conformers accounting for over 98% population were subjected to subsequent calculations.

NMR shielding constants were calculated with the GIAO method at mPW1PW91-SCRF/6-31+G(d,p) level with IEFPCM solvent model in chloroform solvent. The shielding constants obtained were converted into chemical shifts by referencing to TMS at 0 ppm ($\delta_{cal} = \sigma_{TMS} - \sigma_{cal}$), where the σ_{TMS} was the shielding constant of TMS calculated at the same level. For each possible candidate, the parameters a and b of the linear regression $\delta_{cal} = a\delta_{exp} + b$; the correlation coefficient, R^2; the mean absolute error (MAE) defined as $\Sigma_n |\delta_{cal} - \delta_{exp}|/n$; the corrected mean absolute error, CMAE, defined as $\Sigma_n |\delta_{corr} - \delta_{exp}|/n$, where $\delta_{corr} = (\delta_{cal} - b)/a$, were calculated[10, 18]. The DP4+ probabilities of each possible candidate were calculated with the EXCEL spreadsheet provided by Sarotti, *et al*[17].

To provide more accurate geometries for the QIDA, all B3LYP-D3(BJ)/6-31G(d) optimized conformers were reoptimized at M06-2X-D3/def2-SVP level, and energy of optimized conformers were evaluated at M06-2X-D3/def2-TZVP level ($E_{def2TZVP}$). Frequency analysis of all optimized conformations were undertaken at the same level of theory to ensure they were true local minima on the potential energy surface and to obtain the "Thermal correction to Gibbs Free Energy" ($G_{correction}$). Solvation free energies of optimized conformers

were calculated by subtracting energy calculated at M05-2X/6-31G* level in gas (E_{gas}) from energy calculated at M05-2X/6-31G* level with SMD solvation model ($E_{solvent}$) (chloroform for QIDA, and methanol for TDDFT ECD calculation described below[21]. Gibbs free energies of each conformers were calculated according to the following equation:

$$G_{solvent} = E_{def2TZVP} + G_{correction} + (E_{solvent} - E_{gas}) + 1.89\ \text{kcal/mol}$$

Room-temperature (298.15 K) equilibrium populations were calculated according to Boltzmann distribution law. Those conformers accounting for over 98% population were subjected to QIDA and TDDFT ECD calculation described below.

TDDFT ECD calculations were run at CAM-B3LYP/TZVP level of theory in MeOH with IEFPCM solvent model. For each conformer, 30 excited states were calculated. The calculated ECD curves were generated using Multiwfn 3.6 software[25].

Quantitative interproton distance analysis[11, 26]

(1) The acquisition of calculated interproton distances

r_{cal} was the calculated interproton distances (Å) extracted from DFT optimized conformers obtained with method as stated above; For molecules with multiple conformers, r_{cal} were obtained by weighting interproton distances from individual conformers according to their Boltzmann populations:

$$\boldsymbol{r}_{H_a-H_b,\ calcd.} = \left(\sum p_i \times \boldsymbol{r}_{H_a-H_b,i}^{-6}\right)^{-1/6}$$

where $r_{Ha\text{-}Hb,calcd}$ is the interproton distance of interest (between H_a and H_b) in conformer i, and *pi* is the population of conformer i.

To obtain Boltzmann averaged distance between a single proton (H_x) and three chemically equivalent methyl protons (H_{abc}), the following equation is used:

$$\boldsymbol{r}_{H_x-H_{abc},\ calc} = \left\{\frac{\sum_i\left[(\boldsymbol{r}_{H_x-H_{a,i}})^{-6} \times p_i + (\boldsymbol{r}_{H_x-H_{b,i}})^{-6} \times p_i + (\boldsymbol{r}_{H_x-H_{c,i}})^{-6} \times p_i\right]}{3}\right\}^{-1/6}$$

To obtain Boltzmann averaged distance between a single proton (H) and two overlapped methylene protons (H_{ab}), the following equation is used:

$$\boldsymbol{r}_{H_x-H_{ab},\ calcd.} = \left\{\frac{\sum_i\left[(\boldsymbol{r}_{H_x-H_{a,i}})^{-6} \times p_i + (\boldsymbol{r}_{H_x-H_{b,i}})^{-6} \times p_i\right]}{2}\right\}^{-1/6}$$

(2) The acquisition of experimental interproton distances

The 1D selective NOESY spectra was collected using the "selnogpzs" pulse program; Before the measurement, the mixing time for the 1D experiments were optimized by selective irradiation of H_3-19 (Figure S23). During a gradient of mixing time (100–1000 ms, step = 100 ms), the H_3-19/H-11 intensities were recorded. Then the PANIC NOE build-up curve (Figure S24) was plotted. According to the curve, as well as in consideration of the signal/noise ratio under different mixing time, a mixing time of 700 ms was adopted.

$r_{exptl.}$ was experimental interproton distances (Å) determined by using equation below and the above obtained intensities from the 1D NOESY spectra:

$$\boldsymbol{r}_{H_a-H_b,\ exp} = \boldsymbol{r}_{ref} \times \left(\frac{\eta_{H_a-H_b,\ exp}}{\eta_{ref}}\right)^{-1/6}$$

Where $\boldsymbol{r}$ is interproton distance, and η is NOESY signal intensity derived from the 1D NOESY spectra. Peak amplitude normalization for improved cross-relaxation (PANIC) method was employed by setting the irradiated peak in each 1D NOESY spectrum to 1000 intensity units as shown in Figure S25. The distances between methylene protons H-6 and H-7 (2.43 Å), which were insensitive to molecular conformation, was chosen to be the reference distance (r_{ref}).

To obtain the averaged experimental distance between a single proton (H_x) and three chemically equivalent methyl protons (H_{abc}):

(a) When irradiating H_x, set the irradiated peak to 1000, and then average the observed H_x-H_{abc} intensity [($\eta_{Hx\text{-}Habc,exp}$)/3], and then use this intensity to calculated the averaged $r_{Hx\text{-}Habc,exp}$;

(b) When irradiating H_{abc}, set the irradiated peak to 1000, and then use observed H_x-H_{abc} intensity ($\eta_{Hx\text{-}Habc,exp}$) directly to calculated the averaged $r_{Hx\text{-}Habc,exp}$.

To obtain the averaged experimental distance between a single proton (H_x) and two overlapped methylene protons (H_{ab}):

(a) When irradiating H_x, set the irradiated peak to 1000, and then average the observed H_x-H_{ab} intensity [($\eta_{Hx\text{-}Habc,exp}$)/2], and then use this intensity to calculated the averaged $r_{Hx\text{-}Hab,exp}$;

(b) When irradiating H_{ab}, set the irradiated peak to 1000, and then use observed H_x-H_{ab} intensity ($\eta_{Hx\text{-}Habc,exp}$) directly to calculated the averaged $r_{Hx\text{-}Habc,exp}$.

(3) The evaluation of fitness between experimental and calculated interproton distances

The mean absolute deviation (MAD) was defined as:

$$MAD = \frac{\sum_n |\boldsymbol{r}_{cal} - \boldsymbol{r}_{exp}|}{n}$$

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This project was supported financially by the National Natural Science Foundation of China (Nos. 81874298, 81673329). The authors acknowledged Dr. Siying Zhong (University of Bristol) for precious discussions on the use of the QIDA method.

Notes and references

1. (a) W. L. Xiao, R. T. Li, S. X. Huang, J. X. Pu and H. D. Sun, *Nat. Prod. Rep.*, 2008, **25**, 871-891; (b) Y. M. Shi, W. L. Xiao, J. X. Pu and H. D. Sun, *Nat. Prod. Rep.*, 2015, **32**, 367-410.
2. Nanjing University of Chinese Medicine, *Dictionary of traditional Chinese materia medica (Zhong yao da ci dian)*, Shanghai scientific and technical publishers, Shanghai, 2006. Vol. 1, p 3330-3331.
3. (a) C. Q. Liang, Y. M. Shi, R. H. Luo, X. Y. Li, Z. H. Gao, X. N. Li, L. M. Yang, S. Z. Shang, Y. Li, Y. T. Zheng, H. B. Zhang, W. L.

Xiao and H. D. Sun, *Org. Lett.*, 2012, **14**, 6362-6365; (b) C. Q. Liang, Y. M. Shi, X. Y. Li, R. H. Luo, Y. Li, Y. T. Zheng, H. B. Zhang, W. L. Xiao and H. D. Sun, *J. Nat. Prod.*, 2013, **76**, 2350-2354; (c) C. Q. Liang, Y. M. Shi, W. G. Wang, Z. X. Hu, Y. Li, Y. T. Zheng, X. N. Li, X. Du, J. X. Pu, W. L. Xiao, H. B. Zhang and H. D. Sun, *J. Nat. Prod.*, 2015, **78**, 2067-2073; (d) Z. X. Hu, Y. M. Shi, W. G. Wang, X. N. Li, X. Du, M. Liu, Y. Li, Y. B. Xue, Y. H. Zhang, J. X. Pu and H. D. Sun, *Org. Lett.*, 2015, **17**, 4616-4619.

4. L. N. Li, H. Xue, D. L. Ge, K. Kunio, M. Toshio and O. Sadafumi, *Planta Med.*, 1989, **55**, 300-302.
5. K. Yoshikawa, K. Kouso, J. Takahashi, A. Matsuda, M. Okazoe, A. Umeyama and S. Arihara, *J. Nat. Prod.*, 2005, **68**, 911-914.
6. G. Li, S. Kusari, P. Kusari, O. Kayser and M. Spiteller, *J. Nat. Prod.*, 2015, **78**, 2128-2132.
7. Dictionary of Natural Products 27.2 [Online]; CRC Press, Taylor & Francis Group, Posted 2018. http://dnp.chemnetbase.com/faces/chemical/ChemicalSearch.xhtml (accessed Feb 20th, 2019)
8. W. L. Xiao, L. M. Yang, N. B. Gong, L. Wu, R. R. Wang, J. X. Pu, X. L. Li, S. X. Huang, Y. T. Zheng, R. T. Li, Y. Lu, Q. T. Zheng and H. D. Sun, *Org. Lett.*, 2006, **8**, 991-994.
9. (a) J. Li, P. Yang, M. Yao, J. Deng and A. Li, *J. Am. Chem. Soc.*, 2014, **136**, 16477-16480; (b) S. S. Goh, G. Chaubet, B. Gockel, M. C. A. Cordonnier, H. Baars, A. W. Phillips and E. A. Anderson, *Angew. Chem., Int. Ed.*, 2015, **54**, 12618-12621; (c) P. Yang, M. Yao, J. Li, Y. Li and A. Li, *Angew. Chem., Int. Ed.*, 2016, **55**, 6964-6968; (d) G. Chaubet, S. S. Goh, M. Mohammad, B. Gockel, M. C. A. Cordonnier, H. Baars, A. W. Phillips and E. A. Anderson, *Chem. - Eur. J.*, 2017, **23**, 14080-14089; (e) X. Li, P. H. Y. Cheong and R. G. Carter, *Angew. Chem., Int. Ed.*, 2017, **56**, 1704-1718.
10. M. W. Lodewyk, M. R. Siebert and D. J. Tantillo, *Chem. Rev.*, 2012, **112**, 1839-1862.
11. (a) C. P. Butts, C. R. Jones, E. C. Towers, J. L. Flynn, L. Appleby and N. J. Barron, *Org. Biomol. Chem.*, 2011, **9**, 177-184; (b) C. P. Butts, C. R. Jones, Z. Song and T. J. Simpson, *Chem. Commun.*, 2012, **48**, 9023-9025; (c) M. G. Chini, C. R. Jones, A. Zampella, M. V. D'Auria, B. Renga, S. Fiorucci, C. P. Butts and G. Bifulco, *J. Org. Chem.*, 2012, **77**, 1489-1496.
12. (a) W. P. Ding, K. Hu, M. Liu, X. R. Li, R. Chen, X. N. Li, X. Du, P. T. Puno and H. D. Sun, *Fitoterapia*, 2018, **127**, 193-200; (b) Q. Yang, K. Hu, B. C. Yan, M. Liu, X. N. Li, H. D. Sun and P. T. Puno, *Org. Chem. Front.*, 2019, **6**, 45-53.
13. (a) Y. M. Shi, S. L. Cai, X. N. Li, M. Liu, S. Z. Shang, X. Du, W. L. Xiao, J. X. Pu and H. D. Sun, *Org. Lett.*, 2016, **18**, 100-103; (b) J. W. Tang, L. M. Kong, W. Y. Zu, K. Hu, X. N. Li, B. C. Yan, W. G. Wang, H. D. Sun, Y. Li and P. T. Puno, *Org. Lett.*, 2019, **21**, 771-775.
14. Y. M. Shi, K. Hu, G. Pescitelli, M. Liu, X. N. Li, X. Du, W. L. Xiao, H. D. Sun and P. T. Puno, *Org. Lett.*, 2018, **20**, 1500-1504.
15. J. Wu, P. Lorenzo, S. Zhong, M. Ali, C. P. Butts, E. L. Myers and V. K. Aggarwal, *Nature*, 2017, **547**, 436-440.
16. I. Y. Kanal, J. A. Keith and G. R. Hutchison, *Int. J. Quantum Chem.*, 2018, **118**, e25512.
17. N. Grimblat, M. M. Zanardi and A. M. Sarotti, *J. Org. Chem.*, 2015, **80**, 12526-12534.
18. P. H. Willoughby, M. J. Jansma and T. R. Hoye, *Nat. Protoc.*, 2014, **9**, 643-660.
19. A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378-6396.
20. S. Macur, B. T. Farmer and L. R. Brown, *J. Magn. Reson.*, 1986, **70**, 493-499.
21. G. Pescitelli and T. Bruhn, *Chirality*, 2016, **28**, 466-474.
22. Tian Lu, molclus program, Version 1.5, http://www.keinsci.com/research/molclus.html (accessed September 6th, 2018)
23. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
24. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. P. Jr., F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09*; Gaussian, Inc., Wallingford CT2010.
25. T. Lu and F. Chen, *J. Comput. Chem.*, 2011, **33**, 580-592.
26. C. R. Jones, M. D. Greenhalgh, J. R. Bame, T. J. Simpson, R. J. Cox, J. W. Marshall and C. P. Butts, *Chem. Commun.*, 2016, **52**, 2920-2923.