JOC The Journal of Organic Chemistry

Article

pubs.acs.org/joc

Quantitative NMR-Derived Interproton Distances Combined with Quantum Mechanical Calculations of ^{13}C Chemical Shifts in the Stereochemical Determination of Conicasterol F, a Nuclear Receptor Ligand from *Theonella swinhoei*

Maria Giovanna Chini,[†] Catharine R. Jones,[‡] Angela Zampella,[§] Maria Valeria D'Auria,[§] Barbara Renga,[⊥] Stefano Fiorucci,[⊥] Craig P. Butts,*[,‡] and Giuseppe Bifulco*[,†]

[†]Dipartimento di Scienze Farmaceutiche e Biomediche, Università di Salerno, via Ponte don Melillo, 84084 Fisciano (SA), Italy

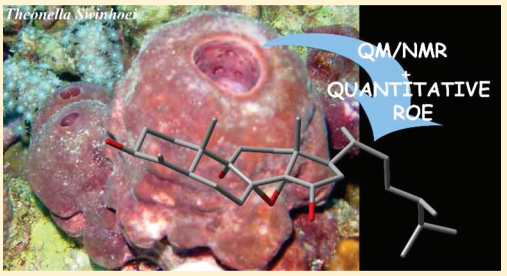
[‡]Department of Chemistry, University of Bristol, Cantocks Close, BS8 1TS Bristol, United Kingdom

[§]Dipartimento di Chimica delle Sostanze Naturali, Università di Napoli "Federico II", via D. Montesano 49, 80131 Napoli, Italy

[⊥]Dipartimento di Medicina Clinica e Sperimentale, Università di Perugia, Nuova Facoltà di Medicina e Chirurgia, Via Gerardo Dottori 1 S. Andrea delle Fratte, 06132 Perugia, Italy

Supporting Information

ABSTRACT: Here we report the first application of combined accurate ROE-distance analysis with DFT calculations of NMR chemical shifts to achieve the relative configuration assignment of a marine natural product, conicasterol F, a new polyhydroxylated steroid isolated from the marine sponge *Theonella swinhoei*. We demonstrate the substantial advantages of this combined approach as a tool for structural studies of natural products, providing a powerful alternative to, or information to underpin, total synthesis when more classical NMR data analysis fails to provide unequivocal results. In this paper, we also describe the isolation and structure elucidation of conicasterol F and its 24-ethyl derivative, theonellasterol I, and their pharmacological evaluation as human nuclear receptor modulators.



INTRODUCTION

Historically, natural products have represented a principal source of therapeutic agents, and so complete knowledge of their constitution and three-dimensional structure is necessary to design and characterize new molecular platforms with potential utility in therapy. In this context, NMR spectroscopy together with modern computational techniques represents an efficient and effective approach to stereochemical determination in solution. It is absolutely critical that the structural information obtained from NMR spectroscopy is as accurate and reliable as possible, because when it fails to provide an unequivocal answer, the traditional fallback position is the time-intensive and expensive solution of total synthesis.

In NMR spectroscopy, the nuclear Overhauser effect (NOE) is undoubtedly a powerful tool for establishing stereochemical and conformational details of chemical structures. However, because many factors may perturb NOE intensities, including spin diffusion, additional cross-relaxation pathways, selective polarization transfer, variation in τ_c between spins, accuracy of signal integration, and conformational flexibility,[1] the analyses of NOE data are generally qualitative (NOE/no NOE) or semiquantitative (strong/medium/weak) when applied to stereochemical determination of small molecules such as natural products.[1] On the other hand, recent improvements in NMR hardware (nonquadrature detection, improved RF generation, and digital receivers, etc.), NOE experimental methods (e.g., 1D-DPFGSE sequences,[2−4] zero-quantum suppression[5]), and data analysis, such as the PANIC method of Macura,[6,7] have made quantitative measurements of NOEs more reliable. On this basis, we have recently reported[8−10] that many of the 'perturbing' factors outlined above do not contribute substantially to NOE experiments conducted on small molecules, and indeed the NOE can be applied quantitatively with surprisingly high levels of accuracy of interproton distances in both rigid and flexible organic test molecules.

Specifically, the accurate NOE-distance methodology compares the relative NOE intensities (and hence relative build-up rates) for pairs of spins in transient NOESY (or ROESY) experiments when the molecule of interest is in the fast tumbling regime and the measurements are made within the initial rate approximation limits.[8] The relative intensity values are obtained from the standardization of each NOE peak intensity versus the irradiated peak in the same selective inversion experiment, because in this way any perturbation, which proportionately affects all spins in a given experiment, is minimized. In fact, Macura et al.[6] and others[7] have highlighted

Received: November 17, 2011
Published: December 27, 2011

dx.doi.org/10.1021/jo2023763 | J. Org. Chem. 2012, 77, 1489−1496

that employing relative, rather than absolute, intensities of NOEs (the so-called 'PANIC' method) from within a single experiment corrects for other forms of relaxation in 1D- or 2D-NOESY experiments and effectively extends the period during which the initial rate approximation holds for the relative values of NOE enhancements.[11]

Also recently, QM/NMR[12,13] has demonstrated success[14,15] as a tool for differentiating diastereoisomers, to understand the mechanism of action of a determinate molecule,[16] as support to the total synthesis,[17] or to determine the structure of unusual natural substances,[18] being able to identify the correct configuration in an efficient and accurate fashion. In the QM/NMR approach, DFT values of NMR parameters, i.e., chemical shifts (^{13}C and ^{1}H) and scalar coupling constants ($^{3}J_{H-H}$ and $^{2,3}J_{C-H}$), are compared with experimental values to assign constitution and configuration. In fact, in many cases if the system is treated at a sufficient level of theory, the correct molecular arrangement will predictably be the isomer characterized by the best match between calculated and measured spectroscopic properties, where the best fit is identified using statistical parameters such as mean absolute error (MAE), total absolute deviation (TAD), or others.[19] In particular, the chemical shift values are especially diagnostic in cases where there are multiple quaternary centers, which strongly limit stereochemical analysis by traditional coupling constant and/or NOE analysis.

On this basis, our aim was to combine these two different validated approaches (accurate NOE-distance analysis and QM/NMR analysis) to obtain a new efficient and robust method for the stereostructural determination of the organic compounds (Figure 1).

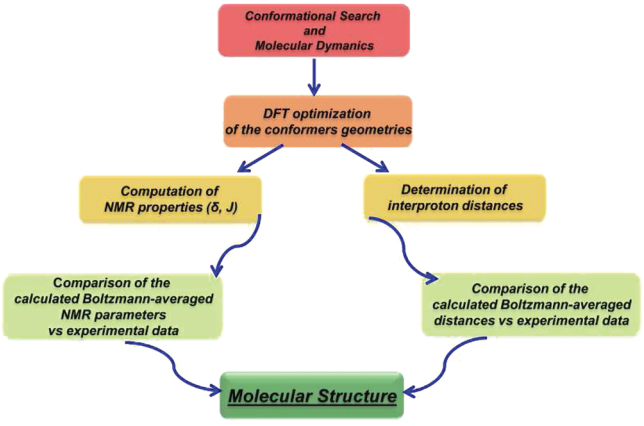


Figure 1. General protocol for combining quantitative interproton distances by NOE/ROE with quantum chemical calculations of NMR parameters in stereochemical determination.

The combination of the two methods allows the stereochemistry to be determined by two experimentally independent methods, providing greater confidence in the final structural assignment. This is especially important in cases where one technique or the other is not completely unequivocal in distinguishing all structural candidates. In this work, we report the first application of our combined method to unknown structures, and in particular the stereostructure assignments of conicasterol F (**1**) and theonellasterol I (**2**) (Figure 2), two new 4-methylene polyhydroxylated steroids isolated from the marine sponge *Theonella swinhoei*. These compounds are especially challenging for stereochemical determination (vide infra) due to the five contiguous, quaternary stereogenic centers from C8−C9−C10−C13−C14.

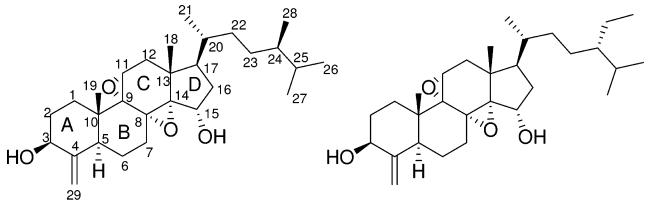


Figure 2. Conicasterol F (**1**) and theonellasterol I (**2**) from the marine sponge *Theonella swinhoei*.

Marine sponges of the genus *Theonella* have attracted great interest from the scientific community for the impressive variety of bioactive secondary metabolites with unusual structures and powerful biological activity. In particular, the recent exploration of a specimen of *Theonella swinhoei* collected at the Solomon islands led to the isolation of new cyclic peptides perthamides C−F,[20,21] and solomonamides,[22] endowed with anti-inflammatory activity. From the same sponge, two sulfated sterols, solomonsterols A and B, and new theonellasterols[23,24] and conicasterols[24] have shown to be potent ligands of human nuclear receptors with pharmacological potential in the treatment of immune-driven inflammatory bowel diseases.[25]

RESULTS AND DISCUSSION

The molecular formula of conicasterol F (**1**) was established as $C_{29}H_{46}O_4$ from HR ESIMS (m/z 465.3532, calcd 465.3556 [M + Li]$^+$) and ^{13}C NMR data (Table 1). The seven degrees of unsaturation implied by the formula were ascribed to six rings and one C=C bond (δ_C 151.6 and 104.5). The ^{1}H NMR spectrum showed signals for six Me groups [(δ_H 0.66 (s), 0.73 (s), 0.78 (d), 0.86 (d), 0.91 (d), and 0.84 (d)], two olefinic H-atoms (δ_H 4.64 and 5.30), and one O-bearing CH group (δ_H 3.67), consistent with a 3-hydroxy-4-methylene-24-methyl steroidal system. The HMBC correlations from CH_2-29 to C3, C4, and C5 and the allylic ^{1}H−^{1}H COSY correlations of CH_2-29 with H-3 and H-5 confirmed the presence of an exocyclic CH_2 group at C4. In addition to the signals arising from the 4-methylene-3β-hydroxy ring A, the ^{13}C NMR spectrum of **1** (Table 1) showed the presence of one oxygen-bearing methine carbon (δ_C 65.0) and three oxygen-bearing quaternary carbons (δ_C 79.8, 61.5, 60.8). The HMBC correlation between the angular methyl Me-18 protons and the carbon at δ_C 79.8 indicated that C-14 was an oxygen-bearing carbon, whereas the position of the secondary alcoholic function at C15 (δ_C 65.0, δ_H 4.18) was deduced by ^{1}H−^{1}H COSY spin system from H17 (δ_H 1.61) to H15 (δ_H 4.18). The HMBC correlations of H-7 at δ_H 1.86 with C-14, C-8 (δ_C 60.8), and C-9 (δ_C 61.5) suggests an epoxide ring between C14 and C-8 and the presence of an oxygen bearing carbon at C9 position which is further substantiated by an HMBC crosspeak Me-19/C-9 (δ_C 61.5).

The additional epoxy functionality at C9/C11 of ring C is suggested in the ^{1}H NMR spectrum by a resonance at δ_H 2.51 (dd, J = 7.3, 2.1 Hz) that is correlated in the HSQC spectrum with C-11 at δ_C 50.1 and which showed COSY crosspeaks exclusively with the protons at δ_H 1.67 and 1.92 assigned to H_2-12. These data thus also account for the additional degree of unsaturation established on the basis of mass data.

dx.doi.org/10.1021/jo2023763 | J. Org. Chem. 2012, 77, 1489−1496

Table 1. ^{1}H and ^{13}C NMR Data (500 and 700 MHz, C_6D_6) for Conicasterol F (1)

conicasterol F (**1**)[a]		
position	δ_H	δ_C
1α	1.34 ovl	30.3
1β	0.88 m	
2α	1.74 ovl	32.2
2β	1.27 m	
3	3.67 dd (4.6, 11.4)	72.6
4	–	151.6
5α	1.99 br d (11.6)	45.3
6α	1.38 m	22.6
6β	1.45 m	
7α	1.86 ovl	27.3
7β	2.26 ddd (1.9, 4.3, 14.3)	
8	–	60.8
9	–	61.5
10	–	39.5
11β	2.51 dd (2.1, 7.3)	50.1
12α	1.92 d (14.0)	35.7
12β	1.67 dd (7.3, 14.0)	
13	–	44.9
14	–	79.8
15β	4.18 dd (1.4, 8.8)	65.0
16α	2.07 m	40.9
16β	1.77 ovl	
17	1.61 m	50.6
18	0.66 s	14.8
19	0.73 s	18.5
20	1.21 m	35.0
21	0.78 d (6.7)	18.7
22	1.24 ovl	33.4
	1.47 ovl	
23	1.23 ovl	30.9
	1.34 ovl	
24	1.25 ovl	39.3
25	1.54 m	32.7
26	0.86 d (6.7)	18.6
27	0.91 d (6.7)	20.6
28	0.84 d (6.5)	15.8
29a	4.64 br s	104.5
29b	5.30 br s	

[a]Coupling constants are in parentheses and given in hertz. ^{1}H and ^{13}C assignments were made based on COSY, HSQC, HMBC, and ROESY experiments as described in the main text.

The stereochemical assignment of the majority of the steroid skeleton is straightforward. The absolute steroidal configuration as depicted in Figure 2 was assumed on biogenetic grounds. The coupling constants between H-3 [δ_H 3.67 (dd, J = 4.6, 11.4 Hz)] and H_2-2, and the ROESY correlations H-3/H-5α, indicated that H-3 was axial and therefore the OH-3 was β-oriented. The axial disposition of H-11 [δ_H 2.51 (dd J = 2.1, 7.3 Hz] and, consequently, the α-orientation of the C9/C11 epoxide ring, was evident from the $^3J_{HH}$ vicinal coupling to H_2-12 and ROESY correlations to axial protons Me-19, H-8, and Me-18. The α-orientation of 15-OH was similarly assigned on the basis of a ROESY correlation between H-15 and Me-18. The configuration at C-24 on the flexible side chain was determined to be the same as conicasterol by comparison of ^{1}H and ^{13}C chemical shifts.[24]

Theonellasterol I (**2**) was isolated as an optically active pale yellow oil, and a molecular formula of $C_{30}H_{48}O_4$, was established by high resolution mass spectrometry. The ^{1}H and ^{13}C NMR spectra of **2** were very similar to those of conicasterol F (Table S5 in Supporting Information), with the data for the steroidal core being essentially identical ($\Delta\delta$ <0.5 ppm). The only difference between these two molecules lies in the steroidal side chain, with an ethyl group replacing the C24 methyl in **1**. On the basis of chemical shift similarities, both the constitution and stereochemistry of the tetracyclic nucleus of theonellasterol I (**2**) was assumed to be the same as that for conicasterol F (**1**). On the other hand, the configuration at C24 was determined by comparison of ^{13}C NMR data with the epimeric steroidal side chains.[26,27]

Assignment of the Relative Configuration of 8−14 Epoxy Ring in Conicasterol F (1). The presence of four contiguous quaternary centers in ring C of **1** severely hampers the stereochemical assignment of the relative configuration of the 8−14 epoxy ring, with the lack of protons at C8 and C14 making conventional $^3J_{HH}$ and NOE analysis impossible. Therefore, this molecule was selected as a suitable case study to test our combination of quantitative ROE-derived interproton distances with QM calculation of NMR parameters. The relative rigidity, the molecular size, and the solubility in nonviscous benzene should ensure appropriate conditions for accurate interproton measurements by ROE.[8] On the other hand, because the only difference between the two possible diasteroisomers of the compound **1** is the stereochemistry of C8 and 14 (Figure 3), substantial chemical shift differences are

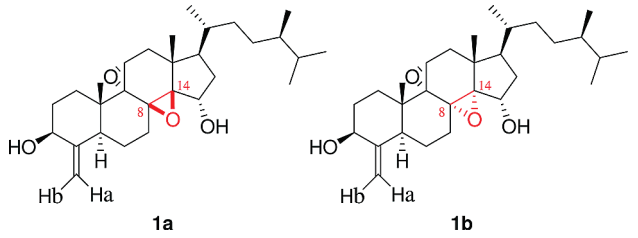


Figure 3. Molecular structure of the two possible diasteroisomers (**1a** and **1b**) of conicasterol F.

expected for the carbon atoms close to the junction between C-D rings, suggesting the potential for stereostructure validation by GIAO calculation of ^{13}C NMR chemical shifts.

Following our protocol (Figure 1), we performed molecular dynamics and Monte Carlo conformational search calculations (see Computational Details) on both possible stereoisomers **1a** and **1b** (Figure 3), using the MMFF[28] force field (MacroModel software package).[29,30] In each case, only a single conformer of the steroid ring system was found. The resulting geometries for **1a** and **1b** were optimized (Figure 4) at the DFT level using the MPW1PW91 functional and 6-31G (d)[31] basis set (Gaussian

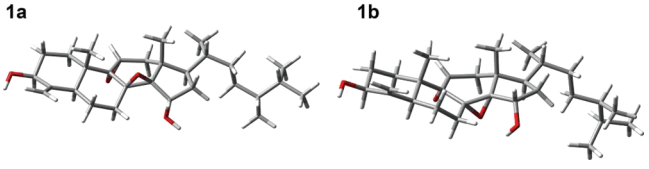


Figure 4. Optimized geometries of diastereoisomers **1a** and **1b** of conicasterol F.

dx.doi.org/10.1021/jo2023763 | J. Org. Chem. 2012, 77, 1489−1496

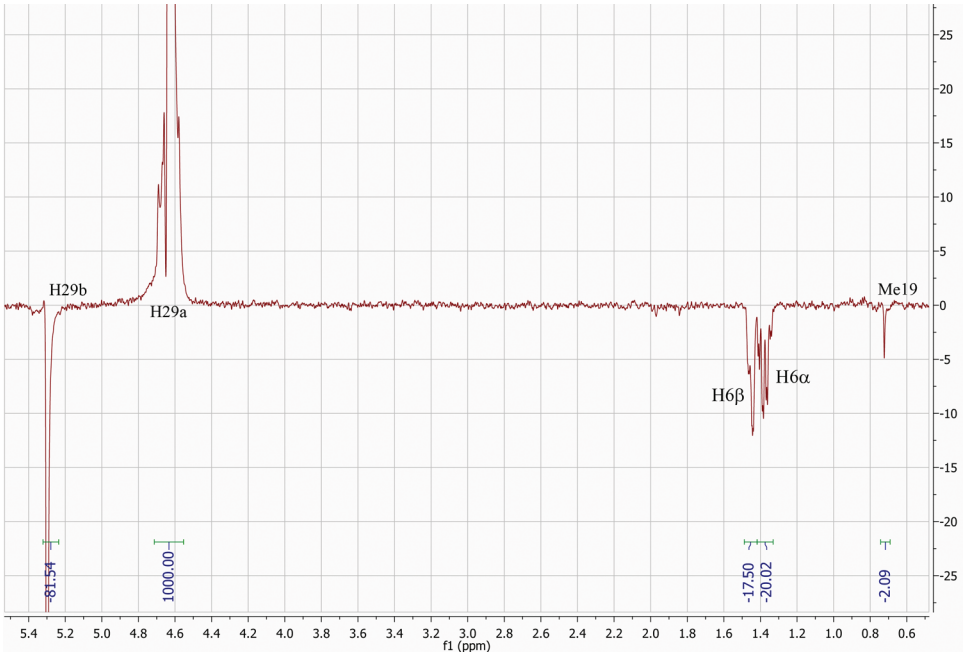


Figure 5. Selective 1D-ROESY spectrum (600 MHz, C_6D_6) of H29a of conicasterol F (**1**).

09 Software Package).[32] Calculated interproton distances were obtained directly from the DFT-optimized geometries without further corrections.

Quantitative experimental interproton distances for **1** were obtained from 1D-ROESY spectra as detailed in our previous work.[8] Figure 5 shows an example of the data obtained, with the selective 1D-ROESY spectrum of H29a, with clear ROE enhancements for H29b, H6α, H6β, and Me19 of **1**.

Selective 1D-ROESY experiments were performed only on the key protons[33] around the junction C-D (See Figure 2), namely protons H7β, H11β, H15β, and Me18. The 1D-ROESY spectra were calibrated in the PANIC fashion[6,7] by setting the integral value for the irradiated peak in each case arbitrarily to 1000[8] and thus standardizing the absolute values of the ROE intensities to this in every 1D-ROESY spectrum. Interproton distances were then calculated using eq 1

$$\frac{\eta_{IS}}{\eta_{ref}} = \frac{r_{IS}^{-6}}{r_{ref}^{-6}} \quad (1)$$

where η_{IS} is the intensity of the NOE for a given proton pair I and S, r_{IS} is the corresponding interproton distance, η_{ref}, and r_{ref} are reference values for a single chosen NOE for which the interproton distance is assumed based on geometric constraint. We chose the methylene protons H29a–H29b as the reference ROE (η_{ref} = 81.54) for the 1D-ROESY data set with the corresponding intermethylene distance of 1.85 Å as the reference distance (r_{ref}). Applying eq 1 in turn to each ROE resonance (η_{IS}) observed in the 1D-ROESY spectra allows the calculation of the corresponding interproton distance (r_{IS}). Methyl groups were treated by using r^{-3} averaging as described in ref 1, as this gives slightly improved fits over r^{-6} averaging on molecules of the size of conicasterol. A subset of these values are shown in Table 2 which were identified as useful for stereochemical structure elucidation, i.e., those where DFT-calculated interproton distances in **1a** and **1b** differed by >0.03 Å (~1%) from each other.

Table 2. Interproton Distances Determined by ROE for H7β, H11β, H15β, Me18, and H29a–b of Conicasterol F (1) in Benzene-d_6 and Comparison with DFT-Calculated Value for 1a and 1b (value in bold was used to calibrate the ROEs)

proton		exptl, R_{ROE}(Å)	1a		1b	
			R_{calcd} (Å)	ABS % error[a]	R_{calcd} (Å)	ABS % error[a]
H29a	**H29b**	**1.85**	**1.85**	–	**1.85**	–
H7β	H15β	2.73	2.45	11.4	2.71	0.5
H11β	H12α	3.07	2.87	6.9	2.92	5.1
H11β	Me19	3.43	3.83	10.4	3.20	7.4
H11β	H1β	2.09	2.26	7.4	2.12	1.7
H15β	H16α	2.94	2.69	9.3	2.87	2.5
H15β	H16β	2.38	2.48	4.0	2.32	2.9
Me18	H15β	3.42	4.37	21.7	3.18	7.5
Me18	H11β	3.67	3.71	1.2	3.54	3.7
Me18	H12β	3.12	2.89	8.1	2.93	6.3
Me18	H16β	3.02	3.09	2.2	3.01	0.4
Me18	H20	2.70	2.61	3.1	2.74	1.6
Me18	Me19	3.31	4.00	17.2	3.28	1.0
H29a	H6α	2.34	2.44	4.2	2.35	0.5
H29a	H6β	2.39	2.33	2.7	2.38	0.5
[b]**MAE**				7.8		3.0
STD				5.9		2.6

[a]|% error| = $|r_{calcd} - r_{ROE}|/r_{calcd}$, absolute differences for calculated versus ROE-derived distances/calculated distances. [b]MAE = $\sum$[% error|]/n.

On the basis of the data in Table 2, it is clear that diastereomer **1b** shows the best fit with the estimated experimental data (MAE 3.0%, standard deviation (STD) 2.6%), while diastereomer **1a** is clearly less satisfactory (MAE 7.8%, STD 5.9%). Critically, the MAE and STD values for **1b** fall within the expected range of errors we have previously reported for this accurate NOE/ROE-distance method (MAE and STD both <4%), while those for **1a** fall well outside this range. As might be expected, the largest contributions to the

dx.doi.org/10.1021/jo2023763 | J. Org. Chem. 2012, 77, 1489–1496

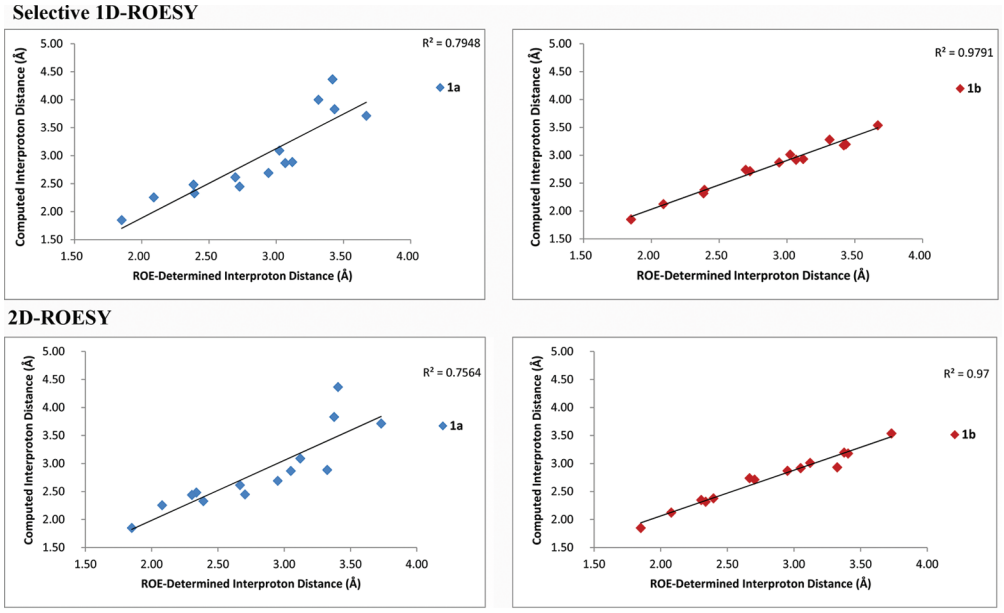


Figure 6. DFT-calculated interproton distances vs experimental from 1D-ROESY experiments (top view) and from 2D-ROESY experiments (bottom view).

MAE for **1a** (underlined in the Table 2) relate to the protons close to the C−D junction (Figure 2), e.g., H7β−H15β, and Me18−H15β, where the conformation of ring C is influenced by the relative stereochemistry of the epoxide rings. However, it must be emphasized that in both of these cases (H7β−H15β, and Me18−H15β), a qualitative ROE/NOE analysis would not have differentiated between the diastereomers, as the ROE peaks would have been present, but with very weak intensities for either diastereomer. 2D-ROESY spectra were also obtained for **1** and confirmed the trend of the 1D-ROESY data. Informative plots of computed vs ROE-determined interproton distances for each diastereomer (Figure 6) clearly show that **1b** provides a far superior fit for both 1D- and 2D-data. However, there is a very slight deterioration in the quality of the overall fit from the 2D-ROESY data, for the 'correct' stereoisomers **1b** (MAE 3.9%, STD 3.5%) (see Table S1, Supporting Information) which is in line with our previous observations.

These encouraging results are confirmed by our QM-NMR structural analysis. In particular, we performed single-point GIAO calculations on the optimized **1a** and **1b** geometries (see above and Figure 4) using the MPW1PW91 functional and the 6-31G (d,p) basis set[31] (Gaussian 09 Software Package).[32] For the ^{13}C chemical shift analysis, we have focused our attention on the key carbon atoms near the C ring because this is the region likely to be influenced by the stereochemistry of the epoxy ring. Moreover, we performed the data evaluation considering the $\Delta\delta$ parameter (differences in experimental vs calculated ^{13}C NMR chemical shifts) and the MAE parameter (MAE $=\sum[|\delta_{exp} - \delta_{calcd}|]/n$, summation through n absolute values of the differences in the corresponding experimental and calculated ^{13}C chemical shifts), which have been successfully used in the characterization of unknown stereostructures by us and by other research groups.[14,15,19] Table 3 lists selected experimental and calculated chemical shifts for carbons near the C ring in **1a** and **1b** (see Table S2, Supporting Information, for chemical shifts of entire molecule).

Once again, the chemical shift data strongly suggest that diastereomer **1b** is the correct structure for **1**, with a very large difference between the average $|\Delta\delta|$ values of the MAE values for **1a** and **1b** (3.7 vs 0.8, respectively). Critically, the $|\Delta\delta|$ MAE for **1b** in Table 3 falls within the MAE calculated for the whole molecule (1.4 ppm, Table S2) and is comparable to the expected error range (≤2.0 ppm) for QM/NMR calculations of this sort, while the $|\Delta\delta|$ MAE in Table 2 for **1a** falls outside these bounds. The graphical representation of the errors in calculated chemical shifts (Figure 7) clearly shows that **1b** is the best match with the experimental data for every carbon with the one exception of C-17 where both diastereomers give calculated values which are within reasonable error limits.

Table 3. Comparison between Experimental ^{13}C Chemical Shifts in Benzene with Calculated ^{13}C Chemical Shifts in Vacuo of Diastereoisomers 1a and 1b and Their $|\Delta\delta|$ Values

		1a		1b	
carbon	δ_{exp}	δ_{calcd}	$\lvert\Delta\delta\rvert$,[a] ppm	δ_{calcd}	$\lvert\Delta\delta\rvert$,[a] ppm
7	27.3	31.4	4.1	28.1	0.8
8	60.8	62.8	2.0	61.3	0.5
9	61.5	62.9	1.4	61.6	0.1
14	79.8	72.8	7.0	79.1	0.7
13	44.9	42.6	2.3	46.1	1.2
12	35.7	41.5	5.8	36.4	0.7
11	50.1	54.1	4.0	49.6	0.5
18	14.8	19.1	4.3	15.9	1.1
15	65.0	75.3	10.3	64.3	0.7
16	40.9	40.2	0.7	41.4	0.5
17	50.6	50.5	0.1	48.1	2.5
20	35.0	33.2	1.8	35.0	0.0
MAE[b] ^{13}C			**3.7**		**0.8**

[a] $|\Delta\delta| = |\delta_{exp} - \delta_{calcd}|$, absolute differences for experimental versus calculated ^{13}C NMR chemical shifts. [b] MAE $= \sum[|\delta_{exp} - \delta_{calcd}|]/n$.

In summary, the evaluation of the GIAO calculated ^{13}C chemical shifts MAEs for **1a** and **1b** geometries (3.7 vs 0.8 ppm, respectively) confirms the stereochemical assignment made on the basis of the accurate ROE-distance analysis. Both methods independently identify a trans C/D ring junction,

dx.doi.org/10.1021/jo2023763 | J. Org. Chem. 2012, 77, 1489−1496

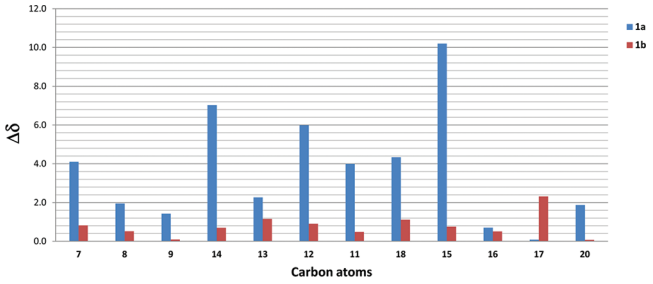


Figure 7. $|\Delta\delta|$ values (parts per million) of the experimental (benzene) versus theoretical carbon chemical shifts (vacuo) for Csp^3 of diastereoisomers **1a** and **1b**.

allowing the assignment of conicasterol F as **1b** depicted in Figure 3.

Pharmacological Evaluation of Conicasterol F (1) and Theonellasterol I (2). Following our recent demonstration[23] that 4-methylene steroids from *Theonella swinhoei* are potent ligands of human nuclear pregnane-X-receptor (PXR) and modulators of farnesoid-X-receptor (FXR), compounds **1** and **2** were challenged in a reporter gene assay using HepG2 cells, a human hepatoma cell line.

As shown in Figure 8, in the FXR transactivation assay, compounds **1** and **2** failed to activate FXR at the concentrations

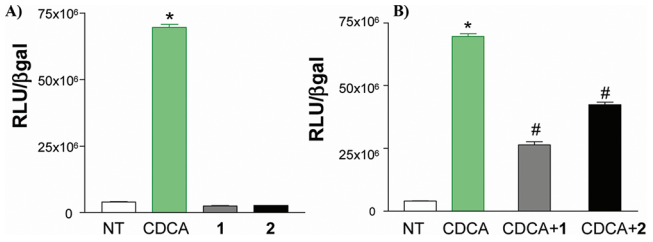


Figure 8. Transactivation of FXR by compounds **1** and **2** in HepG2 cells transfected with pSG5-FXR, pSG5-RXR, pCMV-βgal, and p(hsp27)TKLUC vectors and stimulated for 18 h with (A) chenodeoxycholic acid (CDCA) 10 μM, and **1** and **2**, 10 μM. Transactivation is expressed as a ratio between relative luciferase unit (RLU) and β-galactosidase activity. (B) CDCA, 10 μM, alone or in combination with compounds **1** and **2**, 50 μM. $^*P < 0.05$ versus not treated (NT). $^{\#}P < 0.05$ versus CDCA ($n = 4$).

of 10 μM (panel A) whereas effectively antagonized FXR activation induced by its endogenous ligand, the bile acid chenodeoxycholic acid CDCA, with a discrete difference in their potency (panel B). Notably, as previously demonstrated by us,[23] the presence at position 24 of a methyl group in conicasterol F (**1**) with an opposite configuration with respect to the ethyl group in theonellasterol I (**2**) allows a better fitting in the shallow groove on FXR molecular surface.

In addition to an FXR antagonistic activity, conicasterol F (**1**) effectively induced PXR expression with comparable efficacy to rifaximin in inducing PXR transactivation (Figure 9, panel A). Therefore, conicasterol F (**1**) represents another example of a dual ligand of human nuclear receptors, an antagonist of FXR and a potent agonist of PXR, of marine origin.

CONCLUSIONS

In this paper, we have described our new highly effective method for stereochemical structure determination, based on the combination of the quantitative ROE-derived interproton

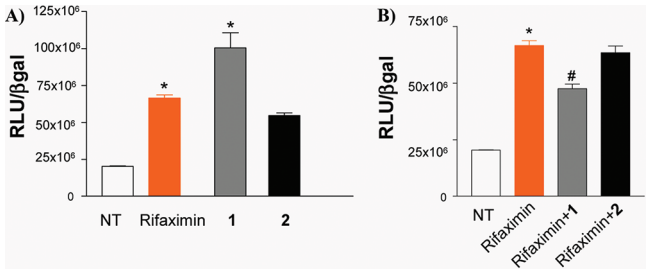


Figure 9. Transactivation of PXR by compounds **1** and **2** in HepG2 cells transfected with pSG5-PXR, pSG5-RXR, pCMV-βgal, and p(cyp3a4)TKLUC vectors and stimulated 18 h with (A) rifaximin 10 μM, and compounds **1** and **2**, 10 μM. Transactivation is expressed as a ratio between relative luciferase unit (RLU) and β-galactosidase activity. (B) Rifaximin 10 μM, alone or in combination with compounds **1** and **2**, 50 μM. $^*P < 0.05$ versus not treated (NT). $^{\#}P < 0.05$ versus rifaximin ($n = 4$).

distances and the QM calculation of the ^{13}C chemical shifts, and its first application to a natural product of unknown configuration. The stereochemical determination of the doubly quaternary C8−C14 epoxy functional group in conicasterol F (**1**), a new 4-methylenesterol derivative, was used as a case study to test our methodology and represents its first application. Experimentally determined interproton distances from 1D and/or 2D-ROESY spectra and experimental ^{13}C chemical shifts are compared with the corresponding calculated values by the evaluation of mean absolute errors and standard deviations and compared both internally against their expected error ranges and against each other through their structure predictions. In this way, a much more reliable stereochemical determination can be made. In this report, the simultaneous and crossed analysis of MAE values derived by ROEs and chemical shift data clearly identify **1b** as the correct steroisomer with the C8−C14 epoxy group on the α face of the marine sterol, giving MAEs of 3.0% (0.10 Å) and an average $|\Delta\delta|$ of 0.8 ppm.

In summary, we have demonstrated our method as a powerful tool in stereochemical structure determination, and we strongly recommend applying this combined methodology prior to, or indeed as a complete alternative to, total synthesis when traditional NMR data analysis is not sufficient to distinguish stereochemical alternative, such as in the case of conicasterol F (**1**).

Further, a pharmacological evaluation of conicasterol F (**1**) and theonellasterol I (**2**) as ligands of two well-known nuclear receptors, PXR and FXR, has been also reported. Farnesoid-X-receptor (FXR) and the pregnane-X-receptor (PXR), two nuclear receptors both functioning as bile acid-activated receptors, have emerged as the main receptors involved in regulating bile acid synthesis, detoxification, and excretion in the liver and gastrointestinal tract.[34−37] From the pharmacology standpoint, a dual ligand, such as conicasterol F (**1**), holds potential in the treatment of liver disorders characterized by cholestasis and/or impaired metabolism of xenobiotics and, because both FXR and PXR exert anti-inflammatory effects in the intestine, in the treatment of inflammatory bowel diseases. In conclusion, this discovery reaffirms the utility of examining natural product libraries for identifying novel receptor ligands potentially useful in the treatment of liver-related immune disorders.

dx.doi.org/10.1021/jo2023763 | J. Org. Chem. 2012, 77, 1489−1496

EXPERIMENTAL SECTION

General Procedures. Specific rotations were measured on a polarimeter. High-resolution ESI-MS spectra were recorded with a QTOF spectrometer. ESI-MS experiments were performed on an API 2000 triple-quadrupole mass spectrometer. NMR spectra were obtained on 500 MHz, and on 600 and 700 MHz NMR spectrometers equipped with an inverse detection cryogenic probe. Chemical shifts (δ, ppm), J in hertz, are referenced to C_6HD_5 as internal standard (δ_H 7.16, δ_C 128.4). HPLC was performed using a pump equipped with an injector and a differential refractometer. Silica gel (200−400 mesh) from Macherey-Nagel Company was used for flash chromatography. The purities of compounds were determined to be greater than 95% by HPLC.

For the assignment of the relative configuration by the determination of the interproton distance by ROE, the NMR sample was prepared in 3 mm tubes, 0.5 mL of C_6D_6 (0.7 mg conicasterol F), under air without degassing. 1D selective transient ROESY spectra were obtained using 1.7039 s acquisition time, 4096 scans for C_6D_6 samples; 2D-ROESY spectra were obtained with 1024 t2 points, 200 t1 points, 0.19 s t2 acquisition time, and 256 scans. All ROESY spectra employed 300 ms mixing times. 1D-ROESY data use 1 s relaxation delays, while 2D-ROESY data employ relaxation delays of 2s. NMR data were processed using MestreNova version7.

Sponge Material and Separation of Individual Sterols. Two taxonomic vouchers (field collection references R3170 and R3159) both assigned to the widespread species *Theonella swinhoei* (order Lithistida, family Theonellidae) were collected on the barrier reef of Vangunu and Malaita Island, Solomon Islands, in July 2004, and reference specimens are on files at the ORSTOM, Centre of Noumea. The samples were frozen immediately after collection and lyophilized to yield 600 and 207 g (dry weight) of R3170 and R3159, respectively. Taxonomic identification was performed by Dr. John Hooper of Queensland Museum, Brisbane, Australia.

Theonella swinhoei (R3159) lyophilized material (207 g) was extracted with methanol (3 × 1.5 L) at room temperature, and the crude methanolic extract was subjected to a modified Kupchan's partitioning procedure as follows. The methanol extract was dissolved in a mixture of MeOH/H_2O containing 10% H_2O and partitioned against *n*-hexane to give 4.5 g of the extract. The water content (% v/v) of the MeOH extract was adjusted to 30% and partitioned against $CHCl_3$ to give 6.0 g of the extract. The aqueous phase was concentrated to remove MeOH and then extracted with *n*-BuOH to give 10.3 g of the extract. The hexane extract, (4.5 g) was chromatographed by silica gel MPLC using a solvent gradient system from CH_2Cl_2 to $CHCl_2$:MeOH 1:1. The fraction eluted with CH_2Cl_2:MeOH 97:3 (76.8 mg) was further purified by HPLC on a Nucleodur 100-5 C18 (5 μm; 10 mm i.d. × 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 5 mL/min) to give 3.4 mg of conicasterol F (**1**) (t_R = 6.0 min).

Theonella swinhoei (R3170) lyophilized material (600 g) was extracted with methanol (3 × 2.7 L) at room temperature, and the methanolic extract, taken to dryness, was subjected to a modified Kupchan's partitioning procedure as described for R3159 sample. The hexane extract was chromatographed in two runs by silica gel MPLC using a solvent gradient system from CH_2Cl_2 to $CHCl_2$:MeOH 1:1. The fraction eluted with CH_2Cl_2:MeOH 97:3 (97 mg) was further purified by HPLC on a Nucleodur 100-5 C18 (5 μm; 10 mm i.d. × 250 mm) with MeOH:H_2O (92:8) as eluent (flow rate 5 mL/min) to give 2.7 mg of theonellasterol I (**2**) (t_R = 8.0 min)

Characteristic Data for Each Sterol. *Conicasterol F (**1**).* $[\alpha]^{25}_D$ +57.2 (*c* 0.44, MeOH); ^{1}H and ^{13}C NMR data in C_6D_6 given in Table 1; ESI-MS: *m*/*z* 465.4 $[M + Li]^+$. HRMS (ESI): calcd for $C_{29}H_{46}LiO_4$: 465.3556; found 465.3532 $[M + Li]^+$.

*Theonellasterol I (**2**).* $[\alpha]^{25}_D$ +39.0 (*c* 0.17, MeOH); ^{1}H and ^{13}C NMR data in C_6D_6 given in Table S5 in Supporting Information; ESI-MS: *m*/*z* 479.4 $[M + Li]^+$. HRMS (ESI): calcd for $C_{30}H_{48}LiO_4$: 479.3733; found 479.3713 $[M + Li]^+$.

Plasmids, Cell Culture, Transfection, and Luciferase Assays. All transfections were made using Fugene HD transfection reagent (Roche). For FXR-mediated transactivation, HepG2 cells, plated in a six-well plate at 5 × 10^5 cells/well, were transfected with 100 ng of pSG5-FXR, 100 ng of pSG5-RXR, 200 ng of pCMV-βgalactosidase, and 500 ng of the reporter vector p(hsp27)-TK-LUC containing the FXR response element IR1 cloned from the promoter of heat shock protein 27 (hsp27). At 48 h post-transfection, cells were stimulated 18 h with 10 μM CDCA or with the compounds **1** and **2** (10 μM) alone or in combination (50 μM) with CDCA.

For PXR-mediated transactivation, HepG2 cells, plated in a six-well plate at 5 × 10^5 cells/well, were transfected with 100 ng of pSG5-PXR, 100 ng of pSG5-RXR, 200 ng of pCMV-βgalactosidase, and 500 ng of the reporter vector containing the PXR target gene promoter (CYP3A4 gene promoter) cloned upstream of the luciferase gene (pCYP3A4promoter-TKLuc). At 48 h post-transfection, cells were stimulated 18 h with 10 μM rifaximin or with the compounds **1** and **2** (10 μM) alone or in combination (50 μM) with rifaximin. Cells were lysed in 100 μL of diluted reporter lysis buffer (Promega), and 0.2 μL of cellular lysates was assayed for luciferase activity using the Luciferase Assay System (Promega). Luminescence was measured using an automated luminometer. Luciferase activities were normalized for transfection efficiencies by dividing the relative light units by β-galactosidase activity expressed from cells cotransfected with pCMV-βgal.

Computational Details. Molecular dynamics and Monte Carlo calculations were both performed on conicasterol F (**1**) on a 4x AMD Opteron SixCore 2.4 GHz. The dynamics calculations of **1** were performed at two different temperatures (600 and 750 K for 8 ns using 1.5 fs as time-step) using the MMFF[28] force field (MacroModel software package).[29,30] During both calculations, a standard constant temperature velocity−Verlet algorithm was used to integrate the equations of motions.[38]

All the obtained structures (numbering ∼150, selected at regular intervals throughout the simulation) from Molecular Dynamics calculations for each isomer were minimized using the Polak-Ribier conjugate gradient algorithm (PRCG, 100 000 steps, convergence threshold 0.005 kJ mol^{-1} $Å^{-1}$), leading to the selection of the lowest energy minimum conformer for both diastereoisomers. The distribution of the resulting geometries were in accordance with the results of a parallel conformational search performed with the Monte Carlo multiple minimum (MCMM) method (MMFFs,[28] 10 000 steps, numbering 150, stored on a similarity and an energy criterion), where the variables used for the calculations included all the possible rotable torsions. The empirical geometries from the conformer search for diastereoisomers **1a** and **1b** were optimized in vacuo at the DFT MPW1PW91 level using the 6-31G(d)[31] basis set (Gaussian 09 software package),[32] and then the optimized structures were used as inputs for the single-point ^{13}C chemical shift calculations performed in vacuo employing the same functional combined with the 6-31G(d,p)[31] basis set. The calculated values of chemical shifts of conicasterol F (**1**) were referenced to the theoretical tetramethylsilane ^{13}C chemical shift value (previously optimized at DFT level), computed at the same level of theory.

ASSOCIATED CONTENT

Supporting Information

^{1}H, HMBC, HSQC, and COSY NMR spectra for conicasterol F (**1**) and theonellasterol I (**2**). Selective 1D-ROESY and 2D ROESY experiments for conicasterol F (**1**). Table of internuclear distances obtained by 2D ROESY experiments. Cartesian coordinates of **1a** and **1b**. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Craig.Butts@bristol.ac.uk; bifulco@unisa.it.

■ ACKNOWLEDGMENTS

This work was supported by grants from MIUR (PRIN 2009), Rome, Italy. We thank the marine natural products and field team of IRD for providing the biological material, Solomon Islands government for collection permits and Dr. John Hooper at Queensland Museum, Brisbane, Australia, for the identification of the sponges.

■ REFERENCES

(1) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, 2nd revised ed.; John Wiley & Sons Inc.: Hoboken, NJ, 2000.
(2) Stonehouse, J.; Adell, P.; Keeler, J.; Shaka, A. J. *J. Am. Chem. Soc.* **1994**, *116*, 6037−6038.
(3) Stott, K.; Keeler, J.; Van, Q. N.; Shaka, A. J. *J. Magn. Reson.* **1997**, *125*, 302−324.
(4) Stott, K.; Stonehouse, J.; Keeler, J.; Hwang, T. L.; Shaka, A. J. *J. Am. Chem. Soc.* **1995**, *117*, 4199−4200.
(5) Thrippleton, M. J.; Keeler, J. *Angew. Chem., Int. Ed.* **2003**, *42*, 3938−3941.
(6) Macura, S.; Farmer, B. T. II; Brown, L. R. *J. Magn. Reson.* **1986**, *70*, 493−499.
(7) Hu, H.; Krishnamurthy, K. *J. Magn. Reson.* **2006**, *182*, 173−177.
(8) Butts, C. P.; Jones, C. R.; Towers, E. C.; Flynn, J. L.; Appleby, L.; Barron, N. *Org. Biomol. Chem.* **2011**, *9*, 177−184.
(9) Butts, C. P.; Jones, C. R.; Harvey, J. N. *Chem. Commun.* **2011**, *47*, 1193−1195.
(10) Butts, C. P.; Jones, C. R.; Harvey, J. N. *Beilstein J. Org. Chem.* **2011**, *7*, 145−150.
(11) As we have noted previously in ref 8, this approach also offers an alternative test for the fulfillment of the initial rate approximation by using a single 1D-NOESY mixing time (and thus not needing a full NOE build-up curve for every selective irradiation, which is extremely time-consuming) on a given sample, i.e., where the initial rate approximation holds true, all of the relative intensities of peaks within the spectra will be constant as a function of mixing time.
(12) Barone, G.; Gomez-Paloma, L.; Duca, D.; Silvestri, A.; Riccio, R.; Bifulco, G. *Chem.—Eur. J.* **2002**, *8*, 3233−3239.
(13) Barone, G.; Duca, D.; Silvestri, A.; L. Gomez-Paloma, L.; Riccio, R.; Bifulco, G. *Chem.—Eur. J.* **2002**, *8*, 3240−3245.
(14) Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, R. *Chem. Rev. (Washington, DC, U. S.)* **2007**, *107*, 3744−3779.
(15) Di Micco, S.; Chini, M. G.; Riccio, R.; Bifulco, G. *Eur. J. Org. Chem.* **2010**, *8*, 1411−1434.
(16) Di Micco, S.; Boger, D. L.; Riccio, R.; Bifulco, G. *Eur. J. Org. Chem.* **2008**, *14*, 2454−2462.
(17) Chini, M. G.; Riccio, R.; Bifulco, G. *Magn. Reson. Chem.* **2008**, *46*, 962−968.
(18) Tähtinen, P.; Saielli, G.; Guella, G.; Mancini, I.; Bagno, A. *Chem.—Eur. J.* **2008**, *14*, 10445−10452.
(19) Smith, S. G.; Goodman, J. M. *J. Am. Chem. Soc.* **2010**, *132*, 12946−12959.
(20) Festa, C.; De Marino, S.; Sepe, V.; Monti, M. C.; Luciano, P.; D'Auria, M. V.; Debitus, C.; Bucci, M.; Vellecco, V.; Zampella, A. *Tetrahedron* **2009**, *65*, 10424−10429.
(21) Sepe, V.; D'Auria, M. V.; Bifulco, G.; Ummarino, R.; Zampella., A. *Tetrahedron* **2010**, *66*, 7520−7526.
(22) Festa, C.; De Marino, S.; Sepe, V.; D'Auria, M. V.; Bifulco, G.; Debitus, C.; Bucci, M.; Vellecco, V.; Zampella, A. *Org. Lett.* **2011**, *13*, 1532−1535.
(23) De Marino, S.; Ummarino, R.; D'Auria, M. V.; Chini, M. G.; Bifulco, G.; Renga, B.; D'Amore, C.; Fiorucci, S.; Debitus, C.; Zampella, A. *J. Med. Chem.* **2011**, *54*, 3065−3075.
(24) Kho, E.; Imagawa, D. K.; Rohmer, M.; Kashman, Y.; Djerassi, C. *J. Org. Chem.* **1981**, *46*, 1836−1839.
(25) Sepe, V.; Ummarino, R.; D'Auria, M. V.; Mencarelli, A.; D'Amore, C.; Renga, B.; Zampella, A.; Fiorucci, S. *J. Med. Chem.* **2011**, *54*, 4590−4599.
(26) Wright, J. L. C.; Mclnnes, A. G.; Shimizu, S.; Smith, D. G.; Walter, J. A.; Idler, D.; Khalil, W. *Can. J. Chem.* **1978**, *56*, 1898−1903.
(27) Horibe, I.; Nakai, H.; Sato, T.; Seo, S.; Takeda, K.; Takatsuto, S. *J. Chem. Soc., Perkin Trans. 1* **1989**, 1957−1967.
(28) Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 720−729.
(29) Mohamadi, F.; Richards, N. G.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440−467.
(30) *MacroModel, version 8.5*, Schrödinger LLC: New York, 2003.
(31) Cimino, P.; Duca, D.; Gomez-Paloma, L.; Riccio, R.; Bifulco, G. *Magn. Reson. Chem.* **2004**, *42*, S26−S33.
(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.02*, Gaussian, Inc., Wallingford, CT, 2009.
(33) The proton 7α and the H_2-12 being in a crowed region of the ^{1}H spectra (see Supporting Information) were ignored because the selective irradiation of the correct proton was impossible to perform.
(34) Pellicciari, R.; Costantino, G.; Fiorucci, S. *J. Med. Chem.* **2005**, *48*, 5383−5403.
(35) Fiorucci, S.; Rizzo, G.; Donini, A.; Distrutti, E.; Santucci, L. *Trends Mol. Med.* **2007**, *13*, 298−309.
(36) Sonoda, J.; Xie, W.; Rosenfeld, J. M.; Barwick, J. L.; Guzelian, P. S.; Evans, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 13801−13806.
(37) Guo, G. L.; Lambert, G.; Negishi, M.; Ward, J. M.; Brewer, H. B. J.; Kliewer, S. A.; Gonzalez, F. J.; Sinal, C. J. *J. Biol. Chem.* **2003**, *278*, 45062−45071.
(38) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637−649.