ORGANIC LETTERS

2013
Vol. 15, No. 1
156–159

Nudicaulins, Yellow Flower Pigments of *Papaver nudicaule*: Revised Constitution and Assignment of Absolute Configuration

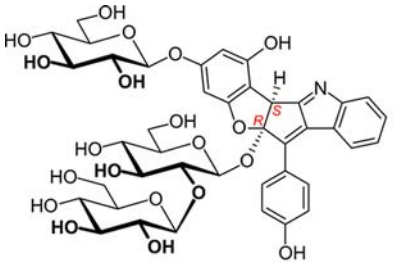
Evangelos C. Tatsis,[†] Anu Schaumlöffel,[‡] Anne C. Warskulat,[†] Georges Massiot,[§] Bernd Schneider,[*,†] and Gerhard Bringmann[*,‡]

Max Planck Institute for Chemical Ecology, Hans-Knöll-Strasse 8, Beutenberg Campus, 07745 Jena, Germany, University of Würzburg, Institute of Organic Chemistry, Am Hubland, 97074 Würzburg, Germany, and USR CNRS-Pierre Fabre No. 3388 ETaC, Center of Research and Development of Pierre Fabre, 3 avenue Hubert Curien, 31035 Toulouse Cedex 01, France

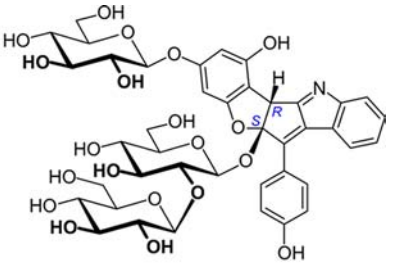
schneider@ice.mpg.de; bringman@chemie.uni-wuerzburg.de

Received November 21, 2012

ABSTRACT



Nudicaulin I



Nudicaulin II

The constitution of the aglycon of nudicaulin has been revised to be a pentacyclic indole alkaloid. The relative and absolute configurations of the two diastereomers, nudicaulins I (3a) and II (3b), have been assigned by NMR, conformational analyses, and interpretation of the experimental ECD spectra by quantum-chemical calculations.

In 1939, Price et al.[1] isolated the pigment responsible for the yellow color of *Papaver nudicaule* flower petals and named it nudicaulin. The first investigations of the nudicaulin structure by elemental analysis, oxidative degradation, and acid hydrolysis suggested the presence of nitrogen and a 4-hydroxyphenyl moiety in a diglucosidic compound. In a later work, a triglucoside was claimed from enzymatic degradation and the spectral properties indicated that nudicaulin was neither a flavonoid nor a carotenoid nor a betaxanthin.[2] Based on NMR spectroscopy and mass spectrometric analyses, Schliemann et al.[3] proposed the structure **1**, with a pentacyclic indole alkaloid skeleton (Figure 1) for the aglycon of the diastereomeric nudicaulins. However, the *N*-containing six-membered ring **e** in this postulated structure appeared to be severely distorted and certainly highly strained. Moreover, the interruption of the conjugated system is not consistent with the UV absorption maximum at ca. 460 nm. Based on further NMR experiments and chemical modifications, we describe here the revision of the previously postulated constitution of the aglycon of these compounds as **2**, using nudicaulin I (**3a**) as a representative example (Figure 1). The AX system of H-6 and H-8 (ring **a**), a four-spin system of H-15−H-18 (**d**), an AA′XX′ system of H-2′/6′ and H-3′/5′ (**f**), and a singlet of H-3 were attributed to the aglycon. HSQC and HMBC correlations (Figures S5−S7) assigned all carbon atoms within the

[†] Max Planck Institute for Chemical Ecology.
[‡] University of Würzburg.
[§] USR CNRS-Pierre Fabre No. 3388 ETaC.

(1) Price, J. R.; Robinson, R.; Scott-Moncrieff, R. *J. Chem. Soc.* **1939**, 1465–1468.
(2) Harborne, J. B. *Phytochemistry* **1965**, *4*, 647–657.
(3) Schliemann, W.; Schneider, B.; Wray, V.; Schmidt, J.; Nimtz, M.; Porzel, A.; Böhm, H. *Phytochemistry* **2006**, *67*, 191–201.

10.1021/ol303211w
Published on Web 12/18/2012

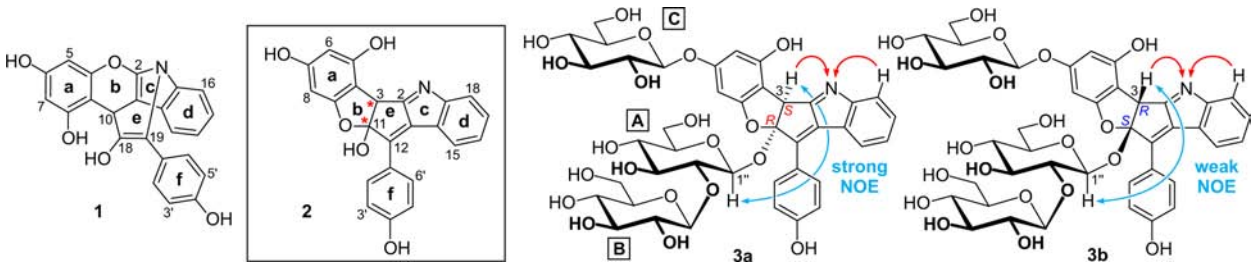


Figure 1. Constitutions of the aglycon core as previously published (**1**)[3] and currently revised (**2**), and structures of the diastereomeric nudicaulins I (**3a**) and II (**3b**). Red arrows indicate $^{1}H,^{15}N$ HMBC correlations.

substructures of rings **a**, **d**, and **f**, which were not subject to revision, and established H-3/C-3 (δ_H 5.55/δ_C 48.9) as a key position connecting these parts of the aglycon. HMBC correlations of H-3 with C-4 (δ 101.7), C-5 (δ 156.0), and C-9 (δ 160.4) linked C-3 via a C−C bond with C-4 to ring **a**. A correlation of H-3 with C-2 (δ 177.1) and C-13 (δ 130.8) connected C-3 to the N-containing ring **c** of the indole unit and also substantiated the position of the nitrogen next to C-2, as also confirmed by its chemical shift at a low field. Two additional HMBC cross signals of H-3, one of them with C-12 (δ 165.0), which is also connected through an HMBC correlation with H-2′/6′ of ring **f**, and the other one with the acetal carbon C-11 (δ 126.4), completed the assignment of the carbon atoms of rings **b** and **e**. The position of the nitrogen atom (δ 172.7) was established by $^{1}H,^{15}N$-HMBC correlations (Figure S14) of H-3 and H-18 with the indole nitrogen. The $^{1}H,^{15}N$-HMBC correlations in combination with a strong signal in the ROESY spectrum (Figure S13) between H-15 and H-2′/6′ indicated a [2′,3′:1]-annelation of the indole moiety to the central cyclopentene ring **e**. The attachment of the carbohydrate units to the aglycon through oxygen atoms and the interunit link between the two glucose units A and B was established by means of HMBC correlations of H-1″ of Glc **A** with C-11, H-1‴′ of Glc **C** with C-7, and H-1‴ of Glc **B** with C-2″ of Glc **A**. The number and attachment positions of the sugar and acyl units of nudicaulins I−VIII[3] remained unaffected by the suggested structural revision of the aglycon.

To further confirm the constitution of the aglycon **2**, the nudicaulins were structurally modified. By hydrogenation, additional protons were introduced into the proton-deficient region of the aglycon, especially in ring **e** (Figure 2a). The successful reaction was confirmed by HRMS (found for **4a**: 874.27482; calcd: 874.27642). The only UV maximum at 255 nm suggested an interuption of the conjugated double-bond system. Thus the ^{1}H NMR spectrum of dihydronudicaulin I (**4a**) displayed aglycon signals shifted to a higher field as compared to the parent compound **3a**, consistent with a reduced number of conjugated double bonds. In addition, a new singlet at δ 4.99 appeared in the ^{1}H NMR spectrum, which correlated with δ_C 55.0 (C-12) in the HSQC spectrum (Figure S29) and showed HMBC cross peaks with C-13 (δ 116.5), C-11 (δ 126.8), C-1′ (δ 130.4), C-2′/6′ (δ 132.4), and C-2 (δ 140.5) (Figure S30). In the $^{1}H-^{1}H$ correlation spectrum optimized for long-range couplings ($^{1}H-^{1}H$ lrCOSY) (Figure S28), the H-12 signal correlated with H-3 (δ 5.07) and H-2′/6′ (δ 7.10). This showed that the quaternary C-12 of **3a** had been converted to a methine in dihydronudicaulin I (**4a**). H-3 still remained the key signal, now showing HMBC cross peaks with C-2, C-4 (δ 108.3), C-5 (δ 155.1), C-9 (δ 161.2), C-11, and C-13. The chemical shift of the acetal carbon C-11 in **4a** (δ 126.8) was the same as that in nudicaulin I (**3a**) (δ 126.7). For complete ^{1}H and ^{13}C NMR data, see Table S3. In summary, analysis of the dihydronudicaulins (**4a**/**b**) further confirmed the constitution of the nudicaulin aglycon (**2**) itself, especially the connection between the indole unit and the rest of the molecule.

In addition, the nudicaulins (**3a**/**b**) were converted to their *O*- and *N*-permethyl derivatives (**5a**/**b**) by applying a Williamson ether synthesis (Figure 2b).[4] By ^{1}H NMR, HSQC (Figure S38), and HMBC spectra (Figure S39) the new *O*- and *N*-methyl signals in the aglycon part of the permethylated nudicaulin were assigned to the positions that correspond to hydroxy and amino groups in the underivatized structures. In the NMR spectra of permethylated nudicaulin I (**5a**), three methyl signals were observed that were categorized as *O*-CH_3 (δ_H 4.01/δ_C 56.1; δ_H 4.05/δ_C 56.6) and *N*-CH_3 (δ_H 4.35/δ_C 35.7) according to their chemical-shift values. HMBC correlations assigned the *O*-CH_3 groups to C-4′ (δ 168.2) in ring **f** and C-5 (δ 157.9) in ring **a**, while the *N*-CH_3 group, as expected, correlated with C-2 (δ 177.6) and C-19 (δ 149.6). Hence, permethylation evidenced the position of the free hydroxy groups and of the nitrogen atom, thus, finally revising the previously postulated[3] constitution of the nudicaulin aglycon as 12-(4-hydroxyphenyl)-3,11-dihydrobenzo-furo[2′,3′:1]-cyclopenta[1,2-*b*]indole-5,7,11-triol (**2**).

The diastereomeric nudicaulins I (**3a**) and II (**3b**) contain two stereogenic centers in the aglycon. This made an investigation of their relative and absolute configurations necessary. A strong cross signal in the ROESY spectrum (Figure S13) between H-3 and H-1″ of Glc **A** of nudicaulin I (**3a**) showed that H-3 and the Glc **A** were oriented to the

(4) Ciucanu, I.; Kerek, F. *Carbohydr. Res.* **1984**, *131*, 209–217.

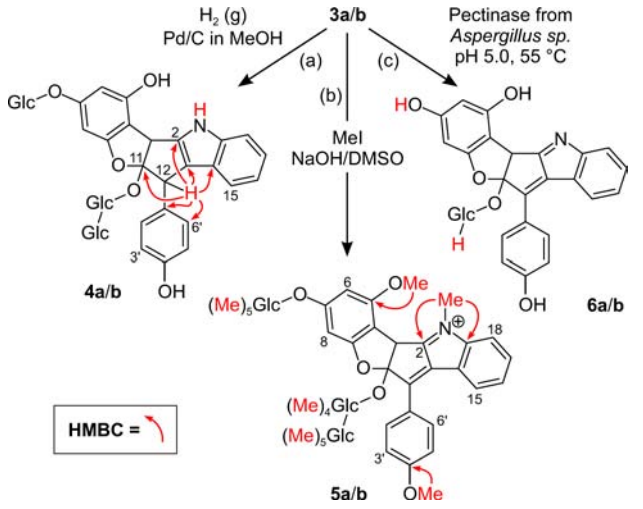


Figure 2. Modification of nudicaulins I and II (**3a**/**b**) and 'new' HMBC interactions (in red). (a) Catalytic hydrogenation. (b) Permethylation. (c) Enzymatic hydrolysis by a pectinase.

same side of the aglycon skeleton, i.e., *cis* to each other (Figure 1). Interestingly, nudicaulin II (**3b**) (Figure S20) showed such a ROESY cross signal, too, which was also in agreement with a relative *cis*-configuration, but it was far less intense. Relative to the integral intensities of the ROE signals between H-2′/6′ and H-3′/5′, which were used as a reference (100%), the integral intensity of the H-3 − H-1″ cross signal of nudicaulin I (**3a**) was 32% and that of nudicaulin II (**3b**) was 19%. A conformational search with B97D/SVP resulted in a minimum structure with a distance *d*(H-3,H-1″) of 2.4 Å for the (3*S*,11*R*)-**3** diastereomer (*cis*), which was consistent with the strong ROE of nudicaulin I (**3a**), while the longer distance *d*(H-3,H-1″), 3.8 Å, for the lowest-energy conformer of the (3*R*,11*S*)-**3** diastereomer (also *cis*) was in accordance with the weaker ROE of nudicaulin II (**3b**). For (3*R*,11*R*)-**3** as a likewise imaginable *trans*-isomer, a distance *d*(H-3,H-1″) of about 4.5 Å was predicted, from which no ROE signal would be expected, because the aglycon core was in between and, therefore, could be excluded (Figure S42).

In the ROESY spectra of dihydronudicaulins I (**4a**) and II (**4b**) (Figures S31 and S36), ROE signals were observed between H-3 and H-1″ of glucose **A**, between H-12 and H-1″ of glucose **A**, between H-12 and H-2′/6′, and between H-3 and H-2′/6′. No ROE signal was observed between H-3 and H-12 suggesting an opposite orientation with respect to the plane formed by the indole moiety and ring **e** (corresponds to *trans*-orientation of Glc **A** and H-12 and *trans*-orientation of the *p*-hydroxyphenyl ring and the benzofuran moiety). According to these data, the hydrogenation had occurred with a high degree of stereoselectivity.

For the elucidation of the absolute configuration of the nudicaulins **3a**/**b**, electronic circular-dichroism (ECD) spectroscopy in combination with quantum-chemical ECD calculations[5] was the method of choice. A first analysis of the experimental UV absorptions of **3a** and **3b** measured in two different solvent systems revealed a strong solvent dependence of the first UV band. In the MeCN−H_2O mixture, this absorption appeared with high intensity around 460 nm, possessing a broad blue-shifted shoulder, while two signals of low intensity around 460 and 385 nm were observed in MeOH (Figure S43).

First excited-state calculations for the *cis*-configured nudicaulins I and II were performed with time-dependent (TD) CAM-B3LYP, including CPCM for solvent effects of MeOH. However, the calculated spectra did not show sufficiently good agreement with the experimental ones (Figure S46). A reason for this could be interactions between the solvent and the glucose moieties that might influence the chiroptical properties.

Therefore, an arbitrarily chosen number of water molecules (12) were added to the structure of the two *cis*-diastereomers, (3*S*,11*R*)-**3** and (3*R*,11*S*)-**3**, using the CAST program.[6] The conformations thus identified were further optimized with B97-D/SVP and then submitted to semiempirical ZINDO/S-CI UV calculations. The results were compared with the UV curves calculated for identical conformations, without water molecules. The ones including water possessed a red-shifted, intensified first absorption band, thus corroborating the substantial effect of the solvent (Figure S45, Table S5). These findings showed that within TDDFT the CPCM approach alone would possibly not be sufficient to describe the excited-state properties of the authentic nudicaulins I and II properly, even though the effects seemed smaller in the experiment when using methanol. In addition to the ZINDO/S-CI tests, the analysis of the excited states of the TDDFT computations showed that within the aglycon core electronic transitions with charge-transfer (CT) character occurred within the same wavelength region that was affected by the solvent. Since more advanced methods to correctly account for these properties would have led to an immense increase in computational costs, the compound had to be simplified synthetically, at least to minimize the solvent effects.

Attempts to remove all three glucose units were unsuccessful, since, apparently, Glc **A** is needed to prevent decomposition of the chemically unstable aglycon core. Therefore, nudicaulins I and II could only be partially deglucosylated to give their monoglucosidic forms **6a**/**b** (Figure 2c), using a commercial pectinase. Compared to the nudicaulins I and II, the NMR spectra of the monoglucosides **6a** and **6b** (Figures S22, S23, S25) were considerably simplified in the carbohydrate regions. The ROESY spectra of the two diastereomeric monoglucosides (Figures S24, S26) confirmed the results that had been obtained for nudicaulins I (**3a**) and II (**3b**), indicating an unchanged aglycon structure after hydrolysis.

(5) For selected reviews, see: (a) Warnke, I.; Furche, F. *WIREs Comp. Mol. Sc.* **2012**, *2*, 150–166. (b) Pescitelli, G.; Di Bari, L.; Berova, N. *Chem. Soc. Rev.* **2011**, *40*, 4603–4625. (c) Autschbach, J. *Chirality* **2009**, *21*, E116–E152. (d) Bringmann, G.; Bruhn, T.; Maksimenka, K.; Hemberger, Y. *Eur. J. Org. Chem.* **2009**, 2717–2727.

(6) Grebner, C.; Becker, J.; Stepanenko, S.; Engels, B. *J. Comput. Chem.* **2011**, *32*, 2245–2253.

The offline CD spectra of **6a** and **6b** were also recorded in methanol, yielding nearly mirror-imaged curves, in which the first Cotton effect now occurred around 385 nm, instead of 460 nm in compounds **3a**/**3b** (Figure S44), showing that a comparison of the spectra calculated for a simplified structure with the experimental ones of the authentic nudicaulins would definitely not have been possible in this case.

Still, the agreement of the spectra of the monoglucosides obtained with CAM-B3LYP was not truly improved as compared to the experimental ones (Figure S47). Using the double-hybrid functional B2GP-PLYP (with the COSMO solvent field for methanol), the energies and rotatory strengths of the first excitations were still not quite well predicted in the case of the (3*R*,11*S*)-configured monoglucoside **6** (Figure S48). By contrast, B2GP-PLYP computations of the hypothetic—since chemically unstable—aglycon **2** provided a much better match with the experimental CD data of the monoglucosides **6a**/**b** (Figure S49).

To understand this discrepancy, coupled-cluster calculations with RICC2/def2-SV(P) were carried out for the first 25 excitations of the (3*R*,11*R*)-configured aglycon **2**[7] and, exemplarily, for the first five excited states of the (3*R*,11*S*)-configured monoglucoside **6**. The main objective was to determine the correct energetic positions of the CT transitions and to validate the impact of the remaining glucose moiety within the molecule upon excitation. Comparison of the RICC2 excitations, however, did not show any substantial differences between **6** and **2**. The B2GP-PLYP spectra obtained for the aglycon also seemed comparable with RICC2, although the first excited state had a clearly different main configuration (Table S6). Deviating from this, however, was the presence of the glucose unit in the B2GP-PLYP calculations leading to stronger changes of the rotatory-strength values within the first three excited states, of which the first one showed significant CT character (Figure S50), which obviously occurred at a wrong position and falsified the sign and intensity of the overall CD effect between 360 and 390 nm. Thus, already one glucose unit significantly hampered the applicability of TD methods, making the coupled-cluster approach mandatory for this kind of structure.

To conclude, the comparison of the experimental CD spectra of the monoglucosides of nudicaulins I (**6a**) and II (**6b**) with the ones computed for the *cis*-aglycon **2** with RICC2/def2-SV(P) allowed an unambiguous attribution of the absolute configuration (Figure 3). The CD curve calculated for (3*S*,11*S*)-configured **2** was in perfect agreement with the experimental spectrum of the monoglucoside of nudicaulin I (**6a**). This determined **6a** to have the (3*S*,11*R*)-configuration.[7] Analogously, the spectra of (3*R*,11*R*)-configured **2** and the monoglucosidic nudicaulin II (**6b**) provided a clear match, assigning the latter to possess the (3*R*,11*S*)-configuration.

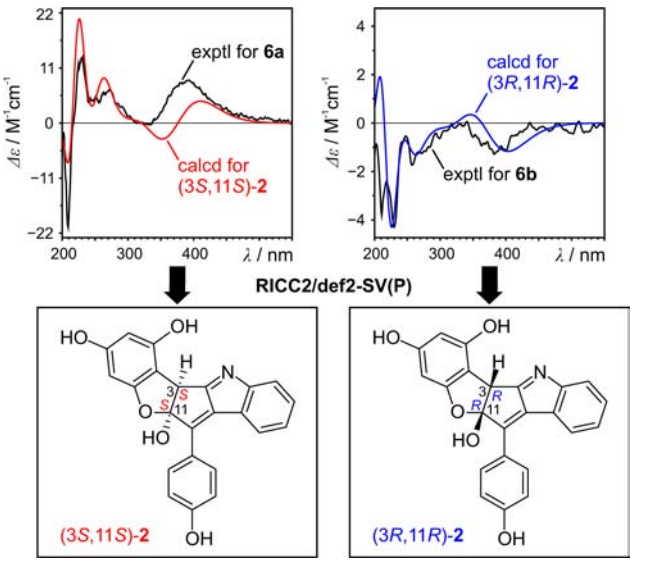


Figure 3. Absolute configurations of **6a**/**b** by comparison of the exptl CD curves with calcd curves for the aglycon model **2**, using RICC2/def2-SV(P) (UV shift = 29 nm, bandwidth σ = 0.40 eV).

From these results, the absolute configurations of the parent compounds nudicaulins I (**3a**) and II (**3b**) were eventually deduced to be (3*S*,11*R*) and (3*R*,11*S*), respectively. This structural assignment became possible only by the experimental decrease of the strong solvent effects of the glucose moieties, which in turn allowed the application of accurate higher-level quantum-chemical methods to predict the excited-state properties. The occurrence of diastereomeric pairs of nudicaulins with racemic (or scalemic) aglycon parts raises a question regarding the biosynthesis, which remains to be investigated.

Acknowledgment. This work was supported by an FP7 Marie Curie Intra-European Fellowship from E.U. to E.C.T. (BIOSYN-NUDICAUL project, Contract Number 221274). We are grateful to B. Amslinger (University of Würzburg) for the CD measurements.

Supporting Information Available. Experimental procedures; tables with ^{1}H and ^{13}C NMR data; MS and UV—vis data; NMR spectra of compounds **3**—**6**; computational details. This material is available free of charge via the Internet at http://pubs.acs.org.

(7) For the formal reason of the CIP convention, the altered priorities at C-11 lead to a change of the stereodescriptor.

The authors declare no competing financial interest.