▌ **Chlorophyll Metabolism**

Chlorophyll Catabolites in Fall Leaves of the Wych Elm Tree Present a Novel Glycosylation Motif

Mathias Scherl,[a] Thomas Müller,[a] Christoph R. Kreutz,[a] Roland G. Huber,[c, d] Engelbert Zass,*[b] Klaus R. Liedl,*[c] and Bernhard Kräutler*[a]

Dedicated to Professor Duilio Arigoni

Abstract: Fall leaves of the common wych elm tree (*Ulmus glabra*) were studied with respect to chlorophyll catabolites. Over a dozen colorless, non-fluorescent chlorophyll catabolites (NCCs) and several yellow chlorophyll catabolites (YCCs) were identified tentatively. Three NCC fractions were isolated and their structures were characterized by spectroscopic means. Two of these, *Ug*-NCC-27 and *Ug*-NCC-43, carried a glucopyranosyl appendage. *Ug*-NCC-53, the least polar of these NCCs, was identified as the formal product of an intramolecular esterification of the propionate and primary glucopyranosyl hydroxyl groups of *Ug*-NCC-43. Thus, the glucopyranose moiety and three of the pyrrole units of *Ug*-NCC-53 span a 20-membered ring, installing a bicyclo[17.3.1]glycoside moiety. This structural motif is unprecedented in heterocyclic natural products, according to a thorough literature search. The remarkable, three-dimensional bicyclo[17.3.1]-glycoside architecture reduces the flexibility of the linear tetrapyrrole. This feature of *Ug*-NCC-53 is intriguing, considering the diverse biological effects of known bicyclo[*n*.3.1]glycosidic natural products.

The annual disappearance of chlorophyll (Chl) has been a remarkable biological enigma until recently.[1] In 1991, *Hv*-NCC-1 (**1 a**) was identified in senescent leaves of barley (*Hordeum vulgare*) as the first colorless linear tetrapyrrole derived from Chl.[2] 1-Formyl-19-oxobilane-type tetrapyrroles, such as **1**, are now also classified as non-fluorescent Chl catabolites (NCCs).[3] Indeed, NCCs were found to accumulate in a variety of senescent leaves, apparently as "final" tetrapyrrolic products of a common pathway of the Chl-breakdown (see Figure 1).[3–5] However, in recent studies, the degradation of Chl was recognized to branch out, for example, to give persistent "hypermodified" fluorescent Chl-catabolites (*hm*FCCs),[6] as well as dioxo-

[a] M. Scherl, Prof. T. Müller, Dr. C. R. Kreutz, Prof. B. Kräutler
Institute of Organic Chemistry and Center of Molecular Biosciences
University of Innsbruck, Innrain 80/82, 6020 Innsbruck (Austria)
E-mail: bernhard.kraeutler@uibk.ac.at

[b] Dr. E. Zass
Laboratory of Organic Chemistry, ETH Zürich
Vladimir-Prelog-Weg 3, 8093 Zürich (Switzerland)
E-mail: zass@retired.ethz.ch

[c] Dr. R. G. Huber, Prof. K. R. Liedl
Institute of General, Inorganic & Theoretical Chemistry
and Center of Molecular Biosciences, University of Innsbruck
Innrain 80/82, 6020 Innsbruck (Austria)
E-mail: Klaus.Liedl@uibk.ac.at

[d] Dr. R. G. Huber
Current address: Bioinformatics Institute
Agency for Science, Technology & Research
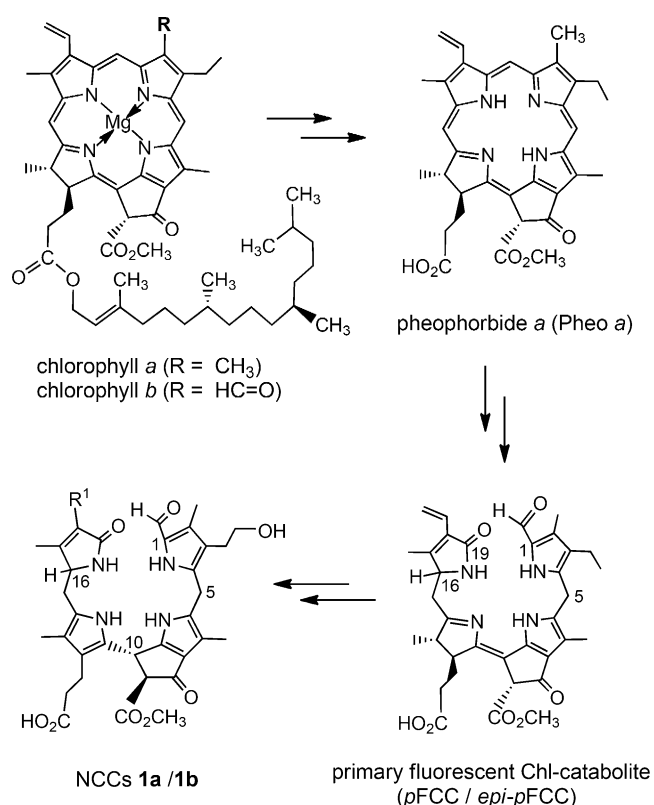30 Biopolis Street, 138671 Singapore (Singapore)

Figure 1. Structural outline of the chlorophyll breakdown to non-fluorescing chlorophyll catabolites (NCCs) via the PaO/phyllobilin pathway.[5] Pheophorbide *a* and primary fluorescing chlorophyll catabolites (*p*FCC/*epi-p*FCC, two C 16-epimers) are general intermediates. NCCs **1 a** (R¹ = CH(OH)CH₂(OH)) and **1 b** (R¹ = CH=CH₂) were first found in leaves of barley (*Hordeum vulgare*)[2] or of the katsura tree (*Cercidiphyllum japonicum*),[11] and were named *Hv*-NCC-1 and *Cj*-NCC-1, respectively.

bilin-type NCCs (DNCCs).[4b, 7] Dioxobilin-type catabolites (such as DNCCs),[7b–d] show a striking structural similarity to bilins, the abundant products of heme breakdown.[8] Furthermore, yellow and pink Chl catabolites (YCCs and PiCCs, respectively) were found as apparent oxidation products of NCCs.[9] In higher plants, Chl-breakdown is, hence, revealed as producing a range of characteristic bilin-type catabolites, named phyllobilins (Figure S1 in the Supporting Information),[4b,c, 5] all arising from the largely common PaO/phyllobilin pathway.[3–5]

Here, we describe an investigation of the Chl catabolites in senescent leaves of the wych elm tree (*Ulmus glabra*), a first representative of the genus *Ulmus* (*Ulmaceae*) to be studied in this respect. In this study, we discovered an NCC with a novel glycosylation topology. Wych elm is a widespread deciduous tree, occurring in mixed native forests in most parts of Europe. It grows as far north as Scotland and Finland, and as far to the south as Greece.[10] The physiology of this Elm tree is of particular interest as it is threatened worldwide by the fungal dutch elm disease, transmitted by bark beetles (*Scolytus sp.*).[10]

Analysis of methanolic extracts of yellow (senescent) fall leaves of a wych elm tree by HPLC allowed the provisional identification of about a dozen colorless, non-fluorescent Chl catabolites (NCCs), as well as several yellow Chl catabolites (YCCs), on the basis of their characteristic UV-absorbance properties.[2, 12] Three of the NCCs were isolated and further structurally characterized (Figure S2 A, B). From 115 g (wet weight) of senescent elm tree leaves, 7.1 mg (8.4 μmol) of *Ug*-NCC-27 (**2**), 12.2 mg (15.1 μmol) of *Ug*-NCC-43 (**3**), and 1.2 mg (1.5 μmol) of *Ug*-NCC-53 (**4**) were obtained, corresponding to about 63 % of the Chl in green leaves (for details see the Supporting Information).

The UV/Vis spectra of the *Ug*-NCCs **2**–**4** displayed an absorbance near 312 nm, similar to those of **1 a** and **1 b** and other natural NCCs,[11, 13] consistent with the presence of an α-formylpyrrole moiety (Supporting Information and Figure S4, Figure S7 therein).[2, 4a] The CD spectra of the *Ug*-NCCs **2**–**4** were also remarkably similar indicating a common configuration at C 10,[4a] as derived for **1 a**, **1 b**[11, 13] and other NCCs occurring in higher plants (Supporting Information and Figure S4).[14] From mass spectra of *Ug*-NCC-27 (**2**) and *Ug*-NCC-43 (**3**), the molecular formulas were deduced as $C_{41}H_{52}N_4O_{15}$ and $C_{41}H_{50}N_4O_{13}$, respectively (Figure S6). Homo- and heteronuclear NMR spectra provided complete signal assignment for the two *Ug*-NCCs, providing the basis for deducing their molecular constitutions (Supporting Information and Table S1). The ^1H NMR spectra of *Ug*-NCC-27 (**2**) and of *Ug*-NCC-43 (**3**) showed signals typical of NCCs; in the spectrum of **3**, the characteristic signals for a vinyl group were present, but they were absent in the spectrum of **2**, in which a t-like signal at 4.56 ppm was consistent with a dihydroxyethyl moiety. The constitutions of **2** and **3** turned out to be the same as those of some NCCs from other plants, for example, *Tc*-NCC-1 and *Tc*-NCC-2 from leaves of the lime tree (*Tilia cordata*).[15] HPLC analysis (with co-injection) of *Ug*-NCC-43 (**3**) and *Tc*-NCC-2 showed a common retention time (Supporting Information and Figure S9). This suggested that these two NCCs are identical, which also implies their common *epi*-type configuration at C 16.[15] In the course of the chloro-

phyll breakdown,[5b] the configuration of the NCCs at C 16 is installed at the stage of their FCC precursors by the enzyme red chlorophyll catabolite reductase (RCCR).[5b, 16] RCCR of *Ulmus glabra* belongs to class-2 RCCRs, which produce catabolites of the so-called *epi*-series.[17]

ESI-MS analysis of *Ug*-NCC-53 (**4**) in the positiveion mode provided a $[M+H]^+$-ion signal at $m/z = 789.20$ consistent with a m/z_{calcd} of $[C_{41}H_{49}N_4O_{12}]^+ = 789.33$, furnishing the molecular formula of $C_{41}H_{48}N_4O_{12}$ (Figure S6). Fragment ions at $m/z = 757.20$ and 666.27 indicated loss of MeOH (from the methyl ester functionality), as well as of C_7H_9NO (ring D), respectively. An ion was conspicuously absent that would correspond to the loss of $C_6H_{10}O_5$ (the sugar moiety), as found in the spectra of **2** and **3** and typical of NCCs glucosylated at the 3^2-position. A 600 MHz ^1H NMR spectrum of a solution of *Ug*-NCC-53 (**4**) in CD_3OD (Figure 2) revealed signals of all 41 carbon-bound pro-
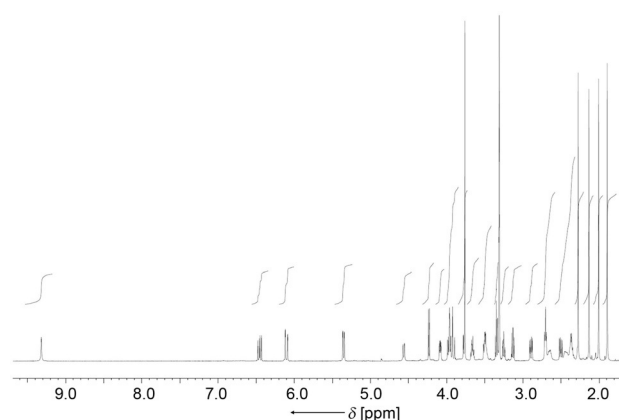

Figure 2. 600 MHz ^1H NMR spectrum of *Ug*-NCC-53 (**4**, CD_3OD, 25 °C), see Table S1 in the Supporting Information for signal assignments.

tons, among them: a singlet (at low field) of the formyl proton, four singlets (at high field) of the four methyl groups attached at the β-pyrrole positions, a singlet at 3.76 ppm (due to a methyl ester function), and the typical signal pattern for a vinyl group around 6 ppm.

$^1H,^1H$ homonuclear correlations from ROESY spectra and $^1H,^{13}C$-heteronuclear correlations from HMBC spectra[18] indicated a glucopyranosyl unit attached at C 3^2 by an oxygen bridge to its anomeric carbon (Figure 3 and Figure S5 in the Supporting Information). The typical downfield shifts of the ^1H and ^{13}C signals for the methylene group $H_2C 3^2$ were also consistent with the peripheral sugar unit. Strikingly, $^1H,^{13}C$-heteronuclear correlations also showed a link of the primary oxygen at C 6′ of the glucopyranosyl moiety to the propionic side chain at ring C. Indeed, a significant shift to high field of the ^{13}C signal of C 12^3, compared to the position of the corresponding carboxyl carbon in NCCs with a free propionic side function (such as the *Ug*-NCCs **2** and **3**), was consistent with the presence of an ester functionality (for details see the Supporting Information, Table S1). The signals of the ^1H-atoms at the C 5′- and C 6′-positions of the sugar moiety of *Ug*-NCC-3 (**4**) were shifted to lower field, when compared to the spectrum of *Ug*-NCC-43 (**3**),
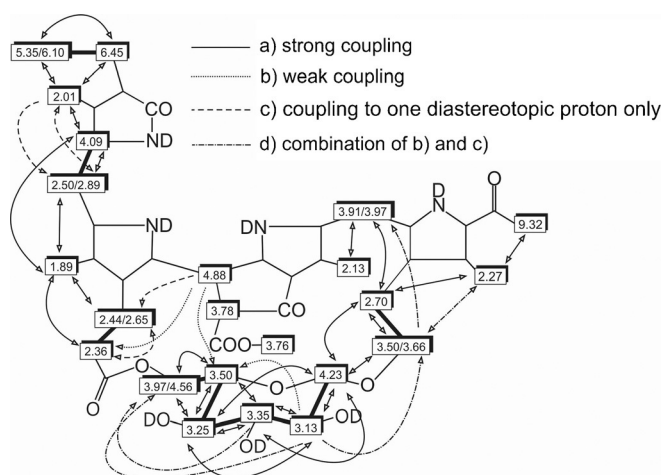
Figure 3. Graphical representations of ^{1}H NMR chemical-shift data of *Ug*-NCC-53 (**4**, in CD$_3$OD, 25 °C) and NOE correlations (from ROESY spectra); bold lines indicate connectivities derived from the COSY spectra; arrows signify NOE correlations, classified as strong (full line) or medium (broken line).

also supporting the presence of an ester linkage at C 6′ of *Ug*-NCC-53 (Table S1). Conformational changes in the propionate chain are the likely cause for high-field shifts of the ^{13}C signals of C 12^1 and C 12^2. The bridging sugar moiety was identified as a β-glucopyranosyl unit by comparing signal positions, ^{1}H,^{1}H-coupling constants, and ^{1}H,^{1}H-NOE correlations in the NMR spectra of **4** with the data of methyl-β-D-glucopyranoside, β-D-glucose (Table S2),[19] and with other NCCs with a peripheral β-glucopyranosyl group at O 3^3, for example, *Bn*-NCC-2[20] or *Tc*-NCC-2.[15] Remarkably, in the ^{1}H NMR spectrum of *Ug*-NCC-53 (**4**) the signal of HC 8^2 (at 3.78 ppm) disappeared only gradually during measurements in CD$_3$OD, indicating slow H/D-exchange at the α-position of the β-keto ester functionality. The proton at C 8^2 coupled with HC 10 (d, $J_{H,H}$ = 3.1 Hz), suggesting a relative *trans*-configuration of HC 8^2 and HC 10, which is typical for stable C 8^2 epimers of NCCs.[14a] An NOE correlation between the protons HC 10 and HC 5′ (of the NCC and glucose moieties, respectively) indicates that the bridging sugar is positioned close to the methine linker between rings B and C of the NCC **4**.

The NCCs *Ug*-NCC-27 (**2**), *Ug*-NCC-43 (**3**), and *Ug*-NCC-53 (**4**) were all glycosylated at the C 3^2-position of ring A, as found in a variety of chlorophyll catabolites from higher plants.[4,5,15] This suggests a broad metabolic significance to this type of functionalization with glucopyranosyl groups in Chl catabolites. Presumably the sugar units are attached to FCCs carrying an OH group at C 3^2, and by enzymes in the cytosol that produce the corresponding glycosylated FCCs as direct NCC precursors (Figure 4).[5b] Indeed, the peripheral attachment of glucopyranosyl groups may be relevant for efficient trans-membrane transport of FCCs and their deposition into the vacuoles, in which the highly stereoselective, non-enzymatic FCC to NCC isomerisation[14a] is presumed to take place.[5b, 21]

Terminal sugar units, as found in the NCCs **2** and **3**, as well as chain-like or branched oligosaccharide moieties in glycosylated bio(macro)-molecules decorate a multitude of natural products.[22] However, in *Ug*-NCC-53 (**4**) the sugar moiety at
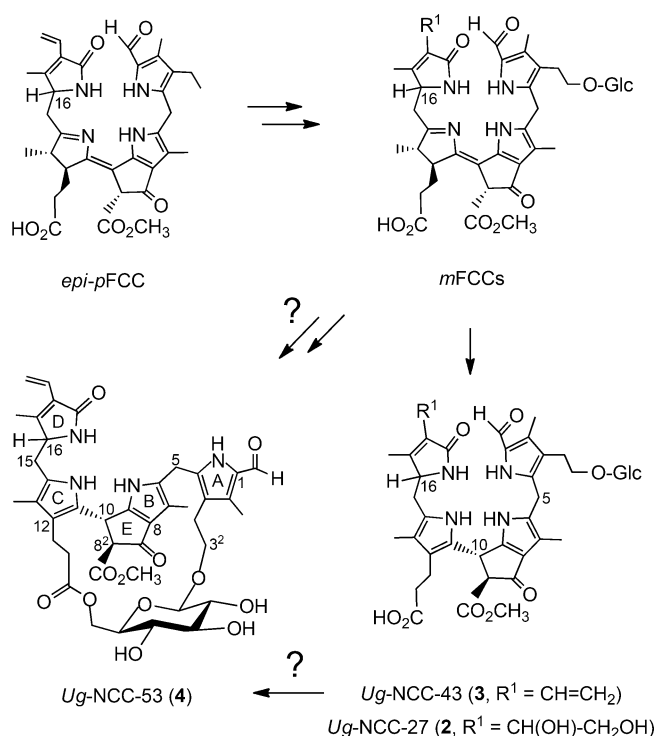

Figure 4. Possible late stages of the chlorophyll breakdown in senescent leaves of the wych elm tree. Hypothetical path from *epi-p*FCC, via modified FCCs (*m*FCCs), to the three *Ug*-NCCs **2**–**4** detected in the elm tree leaves (Glc = β-D-glucopyranose).

ring A is also part of a striking bridging bicyclo[17.3.1]glycoside structure. The implied esterification of the sugar and the propionate group at ring C of this linear tetrapyrrole, clearly reduces the polarity of **4** compared to *Ug*-NCC-43 (**3**). Furthermore, by being part of a 20-membered ring across the pyrrole rings A and C the bridging glycopyranosyl unit also imposes stringent structural restrictions onto the tetrapyrrole core of NCCs, which, otherwise, are conformationally flexible formyloxobilane-type chlorophyll catabolites.

In view of the expected, relevant conformational effect of the unprecedented 1′-6′-glucopyranosyl bridge, the structure of *Ug*-NCC-53 (**4**) was analysed in molecular modelling studies
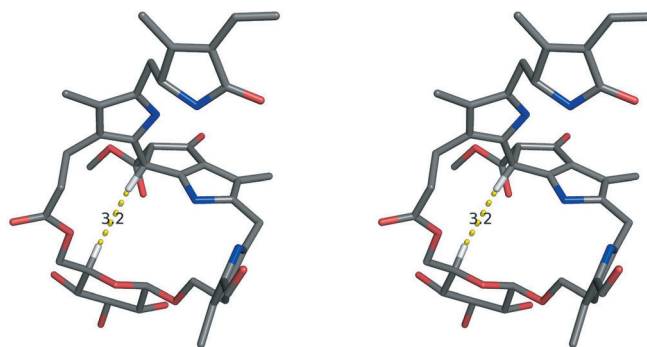

Figure 5. Stereoprojection of a calculated model of the (10 R, 16 R)-isomer of **4**. The calculated distance between HC 10 and HC 5′ of 3.2 Å is highlighted (color code: C gray, O red, N blue, H white).

(Figure 5). In these analyses a 1′-6′-β-ᴅ-glucopyranosyl bridge was used in the chair conformation, and both possible C16-epimers of **4** were taken into consideration, as well as the two (possible) epimers at C10 (which were deduced indirectly to be *R* configured[14a]). Thus, four stereoisomers were modelled with 10*R*/16*R*, 10*R*/16*S*, 10*R*/16*S* and 10*S*/16*S* configurations. Calculated distances were obtained using explicit solvent-molecular-dynamics simulations (see Experimental Section and Figure S11–14 in the Supporting Information). The calculated structure of the 10R,16R-isomer remained stable during the sampling time of 200 ns and displayed a short distance (3.0–3.5 Å) between HC10 and HC5′, consistent with the observation of an NOE correlation between these two H-atoms in **4**. For the 10R,16S-isomer a short distance between HC10 and HC5′ was also calculated for the starting structure, which, however, underwent a reorganization after about 50 ns. In contrast, the calculated distance of HC10 to HC5′ for both 10S-isomers exceeded 7.5 Å. Hence, the modelling studies also supported the independently deduced *R*-configuration at C10 of the glycopyranosyl-bridged NCC **4**.[14a]

To clarify the issue, whether the intriguing motif of a bicyclo-[17.3.1]-glycoside-structure of the catabolite **4** would be a novel structural element in a natural tetrapyrrole, called for a systematic search in the literature. Checking specific organic compounds for novelty in appropriate databases is a routine operation nowadays.[23] In our case, the question of the general topological relevance of a bridging-pyranose moiety necessitated a substructure search for macrocycles related to **4** (bridging a pyranose ring), with particular emphasis on natural products. We utilized both Elsevier Reaxys[24] and Chemical Abstracts Service (CAS) SciFinder.[25]

As the macrocycle, spanned by the bicyclo[17.3.1]glycoside-structure **4**, contains 20 ring atoms, we searched for medium and large rings with 12 to 32 ring members (i.e., covering systems [9.3.1]–[28.3.1]) permitting carbon as well as heteroatom ring members in any combination, with any type of bonds outside the pyranose ring. Other features of **4** are the specific 1,5-bridging of the pyranose ring, the lactone function within the macrocycle, and the presence of heterocyclic rings. Thus, for a necessary structural refinement of search results, stepwise restrictions were envisaged; first to those macrocycles with a properly positioned lactone function (Figure 6), further to compounds with one or more N-atoms in the structure, and finally to those containing a generalized pyrrole-type skeleton (four carbons and one nitrogen, any type of bond, any kind of position/annulation in the ring system). Thus, the number of hits was stepwise reduced from a starting total of 17 692 compound records in *SciFinder* (630 classified as natural products, see Supporting Information and Table S3) and 7849 in *Reaxys* (1 463 natural products) to lactones with one or more nitrogen atoms (413 in *SciFinder*, 321 in *Reaxys*), and finally to only 36 compounds (31 in *SciFinder*, 32 in *Reaxys*) containing an aza-cyclopentyl (pyrrole-type) skeleton.

This first series of searches with increasing restrictions was repeated with macrocyles bridging ᴅ-glucopyranose in the 1,6-position (i.e., bicyclo[*n*.3.1]glycosides as in **4**; Figure 6) or other hexopyranoses: of the 36 aforementioned macrocyles con-
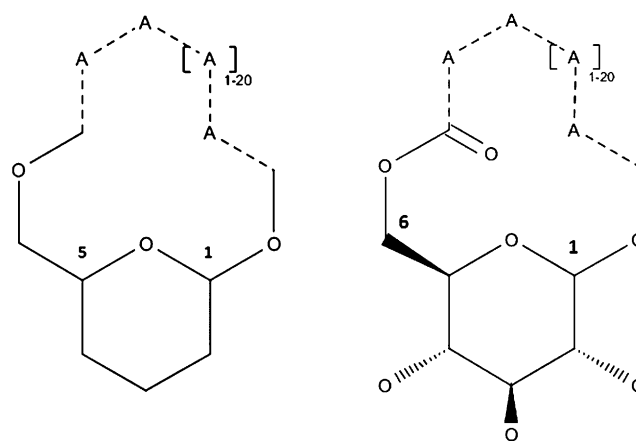

Figure 6. Left: Basic query structure for 1,5-bridged pyranoses (A = any atom except hydrogen, dotted lines: any bond; queries and search results for other bridges are shown in the Supporting Information). Right: Refined query structure for 1,6-bridged ᴅ-glucopyranoses restricted to lactones (see the Supporting Information).

taining a pyrrole-type skeleton, 17 were 1,6-bridged hexopyranoses, the majority (13) being ᴅ-glucopyranoses (see Supporting Information, Tables S4 and S5). A designed glucopyranoside-bridged porphyrin was identified,[26] however, a natural product was not present. Thus, this search indicated that bicyclo[*n*.3.1]glycoside structures bridging five-ring N-heterocycles (related to those in **4**) are unprecedented. However, it furnished about 150 hits for bicyclo[*n*.3.1]glycosides as natural products with bridging by a lactone moiety in the range of organic linkers without N-atoms. Interestingly, among these compounds inhibitors of cell growth[27a] or of high density induced apoptosis[27b] of human cancer cell lines, antibacterial,[28a] antifungal,[28b–d] and antiviral compounds[28e] were found.

A further intriguing question raised by the discovery of the bicycloglycoside structure of *Ug*-NCC-53 (**4**) concerns how and at which stage of Chl-breakdown is **4** generated. As a general rule, modifications reflected by the structures of NCCs are actually incorporated at the stage of FCCs by (mostly) cytosolic enzymes. NCCs with terminal sugar moieties at O3³ have multiple precedence and appear to reflect corresponding glycosylations of their precursor FCCs. In contrast, attachment of a sugar moiety at the propionate group of Chl catabolites by esterification, has only been seen in the persistent "hypermodified" fluorescent Chl catabolites (*hm*FCCs)[4–6] (Figure S1). Note that ester functions at the propionate side chain of FCCs stabilize them against their (acid-induced) isomerization to the corresponding NCCs.[6a, 14b, 29] As a consequence, *hm*FCCs accumulate in the blue luminescent peels of ripening bananas,[6a,c] as well as in de-greened leaves of the banana plant[6b] or of the peace lily, a tropical evergreen.[6d] A similarly esterified FCC was not observed here in the elm tree leaves. Furthermore, a bicycloglycosyl structure, as found in the NCC **4**, imposes stringent structural restraints on the structure of a phyllobilin. Probably, this would interfere with a regular acid-induced isomerization of the FCC to a corresponding NCC.[14a] In the alternative case, in which the bridge motif of **4** was generated at the NCC stage, *Ug*-NCC-43 (**3**) would appear to be a potential biosyn-

thetic precursor of **4** (Figure 4). As NCCs are localized in the vacuole,[5b] such a transformation would be required to take place in this confined environment. However, evidence for the metabolism of an NCC by a vacuolar enzyme is not available.[5b] All in all, the installation of the unique bicycloglycosyl motif of **4** remains puzzling.

Three major non-fluorescent chlorophyll catabolites (NCCs) in naturally de-greened leaves of the wych elm tree (*Ulmus glabra*) were found here to be glycosylated. One of these three 1-formyl-19-oxophyllobilanes, named *Ug*-NCC-53 (**4**), carried an unprecedented bicycloglycosyl structure. This novel ring restrains the flexibility and imposes three-dimensional features on the structure of the NCC **4** (Figure 4), which may be relevant in a still elusive role of this chlorophyll catabolite. This could involve activities related to those found in other bicyclo[*n*.3.1]glycosides (see above) or for example, in defence against bacterial or fungal infections.[30] Clearly, the quest of the physiological relevance of NCCs or of other phyllobilins in senescent leaves and other plant organs[5b, 6, 29] is a challenge to be pursued further.

Experimental Section

Spectroscopy: All plant materials (yellow and green leaves) were collected in November 2010 from a wych elm tree (*Ulmus glabra*) at the main campus of the University of Innsbruck, and were analyzed freshly, or stored at $-80\,°C$ for further use.

UV/Vis: Hitachi U-3000, λ_{max} [nm] (log ε). *^1H- and ^{13}C NMR*: Bruker 600 MHz Avance II+; residual solvent peaks (CD_2HOD: $\delta_H = 3.31$ ppm; $^{13}CD_3OD$: $\delta_C = 49.0$ ppm).[31] *Mass spectrometry*: Finnigan MAT 95 (MS), electrospray ionization (ESI) source, positive ion mode, 1.4 kV spray voltage; Finnigan LCQ Classic (LC-MS), ESI source, positive ion mode, 4.5 kV spray voltage; *m/z* (relative intensity). *HPLC*: An extract from a freshly picked leaf (area of about 15 cm²) of a wych elm tree was analyzed by HPLC (detection at 320 nm, see Supporting Information for further experimental. details, and Figure S2 A, B).

Isolation and structure elucidation of Ug-NCCs: From 115 g of senescent leaves (wet weight), 7.1 mg (8.4 μmol) of *Ug*-NCC-27 (**2**), 12.2 mg (15.1 μmol) of *Ug*-NCC-43 (**3**), and 1.2 mg (1.5 μmol) of *Ug*-NCC-53 (**4**) were isolated, which showed the following spectro-analytical data (see Supporting Information for further details). *Ug*-NCC-27 (**2**): UV/Vis (MeOH, $c = 4.1 \times 10^{-5}$ M): λ_{max} $(\varepsilon_{rel}) = 240$ (1.0), 311 nm (1.0); LC-MS (ESI⁺): *m/z* (%): 879.36 (59, [*M*+K]⁺); 843.39 (14.20), 842.39 (39), 841.42 (100, [*M*+H]⁺, $C_{41}H_{53}N_4O_{15}^+$); *Ug*-NCC-43 (**3**): UV/Vis (MeOH, $c = 6.82 \times 10^{-5}$ M): λ_{max} $(\varepsilon_{rel}) = 248$ (1.00), 313 nm (0.98); LC-MS (ESI⁺): *m/z* (%): 845.43 (48, [*M*+K]⁺); 809.48 (13), 808.48 (55), 807.48 (100, [*M*+H]⁺, $C_{41}H_{51}N_4O_{13}^+$); *Ug*-NCC-53 (**4**): UV/Vis (MeOH, $c = 3.2 \times 10^{-5}$ M): λ_{max} $(\varepsilon_{rel}) = 242$ (sh, 4.39), 312 nm (4.33); MS (ESI⁺): *m/z* (%): 827.20 (50, [*M*+K]⁺); 811.27 (78, [*M*+Na]⁺), 791.20 (27), 790.20 (75), 789.20 (100, [*M*+H]⁺, $C_{41}H_{49}N_4O_{12}^+$).

Molecular modeling: NCC **4** and the C10- and C16-epimeric versions of it were constructed using MOE 2013.08 (Chemical Computing Group Inc., Montreal, QC, Canada). Partial charges were obtained by using the AM1-BCC semi-empirical method,[32] as implemented in the antechamber tool of the AmberTools 13 package.[33] All species were hydrated in octahedral periodic boxes of approximately 3000 TIP3P water molecules.[34] Bond, angle, and torsion potentials were modelled using the generalized AMBER force field

(GAFF) version 1.5.[35] All systems were equilibrated for 100 ns using a van der Waals cut-off of 8.0 Å, particle mesh Ewald electrostatics,[36] a pressure of 1.0 atm by Berendsen weak coupling[37] and a temperature maintained at 300 K by a Langevin thermostat.[38] Shake[39] was enabled on all bonds to hydrogen to allow for a simulation time step of 0.2 fs. Subsequently, 200 ns of sampling were obtained for each system using the GPU implementation of pmemd.[40] One nanosecond running averages of the distances ($H_3C\,8^5$)H 1,2,3–H 5′, HC 10–H 5′, HC 10–$H_{A/B}C\,12^1$, and HC 10–$H_{A/B}C\,12^2$ were computed for the duration of the simulation using "ptraj" from the AmberTools 13 package and are given in the Supporting Information (Figure S11–14).

Literature search: Substructure searches were executed in CAS SciFinder[24] (non-Java structure editor, query structures saved in cxf format; "Explore Substances" – "Chemical Structure" – "Substructure", no restrictions concerning salts, mixtures, isotopes etc.) and Elsevier Reayxs[25] (ChemAxon Marvin Sketch 6.0.6 and earlier versions, query structures saved in mrv format; "Substances,Names,Formulas" – "Structure" – "Substructure on all atoms", no restrictions concerning salts, mixtures, isotopes etc.): final search for the data given here was performed on Jan 10th, 2016 (Reayxs: Version 2.20770.1, last update Jan 7th, 2016; SciFinder: Version December 2015), preceded by earlier searches in April & June 2015, and in June 2014, first exploratory searches in April 2014. In SciFinder, the literature was limited to the CAPLUS database. For details about search strategies, queries, and results see the Supporting Information.

Acknowledgements

[1] a) S. B. Brown, J. D. Houghton, G. A. F. Hendry in *Chlorophylls*, (Ed. H. Scheer), CRC Press, Boca Raton, **1991**, pp. 465–489; b) P. Matile, *Chimia* **1987**, *41*, 376–381.

[2] B. Kräutler, B. Jaun, K. Bortlik, M. Schellenberg, P. Matile, *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1315–1318; *Angew. Chem.* **1991**, *103*, 1354–1357.

[3] B. Kräutler, P. Matile, *Acc. Chem. Res.* **1999**, *32*, 35–43.

[4] a) S. Moser, T. Müller, M. Oberhuber, B. Kräutler, *Eur. J. Org. Chem.* **2009**, *2009*, 21–31; b) B. Kräutler, *Chem. Soc. Rev.* **2014**, *43*, 6227–6238; c) B. Kräutler, *Angew. Chem. Int. Ed.* **2016**, *55*, 4882–4907; *Angew. Chem.* **2016**, *128*, 4964–4990.

[5] a) B. Kräutler, S. Hörtensteiner in *Handbook of Porphyrin Science*, Vol. 28, (Ed. G. C. Ferreira, Kadish, K. M., Smith, K. M., Guilard, R.), World Scientific, New Jersey (USA), **2013**, pp. 117–185; b) S. Hörtensteiner, B. Kräutler, *Biochim. Biophys. Acta Bioenergetics* **2011**, *1807*, 977–988.

[6] a) S. Moser, T. Müller, M.-O. Ebert, S. Jockusch, N. J. Turro, B. Kräutler, *Angew. Chem. Int. Ed.* **2008**, *47*, 8954–8957; *Angew. Chem.* **2008**, *120*, 9087–9091; b) S. Banala, S. Moser, T. Müller, C. Kreutz, A. Holzinger, C. Lütz, B. Kräutler, *Angew. Chem. Int. Ed.* **2010**, *49*, 5174–5177; *Angew. Chem.* **2010**, *122*, 5300–5304; c) S. Moser, T. Müller, A. Holzinger, C. Lütz, S. Jockusch, N. J. Turro, B. Kräutler, *Proc. Natl. Acad. Sci. USA* **2009**,

106, 15538 – 15542; d) B. Kräutler, S. Banala, S. Moser, C. Vergeiner, T. Müller, C. Lütz, A. Holzinger, *FEBS Lett.* **2010**, *584*, 4215 – 4221.

[7] a) F. G. Losey, N. Engel, *J. Biol. Chem.* **2001**, *276*, 8643 – 8647; b) T. Müller, M. Rafelsberger, C. Vergeiner, B. Kräutler, *Angew. Chem. Int. Ed.* **2011**, *50*, 10724 – 10727; *Angew. Chem.* **2011**, *123*, 10912 – 10916; c) B. Christ, I. Süssenbacher, S. Moser, N. Bichsel, A. Egert, T. Müller, B. Kräutler, S. Hörtensteiner, *Plant Cell* **2013**, *25*, 1868 – 1880; d) I. Süssenbacher, S. Hörtensteiner, B. Kräutler, *Angew. Chem. Int. Ed.* **2015**, *54*, 13777 – 13781; *Angew. Chem.* **2015**, *127*, 13981 – 13985.

[8] a) P. R. Ortiz de Montellano, K. Auclair in *The Porphyrin Handbook*, Vol. 12, (Eds.: K. M. Kadish, K. M. Smith, R. Guilard), Academic Press, Amsterdam, **2003**, pp. 183 – 210; b) H. Falk, *Chemistry of Linear Oligopyrroles and Bile Pigments*, Springer, Vienna (Austria), **1989**.

[9] a) S. Moser, M. Ulrich, T. Müller, B. Kräutler, *Photochem. Photobiol. Sci.* **2008**, *7*, 1577 – 1581; b) M. Ulrich, S. Moser, T. Müller, B. Kräutler, *Chem. Eur. J.* **2011**, *17*, 2330 – 2334; c) C. Vergeiner, M. Ulrich, C. Li, X. Liu, T. Müller, B. Kräutler, *Chem. Eur. J.* **2015**, *21*, 136 – 149.

[10] A. P. Moller, *Oikos* **1999**, *85*, 109 – 116.

[11] C. Curty, N. Engel, *Phytochemistry* **1996**, *42*, 1531 – 1536.

[12] B. Kräutler in *The Porphyrin Handbook*, Vol. 13 (Eds.: K. M. Kadish, K. M. Smith, R. Guilard), Elsevier, Amsterdam, **2003**, pp. 183 – 209.

[13] B. Kräutler, B. Jaun, W. Amrein, K. Bortlik, M. Schellenberg, P. Matile, *Plant Physiol. Biochem.* **1992**, *30*, 333 – 346.

[14] a) M. Oberhuber, J. Berghold, K. Breuker, S. Hörtensteiner, B. Kräutler, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6910 – 6915; b) M. Oberhuber, J. Berghold, B. Kräutler, *Angew. Chem. Int. Ed.* **2008**, *47*, 3057 – 3061; *Angew. Chem.* **2008**, *120*, 3100 – 3104.

[15] M. Scherl, T. Müller, B. Kräutler, *Chem. Biodiversity* **2012**, *9*, 2605 – 2617.

[16] S. Rodoni, W. Mühlecker, M. Anderl, B. Kräutler, D. Moser, H. Thomas, P. Matile, S. Hörtensteiner, *Plant Physiol.* **1997**, *115*, 669 – 676.

[17] a) K. L. Wüthrich, L. Bovet, P. E. Hunziker, I. S. Donnison, S. Hörtensteiner, *Plant J.* **2000**, *21*, 189 – 198; b) S. Hörtensteiner, S. Rodoni, M. Schellenberg, F. Vicentini, O. I. Nandi, Y. L. Qui, P. Matile, *Plant Biol.* **2000**, *2*, 63 – 67.

[18] a) R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One & Two Dimensions*, Clarendon Press, Oxford, **1987**; b) H. Kessler, M. Gehrke, C. Griesinger, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 490 – 536; *Angew. Chem.* **1988**, *100*, 507 – 554.

[19] E. Pretsch, P. Bühlmann, C. Affolter, *Structure Determination of Organic Compounds*, Springer, Heidelberg, **2000**.

[20] W. Mühlecker, B. Kräutler, *Plant Physiol. Biochem.* **1996**, *34*, 61 – 75.

[21] P. Matile, S. Hörtensteiner, H. Thomas, B. Kräutler, *Plant Physiol.* **1996**, *112*, 1403 – 1409.

[22] a) A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, M. E. Etzler in *Essentials of Glycobiology*, CSH Press, Cold Spring Harbor, NY, USA, **2009**; b) Y. Funakoshi, T. Suzuki, *Biochim. Biophys. Acta* **2009**, *1790*, 81 – 94.

[23] a) E. Zass, *Heterocycles* **2010**, *82*, 63 – 86; b) J. N. Currano, D. L. Roth, *Chemical Information for Chemists—A Primer*, RSC Publishing, Cambridge, **2014**; c) E. Zass, D. A. Plattner, A. K. Beck, M. Neuburger, *Helv. Chim. Acta* **2002**, *85*, 4012 – 4045; d) D. Seebach, E. Zass, W. B. Schweizer, A. J. Thompson, A. French, B. G. Davis, G. Kyd, I. J. Bruno, *Angew.*

Chem. Int. Ed. **2009**, *48*, 9596 – 9598; *Angew. Chem.* **2009**, *121*, 9774 – 9776; e) D. Seebach, A. K. Beck, S. Capone, G. Deniau, U. Grošelj, E. Zass, *Synthesis* **2009**, 1 – 32.

[24] https://www.reaxys.com/documentation/about query.shtml; https://www.elsevier.com/solutions/reaxys (last accessed Apr. 7th, 2016).

[25] https://scifinder.cas.org; http://www.cas.org/products/scifinder (last accessed April 7th, 2016).

[26] E. Davoust, R. Granet, P. Krausz, V. Carre, M. Guilloton, *Tetrahedron Lett. Tetrah. Lett.* **1999**, *40*, 2513 – 2516.

[27] a) T. Nagano, J. Pospisil, G. Chollet, S. Schulthoff, V. Hickmann, E. Moulin, J. Herrmann, R. Müller, A. Fürstner, *Chem. Eur. J.* **2009**, *15*, 9697 – 9706; b) M. Gavagnin, M. Carbone, P. Amodeo, E. Mollo, R. M. Vitale, V. Roussis, G. Cimino, *J. Org. Chem.* **2007**, *72*, 5625 – 5630.

[28] a) R. T. Desmond, A. N. Magpusao, C. Lorenc, J. B. Priestley, M. W. Peczuh, *Beilstein J. Org. Chem.* **2014**, *10*, 2215 – 2221; b) B. Ruttens, P. Blom, S. Van Hoof, I. Hubrecht, J. Van Der Eycken, B. Sas, J. Van Hemel, J. Vandenkerckhove, *J. Org. Chem.* **2007**, *72*, 5514 – 5522; c) B. Fogliani, P. Raharivelomanana, J.-P. Bianchini, S. Bouraima-Madjebi, E. Hnawia, *Phytochemistry* **2005**, *66*, 241 – 247; d) X.-C. Li, A. S. Joshi, H. N. ElSohly, S. I. Khan, M. R. Jacob, Z. Zhang, I. A. Khan, D. Ferreira, L. A. Walker, S. E. Broedel, R. E. Raulli, R. L. Cihlar, *Nat. Prod.* **2002**, *65*, 1909 – 1914; e) H.-X. Xu, H. Wan, H. Dong, P. P. But, L. Y. Foo, *Biol. Pharm. Bull.* **2000**, *23*, 1072 – 1076.

[29] S. Moser, T. Müller, A. Holzinger, C. Lütz, B. Kräutler, *Chem. Eur. J.* **2012**, *18*, 10873 – 10885.

[30] a) S. G. Krattinger, E. S. Lagudah, W. Spielmeyer, R. P. Singh, J. Huerta-Espino, H. McFadden, E. Bossolini, L. L. Selter, B. Keller, *Science* **2009**, *323*, 1360 – 1363; b) N. Yao, J. T. Greenberg, *Plant Cell* **2006**, *18*, 397 – 411.

[31] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512 – 7515.

[32] A. Jakalian, B. L. Bush, D. B. Jack, C. I. Bayly, *J. Comput. Chem.* **2000**, *21*, 132 – 146.

[33] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668 – 1688.

[34] M. W. Mahoney, W. L. Jorgensen, *J. Chem. Phys.* **2000**, *112*, 8910 – 8922.

[35] J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157 – 1174.

[36] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089 – 10092.

[37] H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684 – 3690.

[38] S. E. Feller, Y. H. Zhang, R. W. Pastor, B. R. Brooks, *J. Chem. Phys.* **1995**, *103*, 4613 – 4621.

[39] H. C. Andersen, *J. Comp. Physiol.* **1983**, *52*, 24 – 34.

[40] R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **2013**, *9*, 3878 – 3888.