



**HELMHOLTZ
ZENTRUM FÜR
INFEKTIONSFORSCHUNG**

This is a pre- or post-print of an article published in

Kayser, H., Wray, V., Nitz, M.

Structure of a novel farnesylated bilin from an insect -

**Formation by α -cleavage of heme A of mitochondrial
cytochrome c oxidases?**

(2014) FEBS Journal, 281 (10), pp. 2366-2376.

Structure of a novel farnesylated bilin from insect: formation by α -cleavage of heme A of mitochondrial cytochrome c oxidases?

Hartmut Kayser^{a1}, Victor Wray^b, Manfred Nimtz^b

^aInstitut für Allgemeine Zoologie und Endokrinologie, Universität Ulm, Helmholtzstrasse 8/1,
89081 Ulm, Germany

^bHelmholtz-Zentrum für Infektionsforschung GmbH, Molekulare Strukturbiologie,
Inhoffenstrasse 7, 38124 Braunschweig, Germany

¹To whom correspondence should be addressed.

E-mail: hartmut.kayser@uni-ulm.de

Tel.: +49 731 50 15110

Fax: + 49 731 50 32609

E-mail addresses of co-authors:

victor.wray@helmholtz-hzi.de

manfred.nimtz@helmholtz-hzi.de

The authors declare no conflict of interest.

Key words: bilins; structure elucidation; heme A; mitochondria; insects.

Running title: Structure of a novel farnesylated bilin

Abstract

Biliproteins are present in almost all forms of life with many of them playing vital roles in photobiology. The bilin ligand of a recently characterised 500 kDa-biliprotein from an insect has been isolated and its structure elucidated employing chemical and spectroscopic techniques (UV-VIS, IR, MS, NMR, CD). This blue pigment, named CV-bilin, represents a unique high-molecular weight derivative of biliverdin IX α with unusual 10*E* configuration and a molecular mass of 852 corresponding to C₄₈H₆₀N₄O₁₀. The high mass of this open-chain tetrapyrrole is due to the presence of an epoxi-dihydroxyethylfarnesyl substituent at C-18 and a hydroxymethyl substituent at C-13. This substitution pattern exactly reflects that of heme A of mitochondrial cytochrome c oxidases with a hydroxyethylfarnesyl chain and a formyl group at corresponding positions of the cyclic tetrapyrrole. Since no other natural product is known to show these structural features (heme O, the precursor of heme A, has a methyl group at C-13) this bilin is presumed to be derived from heme A by cleavage of the α -methine bridge and oxidative modifications at C-13 and the hydroxyethylfarnesyl chain. Possibly, a bilin structurally related to this insect bilin is produced also in other organisms as a result of mitochondrial turnover or degradation. Since CV-bilin in complex with a specific protein is accumulated at the end of the larval life, stored in the pupa and finally transferred to the oocytes, a possible role of the free or protein-bound pigment in egg or embryonic development is discussed.

Introduction

Bilins are linear tetrapyrrole pigments occurring in almost all forms of life. In vertebrates, bilins or bile pigments are known as degradation products of hemoglobin that are excreted. In higher plants, fungi, red algae and cyanobacteria, bilins act as sensors for light-controlled developmental processes and as light-harvesting accessory pigments in photosynthesis, where they are covalently bound to specific proteins as phytochromes and phycobiliproteins, respectively [1]. These bilins are derived from the IX α isomer of biliverdin by structural changes that define the spectral range of light absorption [2].

Bilins are also widespread among insects; their possible roles, however, are largely unknown apart from being pigments that contribute to body coloration, signalling, camouflage and similar vital functions [3]. Insects produce either biliverdin IX α or IX γ , depending on taxonomy, with the γ -isomer almost exclusively found in lepidoptera. These pigments are present as biliproteins in diverse tissues and may be produced in larvae as well as adults. In all examined insects, bilins are non-covalently associated with protein; the crystal structures of two lepidopteran biliproteins have been solved [4,5]. The specific mode of association induces high chirality of the non-chiral bilin chromophores indicating enantioselectivity of binding of the cyclic helical conformation to the respective protein [4,6].

Recently a new type of high-molecular-weight biliprotein has been isolated from the moth *Cerura vinula* and was characterized biochemically and by molecular imaging [7]. The apoprotein was identified as an arylphorin, representing a conserved hexameric protein family that is widespread among arthropods. In the present study, we have isolated the associated bilin, referred to here as CV-bilin, and solved its structure by employing chemical, biochemical and spectroscopic techniques, including multidimensional NMR. The bilin was shown to be unusual as: (i) it belongs to the IX α series of biliverdin isomers, not IX γ that is typical for lepidoptera,

and (ii) its structure is novel with a unique hydroxyethylfarnesyl substituent and a hydroxymethyl group. Type and positions of these substituents exactly reflect those of heme A, an essential redox-cofactor of mitochondrial cytochrome c oxidases [8]. These features strongly suggest that CV-bilin originates from heme A by cleavage of the α -methine bridge and oxidative modifications, mainly at the hydroxyethylfarnesyl chain. Based on structural similarity it is speculated that CV-bilin may play a juvenile-hormone like role in early development.

Results

UV-VIS and IR spectroscopy

Compared to classical biliverdin isomers the UV-VIS spectrum of CV-bilin displayed a bilatriene spectrum that was blue-shifted by 20-50 nm, depending on solvent (Table 1; Fig. 1-A1). Furthermore, its UV/VIS ratios are lower. These features may indicate a chromophore shorter by one conjugated double bond and/or an unusual geometry. The IR spectrum (Fig. 2) was characterized by strong peaks at 2956 cm^{-1} , 2927 cm^{-1} and 2857 cm^{-1} , assigned to C-H stretching, and a strong broad band at 3385 cm^{-1} , assigned to O-H stretching. Hence, CV-bilin seems to be hydroxylated and linked to an aliphatic chain, a feature not yet known for bilins.

Mass spectroscopy

High-resolution mass spectra of the free CV-bilin revealed a molecular mass of 852, corresponding to $\text{C}_{48}\text{H}_{60}\text{N}_4\text{O}_{10}$. Prominent molecular ions were at m/z 851.4228 (calc. 851.4231 for $[\text{M}-\text{H}]^+$) and at 875.4202 (calc. 875.4202 for $[\text{M}+\text{Na}]^+$) (Fig. 3). Losses of three molecules of water were observed, but no loss of the presumed aliphatic chain; hence its linkage by an ester bond was very unlikely. In accordance with this, CV-bilin was stable under conditions of saponification as shown by mass spectrometry and HPLC. Comparison of the molecular composition of CV-bilin with that of the biliverdin isomers ($\text{C}_{33}\text{H}_{34}\text{N}_4\text{O}_6$) revealed a difference

of $C_{15}H_{26}O_4$, suggesting the presence of a sesquiterpenoid-based substituent and four additional oxygen atoms.

Due to the presence of several oxygen atoms, CV-bilin behaved more polar than the biliverdin isomers: on RP-HPLC its retention time was 12.3 min compared to 14.6 min for biliverdin IX α , for example. This behavior of the free bilins was much more pronounced on silica gel TLC (data not shown).

Ester formation

The mass spectrum of the major methylation product of CV-bilin provided a molecular mass of 880, corresponding to $C_{50}H_{64}N_4O_{10}$, as expected for a dimethyl ester: molecular ions were present at m/z 879.45437 (calc. 879.45442) for $[M-H]^+$, and at m/z 903.45068 (calc. 903.45201) for $[M+Na]^+$. A second methylation product, which was less polar on RP-HPLC, was present in some samples: its molecular mass of 908, corresponding to $C_{51}H_{64}N_4O_{11}$, was indicated at m/z 907.44864 (calc. 907.44933 for $[M-H]^+$) and m/z 931.44575 (calc. 931.44692 for $[M+Na]^+$), respectively. Based on results presented below, this second product represented an artifact originating from oxidation of a hydroxyl to carboxylic acid followed by methylation to produce a trimethyl ester.

Acetylation

Acetylation of CV-bilin yielded at least four products less polar than the original compound on RP-HPLC (Fig. 4). These were identified as: (a) diacetate methyl ether (r.t. 18.4 min): molecular mass 950, corresponding to $C_{53}H_{66}N_4O_{12}$ (found m/z 973.457; calc. 973.457 for $[M+Na]^+$); (b) diacetate (r.t. 17.4 min): molecular mass 936, corresponding to $C_{52}H_{64}N_4O_{12}$; found m/z 959.441; calc. 959.442 for $[M+Na]^+$; and (c) monoacetate methyl ether (r.t. 16.0 min): molecular mass 908, corresponding to $C_{51}H_{64}N_4O_{11}$ (found m/z 931.447; calc. 931.447 for $[M+Na]^+$). No useful mass spectrum was obtained from the most polar fraction (product d; r.t. 14.5 min) presumably

representing a monoacetate. In a different preparation another product was tentatively identified as diacetate carboxylic acid: molecular mass 950, corresponding to $C_{52}H_{62}N_4O_{13}$ (found m/z 951.44; calc. 951.439 for $[M+H]^+$). Taken together, these derivatives of CV-bilin revealed, in accordance with results from mass spectrometry, the presence of three hydroxyl groups with different reactivity yielding a monoacetate and a diacetate that in addition become readily methylated, and that one of the hydroxyl groups is easily oxidized to the corresponding acid (see above),.

Degradation

Information of the tetrapyrrole skeleton was obtained by microchemical degradation of CV-bilin, both unlabeled and [^{14}C -ALA]-labeled (Fig. 5). Treatment of the free bilin with chromic acid provided methyl vinyl maleimide (R_f value of 0.55) (less in quantity in comparison to biliverdin IX α) and a polar product zone close to the application site. This polar product was the only one when chromate was used. Degradation of the dimethyl ester of CV-bilin with chromate yielded a pyrrole dialdehyde methyl ester that was also obtained from the dimethyl ester of biliverdin IX α . In addition to this known dialdehyde, a second dialdehyde, unique to CV-bilin dimethyl ester, was obtained, which was more polar (R_f values of 0.24 vs. 0.38). The two dialdehydes were formed in about equal quantities, as estimated from the intensity of staining and radioactivity (Fig. 5), respectively. For comparison, chromate degradation of the γ -isomer of biliverdin dimethyl ester yielded methyl vinyl pyrrole dialdehyde as the sole product as expected. Treatment of CV-bilin dimethyl ester with chromic acid yielded, in addition to methyl vinyl maleimide and hematinic acid methyl ester, a more polar product presumably arising from oxidation of the new pyrrol dialdehyde ester to the corresponding maleimide. Taken together, these studies revealed that CV-bilin possesses the structural skeleton of biliverdin IX α with one of the inner rings substituted by a polar group, most likely a hydroxyl group, as suggested from the IR and mass spectra.

NMR spectroscopy

The full structure of CV-bilin was elucidated using multidimensional ^1H NMR and ^{13}C NMR spectroscopy on a 600 MHz instrument with cryoprobe head. The ^1H 2D COSY and TOCSY readily identified the nature of the substituents present in the tetrapyrrole unit and their positions were established from through-bond correlations in the 2D ^1H - ^{13}C HMBC spectrum and through-space correlations in the 2D ^1H - ^1H ROESY spectrum (Figs. 6 and S1 - S5). These spectra allowed a full assignment of data for this substructure given in Table S1. The nature of the unusual hydroxymethyl substituent at C-13 was evident from the integral in the ^1H spectrum and the unusually high-field ^{13}C shift (54.3 ppm) and low field ^1H shift (4.60 ppm) that is characteristic of such a substituent in a pyrrole ring system. Its position at C-13 follows from its NOEs to the methylene groups H-12^I and H-12^{II}.

The configuration of the inter-pyrrole linkages follows from the NOEs observed to H-5, 10 and H-15 of this unit. Thus H-5 shows through-space interactions with both the olefinic side chain at C-3 and the methyl group at C-7. Similarly C-15 interacts with the side chains at C-13 and C-17. In contrast H-10 only shows correlations to the side-chain at C-8, but not C-12, indicating the *E*-configuration of the respective double bond.

The same set of spectra also established the structure of the side-chain attached to C-18 of the tetrapyrrole (Table S1). Long-range ^1H - ^{13}C correlations and NOEs unambiguously identified the position of the initial unit C1'-C3' of the side chain at C-18 of the tetrapyrrole, the central unit C4' to C-10', and terminal unit C-11' to C-14'. Integration of the signal at 3.94 ppm in the ^1H spectrum of the latter and inspection of the HMBC spectrum indicated this belonged to a free hydroxymethyl group as it did not show any long-range ^1H - ^{13}C correlations compatible with an ether or ester linkage. The presence of only one further oxygen atom in the side-chain follows from the structural formula determined from the accurate molecular mass and implies the presence of an epoxide to satisfy the ^{13}C chemical shifts of C-8' and C-9'. This is confirmed by

the high field shift of H-8' characteristic of such a system. The orientation of the central unit followed from long-range ^1H - ^{13}C correlations of the methylene protons H-2 with the olefinic carbon C-4'. The *E*-configuration of the double bonds at C-4' and C-12' and the relative stereochemistry of the epoxide ring followed from the NOE data. The absolute stereochemistry of C-1' of the side chain was not determined.

Thus in summary, the combined chemical and spectroscopic data are only compatible with the structure shown in Figure 7A, indicating that the molecule is assembled from modified farnesol and tetrapyrrole units. The structure suggests that the hydroxymethyl group of farnesol has been added to the vinyl group at C-18 of the tetrapyrrole thus creating an asymmetric center at C-1' of the hydroxyethylfarnesyl side-chain. The structure is in accordance with all features of CV-bilin described above: the shortened chromophore with *10E* configuration, the presence of an aliphatic chain and of three hydroxyl groups, one of which is located on one of the inner pyrrole rings giving rise to the new polar pyrrole degradation products.

A second, minor bilin component, which was more polar than the main component on RP-HPLC, was isolated showing data very similar to those of the main component. From the high-resolution ESI MS it had a molecular mass of 866 with the molecular formula $\text{C}_{48}\text{H}_{58}\text{N}_4\text{O}_{11}$. In contrast to the main component that readily lost three molecules of water in the ESI MS the minor component lost only two molecules of water. The ^1H NMR spectrum of a mixture of this derivative and the main bilin component showed decreased intensity for the signals corresponding to H-14' (loss of CH_2OH group) and chemical shifts of H-13'¹ and H-12'. The combined data indicated the derivative was closely related to the major compound and possessed a carboxylic acid residue at C-14' of the ethylfarnesyl side-chain (Fig. 7B). Formation of this minor component was found to arise as an artifact through oxidation of the terminal hydroxyl group. As described above, the corresponding trimethyl ester, less polar than the dimethyl ester, has been identified by MS of the methylation products separated by RP-HPLC.

CD spectroscopy

Free CV-bilin displayed little optical activity. In methanol, the CD spectrum showed a weak negative band in the visible region at 625 nm and a weak positive band in the Soret region at 370 nm (Fig. 1-A2). The CD spectrum in HCl/methanol was comparable (not shown). By contrast, the native protein complex of CV-bilin with its distinct VIS peak (Fig. 1-B1), displayed high optical activity of opposite sign (Fig. 1-B2) indicating that the chromophore was locked by the chiral peptide environment in an enantioselective way. Upon denaturation of the biliprotein with urea the CD spectrum changed to that of the free bilin (not shown).

Developmental aspects

CV-bilin, as a protein complex, is present in high concentration in the hemolymph of final instar larvae. Its concentration was found to increase from day 1 of this instar. The bilin was also identified by MS (m/z 875 $[M+Na]^+$) in the bluish fat bodies of males prior to adult emergence. In females, the oocytes contained intensely green coloured material suggesting the presence of large amounts of bilin; its identification by MS, however, was hampered by excess contaminating material. The presumed bilin is also present in deposited eggs.

Discussion

CV-bilin with a modified hydroxyethylfarnesyl substituent in particular and a hydroxymethyl substituent represents the most unusual linear tetrapyrrole isolated from nature to date. There are only two other molecules, cyclic tetrapyrroles, with a (unmodified) hydroxyethylfarnesyl substituent at the corresponding position as in CV-bilin: heme A and heme O, specific and essential cofactors of cytochrome c oxidases [9,10]. These hemes are biosynthesized from heme B first by the transfer of a farnesyl group to the vinyl group at C-18 to give heme O followed by

a two-step oxidation of the methyl group at C-13 to create a formyl group, which is unique to heme A (Fig. 8B) [8].

Bilins, as open-chain tetrapyrroles, originate by oxidative cleavage of the heme ring, catalysed by heme oxygenases in an evolutionary conserved mode, at one of the four methine bridges thus giving rise to four possible bilin isomers [11]. In animals, plants and microorganism, cleavage mostly occurs at the α -bridge thus generating biliverdin IX α . Since CV-bilin is an unusually modified bilin exactly reflecting the substitution pattern of heme A (Fig. 8), it is very likely that this bilin originates from cleavage of the α -methine bridge of heme A. As there is no obvious high biosynthetic demand for mitochondrial cytochromes at the developmental stage, when CV-bilin is accumulated, cleavage of heme A may immediately follow its synthesis via heme B. If so, CV-bilin would be the final product of a direct route of synthesis of open-chain tetrapyrroles. This view is supported by studies of another lepidopteran insect, where radiolabel from the tetrapyrrole precursor 5-ALA was incorporated into biliverdin IX γ without a lag phase [12]. Synthesis of CV-bilin and its carrier protein are assumed to be tightly coupled processes taking place in the larval fat body as the major source of metabolic products. In this tissue, synthesis of CV-bilin may be independent of the synthesis of mitochondrial cytochromes.

Five structural modifications are required to convert heme A to CV-bilin: (1) heme cleavage at the α -bridge; (2) reduction of the formyl group to a hydroxymethyl group; (3) epoxidation at the C-8'/C-9' double bond of the hydroxyethylfarnesyl chain; (4) hydroxylation of one of the terminal methyl groups of the hydroxyethylfarnesyl chain resulting in the hydroxymethyl group at C-13', and (5) conversion of 10Z to 10E configuration. In general, heme cleavage results in a biliverdin isomer with all-*syn-Z* configuration as the most stable form. CV-bilin, however, is 10E, an unusual feature possibly forced by the presence of the bulky ethylfarnesyl substituent. The sequence of these modifications leading from heme A to CV-bilin remains to be elucidated as well as the first product associating with the protein; possibly, some or even all steps take

place in the biliprotein complex. Interestingly, the two oxidative modifications of the side-chain of CV-bilin are not unusual and well documented for farnesol as the precursor of insect juvenile hormones [13]. While an epoxide at C-10/C-11 of farnesol (corresponding to C-12'/C-13' in CV-bilin) is a structural feature common to all juvenile hormones, a second epoxide is present at C-6/C-7 (corresponding to C-8'/C-9' in CV-bilin) in juvenile hormone bisepoxide from higher diptera [14], and at C-2/C-3 (corresponding to C-4'/C-5' in CV-bilin) in the juvenile hormone skipped bisepoxide isolated from a bug [15]. Furthermore, hydroxylations of methyl groups of farnesol at C-4, C-8 and C-12 have been reported [16,17], and a cytochrome P450 dependent monohydroxylase from hormone-producing tissue has been shown to hydroxylate *in vitro* several sesquiterpenoids, including farnesol, at C-12 [18]; this corresponds to formation of the hydroxymethyl group at C-13' of the hydroxyethylfarnesyl substituent of CV-bilin.

Taken together, there is a close structural similarity between the modified farnesyl chain of CV-bilin and juvenile hormones. There is one striking difference, however: in CV-bilin farnesol is (formally) added through its hydroxyl group at C-1 to the vinyl group of the tetrapyrrole, while in juvenile hormones this hydroxyl is oxidized and converted to the methyl ester, which (in addition to the epoxide) is relevant for hormonal activity. On the other hand, evidence has been accumulated that juvenile hormone acid may also act as a hormone exerting specific roles [19]. In CV-bilin, the free terminal allylic hydroxyl group at C-14' of the side-chain is easily oxidized to the carboxylic acid level, as we show here, thus creating similarity to juvenile hormone acid. As discussed above, the position of the epoxide is not crucial for hormonal activity as it varies depending on insect species and/or suggested roles [13].

The farnesyl chain of CV-bilin is covalently linked via a C-C bond to the tetrapyrrole and hence very unlikely to become liberated to eventually act like a juvenile hormone. On the other hand, the farnesyl chain with its free terminus that may carry a carboxyl group (at C-13'), possibly in methylated form, and the proximal epoxide might be recognized as a juvenile hormone by a

receptor protein in a way very recently shown for the binding of synthetic juvenile hormone analogues to the putative receptor, Methoprene-tolerant [20]. While this is a purely speculative view it may stimulate ideas, however, to uncover any developmental role(s) of CV-bilin. The idea of a role in cellular regulation receives support from studies of *Xenopus laevis* suggesting that biliverdin IX α , presumably in complex with a protein, exerts a critical function in embryonic development [21], and by the isolation of a bilirubin-albumin complex from rats and humans that stimulates cell growth and tissue regeneration [22]. In any case, the high production of CV-bilin at the end of larval development requires massive allocation of metabolic resources and energy to heme synthesis, subsequent heme cleavage, several oxidation steps and association with a specific protein (which is also produced in high quantity) followed by uptake of the biliprotein from the hemolymph into the fat body in the pupa and finally the transfer into developing oocytes. This high investment strongly argues for an essential role of free CV-bilin or its protein complex in egg formation or embryonic development. Such a role may not be unique to *C. vinula* as this unusual bilin has recently also been identified in other phylogenetically related moth species (Kayser & Nimtz, unpublished data).

The addition of farnesol to the vinyl group creates a chiral centre at C-1' of the resulting hydroxyethylfarnesyl substituent. The absolute stereochemistry is unknown and may not be determined for free CV-bilin in solution as this chiral centre is susceptible to racemisation, as shown for heme A [9]. In its natural complex with cytochrome c oxidase, however, the absolute configuration has been solved to be *S* by X-ray structure analysis [23]. In analogy to heme A, the overall conformation of CV-bilin and the stereochemistry at C-1' are thought to be determined by interaction with the protein. The methyl ether formation, observed during acetylation, most probably takes place at the hydroxyl group at C-1' of the hydroxyethylfarnesyl substituent of CV-bilin, as it has been shown for heme A, its presumed precursor [9].

The basis for the weak CD spectrum displayed by the isolated CV-bilin and after denaturation of its protein complex is unknown. As the chromophore *per se* has no chiral centre it must be assumed that some structural features like the chirality at C-1' and the farnesyl side-chain act together to slightly shift the isoenergetic equilibrium of the *M*- and *P* conformations of the bilin in favour of the left-handed *M* enantiomer [24]. Whether the low UV-VIS peak ratio of free CV-bilin, compared to biliverdin IX α , is due to its 10*E* geometry or to some extended conformation of the chromophore is also not clear (cf. [25]). The weak optical activity of free bilin is in contrast to the native protein-bound CV-bilin, which displays high optical activity of opposite sign meaning a high preference of binding for the *P* enantiomer. The sharp VIS peak of the biliprotein may indicate a protonated and partially extended chromophore, which is presumed to be 10*E* due to the obvious high stability of this geometry. Sound answers to all these speculative views will only be possible once the crystal structure of CV-biliprotein has been solved. A comparable situation has been described for the bilin-binding protein of *P. brassicae*, whose X-ray structure allowed for the first time the correlation of the CD sign with the helical conformation of the bilin ligand [4]. Here, the right-handed *P* cyclic helical enantiomer of biliverdin IX γ is selectively locked and the VIS peak is rather broad, not sharp as in case of CV-biliprotein, indicating that the geometries of the two bilins are different in their protein complexes.

In conclusion, we have isolated and identified a novel bilin with a structure that strongly points to heme A as a presumed precursor. As heme A is a functional key component of mitochondrial cytochrome c oxidases their degradation is expected to produce a farnesylated bilin structurally related to CV-bilin in all cells performing mitochondrial respiration. Transformation of heme A into bilins like CV-bilin would then appear as a way of reuse of costly products that may even function in a different context. At present, however, nothing is known about the degradation of heme A (and heme O, its precursor) in biological systems [26].

Experimental Procedures

Chemicals

If not stated otherwise chemicals were purchased from Merck (Darmstadt, Germany) and Sigma-Aldrich (Steinheim, Germany) at highest available purity. Precoated silica gel plates were from Merck. [4-¹⁴C]-5-Aminolevulinic acid (5-ALA) hydrochloride (specific activity 50.1 mCi/mmol) was obtained from DuPont NEN Products (Bad Homburg, Germany).

Bilin isolation

The blue biliprotein was purified from hemolymph of final instar larvae of the puss moth, *C. vinula* L. (Notodontidae), as described [7]. The non-covalently bound bilin was liberated from the protein by the addition of two volumes of methanol. After storage at 5 °C and centrifugation (2000g for 10 min) the bilin was isolated from the supernatant using either solid-phase or liquid-phase extraction, as follows. Method 1: water was added to dilute methanol to <25%, the bluish solution was passed over a Waters Sep-Pak C18 (36 mg) cartridge and, after washing with 0.5 mL of 25% methanol/water (v/v), the bilin was eluted with 100% methanol. Method 2: chloroform was added and, after mixing, water was added to obtain phase separation with the bilin retained in the aqueous phase. The chloroform phase, containing lipid, was discarded, and extraction with chloroform was repeated. A small amount of solid citric acid was added together with chloroform to transfer the bilin to the chloroform phase by vigorous shaking. The volumes of the bilin solutions in methanol and chloroform, respectively, were reduced by using a vacuum-centrifuge (Christ, model Alpha RVC) operated at 45 °C.

Chemical reactions

Microchemical degradation of bilins by chromic acid and chromate, respectively, on thin-layer plates was performed according to [27]. Methyl esters were prepared with 5 % HCl in methanol

(v/v), or with the BOP reagent [28,29]. Acetylation was performed with acetic acid anhydride in pyridine (1:3; v/v). Saponification was done in 0.1 N NaOH in methanol for 3 h at 65 °C followed for 12 h at room temperature. All other operations were carried out at room temperature.

Preparation of biliverdin isomers

Biliverdin isomers were prepared as their dimethyl esters by coupled oxidation of hemin and separated by thin-layer chromatography (TLC) on silica gel [30]. Their identity was confirmed by MS and oxidative degradation [27]. Biliverdin IX γ was also isolated from the wings of the butterfly *P. brassicae* [31]. Commercial biliverdin IX α hydrochloride was used as a reference.

Chromatography

For reverse-phase high performance liquid chromatography (RP-HPLC) of free and derivatised bilins, a Shimadzu gradient system was used consisting of two LC-6A pumps, a SPD-6AV UV-VIS spectrophotometric detector and a SCL-6A system controller. A Macherey-Nagel Nucleosil C18 column (particle size 5 μ m; 120 mm x 5 mm i.d.) was used at room temperature with an acetonitrile/water gradient (25 % to 80 % acetonitrile in 25 min) in the presence of 0.1% trifluoroacetic acid as mobile phase; the flow rate was 1.8 mL/min. Elution was monitored at 373 nm. Thin-layer chromatography (TLC) of bilin derivatives was performed on silica gel 60 F₂₅₄ plastic sheets or glass layers (4 x 8 cm and 5 x 10 cm, respectively; Merck).

Spectroscopy

UV-VIS spectra were recorded on a Perkin Elmer Lambda 15 spectrophotometer. IR spectra were measured as a solid film on a Bruker Hyperion microscope coupled with a VERTEX 70 Fourier Transform Infrared (FTIR) spectrometer. Circular dichroism (CD) spectra were obtained with an Applied Photophysics Chirascan photometer. High resolution and high accuracy mass spectra (MS) were acquired on a Bruker APEXIII Fourier Transform Mass Spectrometer

equipped with an electrospray ion source operated in both positive and negative ion modes. High resolution ESI mass spectra were also recorded on a Thermo Science LTQ Orbitrap mass spectrometer. NMR were recorded in 1D (^1H) and 2D (COSY, TOCSY; ROESY, HMQC and HMBC) at 300 K on Bruker AVANCE DMX 600 or AVANCE III with cryoprobe head NMR spectrometers locked to the major deuterium resonance of the solvent, methanol- d_4 or methanol- d_4 containing a small amount of benzene- d_6 . Chemical shifts were referenced to the residual proton signal of the solvent (^1H : 3.35 and ^{13}C : 49.0 ppm).

Biosynthetic labelling

[4- ^{14}C]-5-Aminolevulinic acid (5-ALA) was injected into final instar larvae at a dose of $2 \times 0.1 \mu\text{Ci}$ per insect on two successive days; hemolymph was collected two days thereafter. The biliprotein was purified using a micro-scale modification of the standard procedure. The bilin was isolated using Method 1, and an aliquot was methylated. [^{14}C -5-ALA]-Labeled biliverdin IX γ was prepared from developing wings of *P. brassicae* [31]. [^{14}C]-labeled products of microchemical degradation [27] on TLC plates were detected by autoradiography using a Phospho Imager Fuji-Film FLA3000 (Fuji Photo Film Co., Tokyo, Japan)

Acknowledgements

Infrared spectra and initial mass spectra were kindly recorded by Drs. Serge Moss, Christian Guenat and Joachim Blanz (all Novartis, Basle, Switzerland). Professor Helma Wennemers is acknowledged for giving access to the CD spectrometer and Dr. Rolf Kramer (both ETH Zürich, Switzerland) for introducing H.K. into its use. The authors are grateful to Christel Kakoschke and Andrea Abrahamik (both HZI, Braunschweig, Germany) for technical assistance. Prof. em. Heinz Falk (University of Linz, Austria) is gratefully acknowledged for his interest in this work and for helpful comments on stereochemical aspects of bilins.

Author Contributions

V.W. performed NMR spectroscopy, analyzed related data and assigned structures; M.N.

performed mass spectroscopy and analyzed related data; H.K. planned the project, performed all other work, analyzed data and wrote the manuscript.

References

1. Lamarter T (2004) Evolution of cyanobacterial and plant phytochromes. *FEBS Lett* **573**, 1-5.
2. Frankenberg N & Lagarias JC (2003) Biosynthesis and biological functions of bilins. In *The Porphyrin Handbook*. (Kadish KM, Smith KM & Guillard R, eds), pp. 211-236, Academic Press, Amsterdam.
3. Kayser H (1985) Pigments. In *Comprehensive Insect Biochemistry, Physiology and Pharmacology* (Kerkut GA & Gilbert LI, eds), vol. 10, pp. 367-415. Pergamon Press, Oxford.
4. Huber R, Schneider M, Mayr I, Müller R, Deutzmann R, Suter F, Zuber H, Falk H & Kayser H (1987) Molecular structure of the bilin binding protein (BBP) from *Pieris brassicae* after refinement at 2.0 Å resolution. *J Mol Biol* **198**, 499-513.
5. Holden HM, Rypniewski WR & Law JH (1987) The molecular structure of insecticyanin from the tobacco hornworm *Manduca sexta* L. at 2.6 Å resolution. *EMBO Journal* **6**, 1565-1570.

6. Scheer H & Kayser H (1988) Conformational studies of biliproteins from the insects *Pieris brassicae* and *Cerura vinula*. *Z Naturforsch* **43c**, 84-90.
7. Kayser H, Mann K, Machaidze G, Nimtz M, Ringler Ph, Müller SA & Aebi U (2009) Isolation, characterisation and molecular imaging of a high-molecular-weight insect biliprotein, a member of the hexameric arylphorin protein family. *J Mol Biol* **389**, 74-89.
8. Brown KR, Allan BM, Do P & Hegg EL (2002) Identification of novel hemes generated by heme a synthase: evidence for two successive monooxygenase reactions. *Biochem* **41**, 10906-10913.
9. Caughey WS, Smythe GA, O'Keeffe DH, Maskasky JE & Smith ML (1975) Heme a of cytochrome c oxidase. Structure and properties: comparisons with hemes b, c, and s and derivatives. *J Biol Chem* **250**, 7602-7622.
10. Wu W, Chang CK, Varotsis C, Babcock GT, Puustinen A & Wikström M (1992) Structure of the heme o prosthetic group from the terminal quinol oxidase of *Escherichia coli*. *J Am Chem Soc* **114**, 1182-1187.
11. Kikuchi G, Yoshida T & Noguchi M (2005) Heme oxygenase and heme degradation. *Biochem Biophys Res Commun* **338**, 558-567.
12. Kayser H & Krull-Savage U (1984) Development-specific incorporation of (¹⁴C)5-aminolevulinate and (³H)leucine into cytochrome c and a biliprotein in the butterfly, *Pieris brassicae*. Correlation with the ecdysteroid titer in the pupa. *Z Naturforsch* **39c**, 948-957.

13. Goodman WG & Granger NA (2005) The Juvenile Hormones. In *Comprehensive Molecular Insect Science* (Gilbert LI, Iatrou K & Gill S, eds). vol 3, pp. 319-408. Elsevier, Amsterdam.
14. Richard DS, Applebaum SW, Sliter TJ, Baker FC, Schooley DA, Reuter CC, Henrich VC & Gilbert LI (1989) Juvenile hormone bisepoxide biosynthesis *in vitro* by the ring gland of *Drosophila melanogaster*: a putative juvenile hormone in the higher Diptera. *Proc Natl Acad Sci USA* **86**, 1421-1425.
15. Kotaki T, Shinada T, Kaihara K, Ohfuné Y & Numata H (2009) Structure determination of a new juvenile hormone from a heteropteran insect. *Org Lett* **11**, 5234-5237.
16. Mauchamp B, Darrouzet E, Malosse C & Couilland F (1999) 4'-OH-JH-III: an additional hydroxylated juvenile hormone produced by locust corpora allata *in vitro*. *Insect Biochem Molec Biol* **29**, 475-480.
17. Darrouzet E, Mauchamp B, Prestwich GD, Kerhoas L, Ujváry I & Couillaud F (1997) Hydroxy juvenile hormones: new putative juvenile hormones biosynthesized by locust corpora allata *in vitro*. *Biochem Biophys Res Commun* **240**, 752-758.
18. Sutherland TD, Unnithan GC, Andersen JF, Evans PH, Murataliev MB, Szabo LZ, Mash EA, Bowers WS & Feyereisen R (1998) A cytochrome P450 terpenoid hydroxylase linked to the suppression of insect juvenile hormone synthesis. *Proc Natl Acad Sci USA* **95**, 12884-12889.
19. Gilbert LI, Granger NA & Roe RM (2000) The juvenile hormones: historical facts and speculations on future research directions. *Insect Biochem Molec Biol* **30**, 617-644.

20. Charles J-Ph, Iwema T, Epa VC, Takaki K, Rynes J & Jindra M (2011) Ligand-binding properties of a juvenile hormone receptor, Methoprene-tolerant. *Proc Natl Acad Sci USA* **108**, 21128-21133.
21. Falchuk KH, Contin JM, Dziedzic TS, Feng Z, French TC, Heffron GJ & Montorzi M (2002) A role for biliverdin IX in dorsal axis development of *Xenopus laevis* embryos. *Proc Natl Acad Sci USA* **99**, 251-256.
22. Condezo-Hoyos L, Abderrahim F, Conde MV, Susin C, Díaz-Gil JJ, González MC & Arribas SM (2009) Antioxidant activity of liver growth factor, a bilirubin covalently bound to albumin. *Free Radical Biol Med* **46**, 656-662.
23. Yamashita E, Aoyama H, Yao M, Muramoto K, Shinzawa-Itoh K, Yoshikawa S & Tsukihara T (2005) Absolute configuration of the hydroxyfarnesylethyl group of haem A, determined by X-ray structural analysis of bovine heart cytochrome c oxidase using methods applicable at 2.8 Å resolution. *Acta Cryst* **D61**, 1373-1377.
24. Boiadjiev SE & Lightner DA (1999) Optical activity and stereochemistry of linear oligopyrroles and bile pigments. *Tetrahedron: Asymmetry* **10**, 607-655.
25. Krois D (1991) Geometry versus basicity of bilatrienes: stretched and helical protonated biliverdins. *Monatsh Chemie* **122**, 495-506.
26. Hederstedt L (2012) Heme A biosynthesis. *Biochim Biophys Acta* **1817**, 920-927.
27. Rüdiger W (1969) Chromsäure- und Chromatabbau von Gallenfarbstoffen. *Hoppe-Seyler's Z Physiol Chem* **350**, 1291-1300.

28. Castro B, Dormoy JR, Evin G & Selve C (1975) Reactifs de couplage peptidique IV (1) – L’hexafluorophosphate de benzotriazolyl N-oxytrisdiméthylamino phosphonium (B.O.P.). *Tetrahedron Lett* **16**, 1219-1222.
29. Kim MH & Patel DV (1994) “BOP” as a reagent for mild and efficient preparation of esters. *Tetrahedron Letters* **35**, 5603-5606.
30. Bonnett R & McDonagh AF (1973) The meso-reactivity of porphyrins and related compounds. Part VI. Oxidative cleavage of the haem system. The four isomeric biliverdins of the IX series. *J chem Soc Perkin Trans I*, 881-888.
31. Sehringer B & Kayser H (2006) Butterfly wings, a new site of porphyrin synthesis and cleavage. Studies on the expression of the lipocalin bilin-binding protein in *Pieris brassicae*. *Insect Biochem Mol Biol* **36**, 482-491.

Supporting information

Table S1. NMR data of CV-bilin.

Figures S1-S5. 2D NMR spectra of CV-bilin.

Table 1. Absorption maxima and peak ratios of CV-bilin and biliverdin (BV) isomers.

Bilin	Solvent	UV [nm]	VIS [nm]	UV/VIS ratio
CV-Bilin	Methanol	368	624	2.8
	5% HCl/Methanol	367	652	1.6
	Chloroform	375	620	2.1
BV IX α	Methanol	377	666	3.7
	5% HCl/Methanol	375	696	2.1
	Chloroform	378	660	3.3
BV IX γ	Methanol	369	643	3.3
	5% HCl/MeOH	361	697	2.1
	Chloroform	375	640	3.3

Figure legends

Figure 1. UV-VIS spectra (A1) and CD spectra (A2) of free CV-bilin in 50% MeOH/H₂O and of its native biliprotein complex (B1 and B2, respectively) in 50 mM Tris/HCl, pH 7.4.

Figure 2. FTIR spectrum of CV-bilin. Diagnostic absorptions are at 3385 cm⁻¹ and around 2900 cm⁻¹.

Figure 3. ESI MS of CV-bilin recorded in positive mode as [M+H]⁺. The molecular mass of 852 corresponds to C₄₈H₆₀N₄O₁₀. The masses at m/z 835.39, m/z 817.37 and m/z 799.37 are due to successive losses of water.

Figure 4. RP-HPLC of free and acetylated CV-bilin. Upper panel: free CV-bilin, retention time of 12.3 min, from the supernatant after apoprotein precipitation with methanol. Lower panel: acetylated CV-bilin with products eluting at 14.5 min, 16.0 min, 17.4 min and 18.4 min (for further details see text). Figures represent screen shots.

Figure 5. TLC autoradiograms of products of chromic acid degradation of [¹⁴C]CV-bilin (CV) and of [¹⁴C] biliverdin IX γ from *P.brassicae* (PB). Samples: Free, unesterified bilin; Est, bilin dimethylester. Products: 1, methyl vinyl maleimide; 2, methanol adduct of 1; 3, hematinic acid methylester; 4, CV-bilin specific polar derivative of 3; 5, hematinic acid. S, sample application site. M, markings of TLC plate.

Figure 6. 1D ¹H NMR spectrum (600 MHz) of CV-bilin in CD₃OD containing a small amount of C₆D₆. (For 2D NMR spectra see Figs. S1-S5.)

Figure 7. Structure of CV-bilin (A) and of the modified side-chain of the minor, more polar product (B). The tetrapyrrole system shows 10*E* configuration. The numbering of the side-chain at C-18 of the tetrapyrrole is that used in the NMR data table (Table S1).

Figure 8. Structure of CV-bilin (A) shown in 10*Z* configuration (compare to Fig. 7A) to facilitate structural comparison with heme A (B).

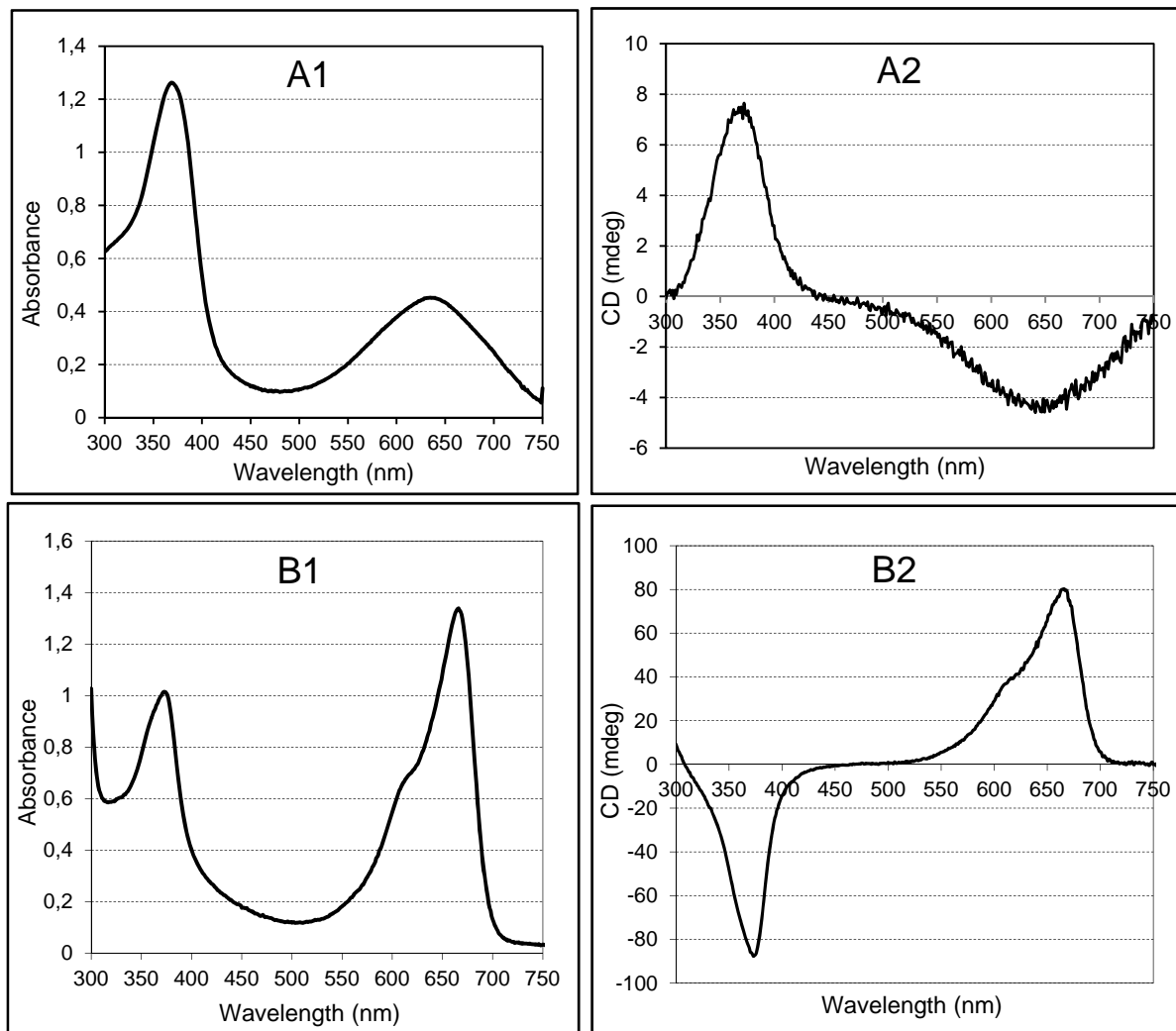


Fig.1

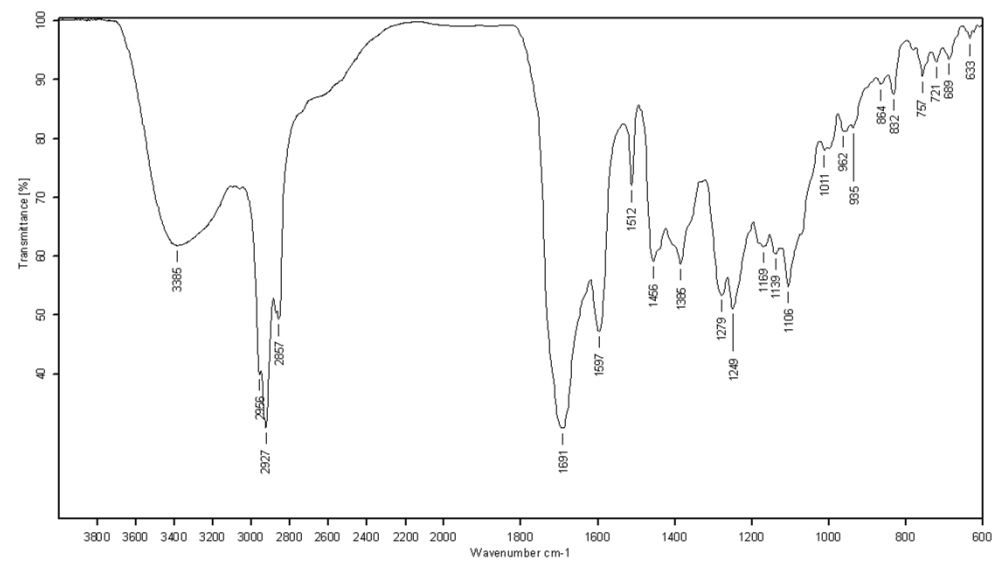


Fig.2

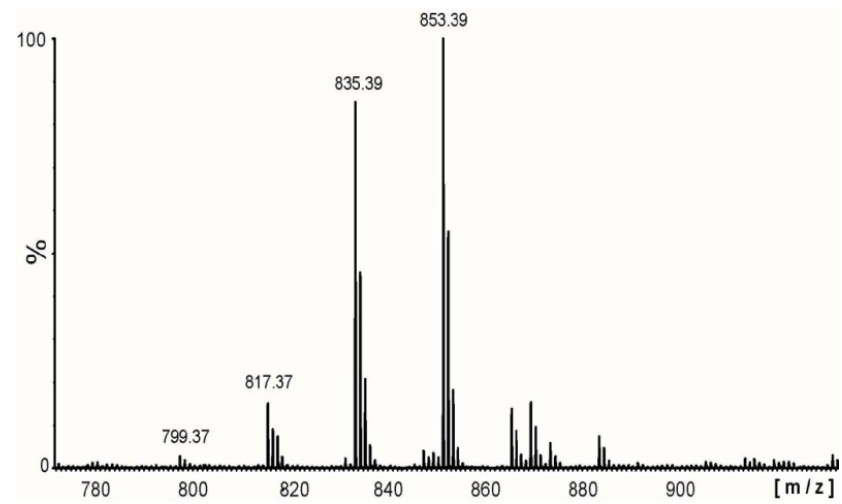


Fig.3

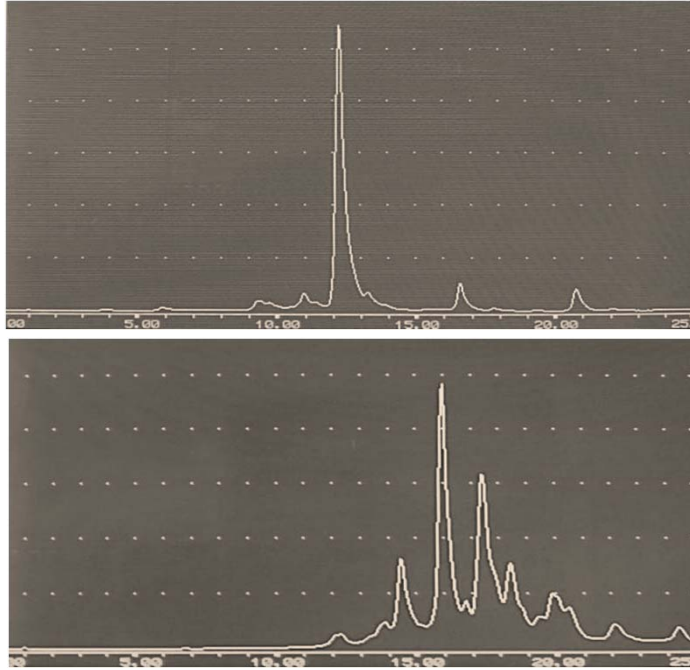


Fig. 4

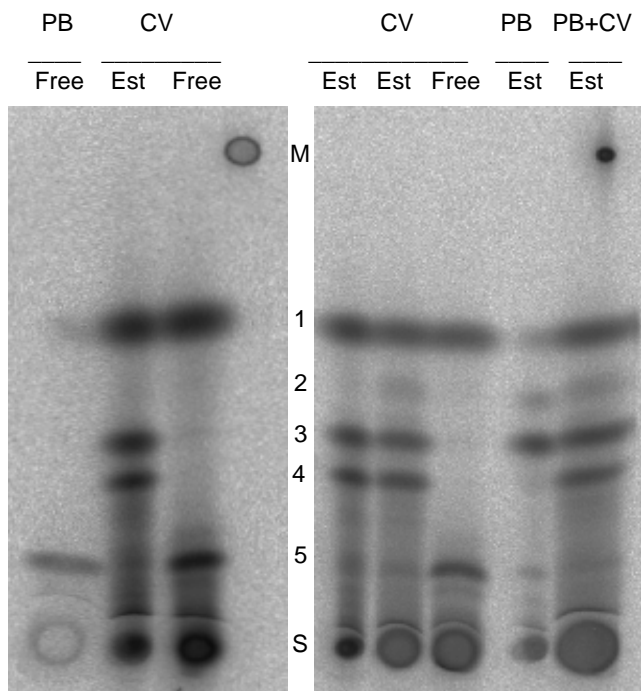


Fig. 5

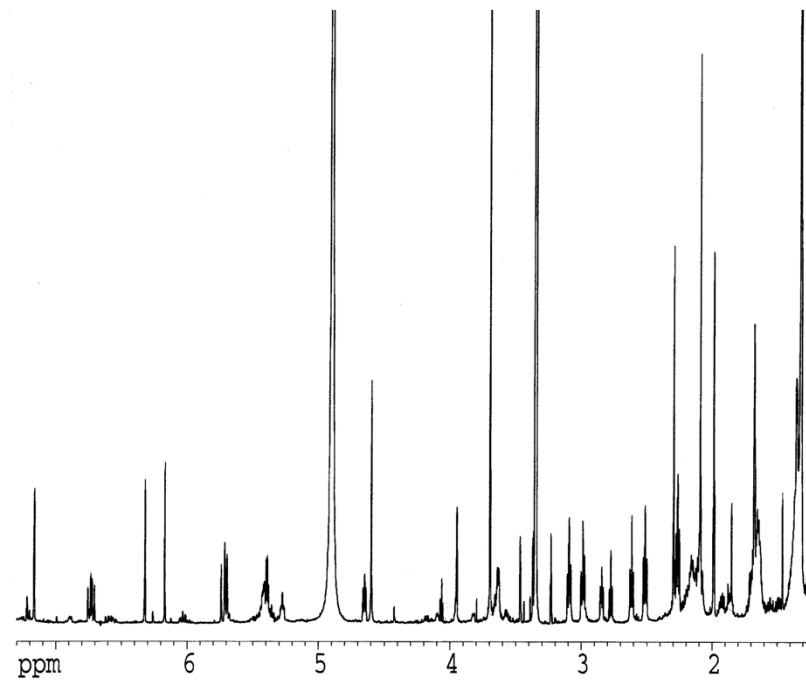


Fig.6

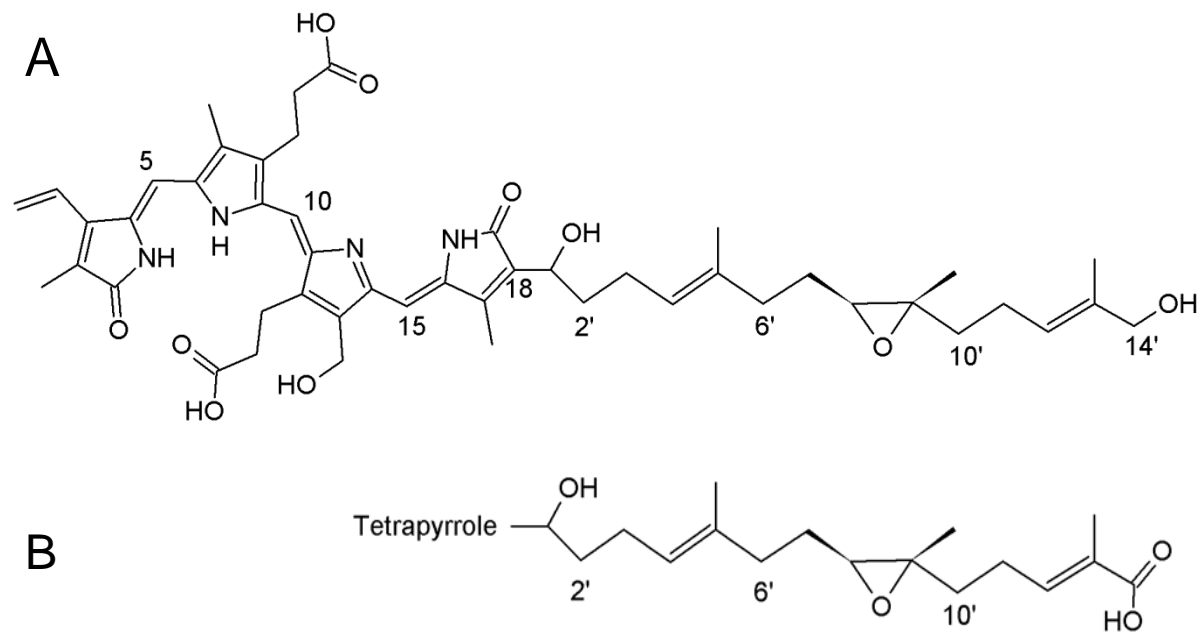


Fig.7

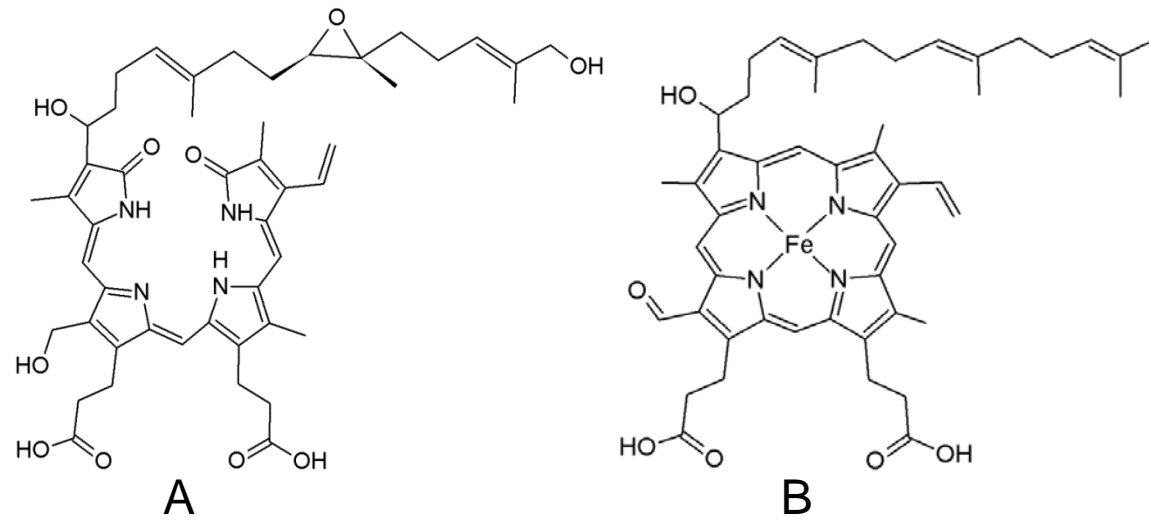


Fig.8

Supporting information to the manuscript

Structure of a novel farnesylated bilin from insect: formation by α -cleavage of heme A of mitochondrial cytochrome c oxidases?

Hartmut Kayser^{a1}, Victor Wray^b, Manfred Nimtz^b

^aInstitut für Allgemeine Zoologie und Endokrinologie, Universität Ulm, Helmholtzstrasse 8/1,
89081 Ulm, Germany

^bHelmholtz-Zentrum für Infektionsforschung GmbH, Molekulare Strukturbiologie,
Inhoffenstrasse 7, 38124 Braunschweig, Germany

¹To whom correspondence should be addressed.

E-mail: hartmut.kayser@uni-ulm.de

Table S1. NMR data of CV-bilin.

Figures S1-S5. 2D NMR spectra of CV-bilin.

Table S1. NMR data of CV-bilin.Spectra were recorded in CD₃OD containing a small amount of C₆D₆.

	¹ H	¹³ C	¹ H - ¹³ C correlations	NOE
Tetrapyrrole				
1		174.3		
2		129.8		
3		141.7		
4		nd		
5	6.18 s	100.1	C-3	H-7 ^I , H-3 ^I
6		157.3		
7		131.0		
8		142.7		
9		145.7		
10	7.14 s	118.0	C-8, C-9, C-11, C-12	H-8 ^I not H-12 ^I
11		138.4		
12		140.1		
13		129.7		
14		145.2		
15	6.33 s	100.1	C-13, C-17	H-13 ^I , H-17 ^I
16		142.7		
17		144.4		
18		133.6		
19		173.9		
2 ^I	1.99 s	9.5	C-1, C-2, C-3	
3 ^I	6.75 dd	127.4	C-2, C-3, C-4	
3 ^{II}	5.74 d, 5.72 d	123.1	C-3	
7 ^I	2.11 s	9.1	C-6, C-7, C-8	H-5
8 ^I	2.98 t	21.6	C-7, C-8, C-9, C-8 ^{III}	H-10
8 ^{II}	2.50 t	38.5	C-8, C-8 ^{III}	
8 ^{III}		178.9		
12 ^I	3.08 t	21.4	C-11, C-12, C-13, C-12 ^{III}	H-13 ^I
12 ^{II}	2.59 t	39.6	C-12, C-12 ^{III}	H-13 ^I
12 ^{III}		179.3		
13 ^I	4.60 s	54.3	C-12, C-13, C-14	H-12 ^I , H-12 ^{II} , H-15
17 ^I	2.30 s	9.7	C16, C-17, C-18, C-19 (w)	H-15

Side-chain at C-18				
1' ^a	4.62 dd	66.9	AC-17, AC18, AC-19, C-2, C-3	
2'	1.88, 1.83 m	36.6	C-4	
3'	2.15 m	25.3	C-4, C-5, C-6	
4'	5.27 dd	125.5	C-2, C-5 ^l , C-6	H-6
5'	---	135.4		
5' ^l	1.68 s	15.9	C-4, C-5, C-6	
6'	2.16	37.1		H-4
7'	1.63 m	26.8	C-8	
8'	2.78 t	64.6	C-6, C-7, C-10	H-10
9'	---	62.6		
9' ^l	1.29 s	16.6	C-8, C-9, C-10	
10'	1.67 m/1.55 m	39.5	C-8, C-9, C-11, C-12	H-8
11'	2.17 m	24.6 or 28.1		
12'	5.41 dd	125.6	C-10, C-13 ^l	H-14
13'	---	136.4		
13' ^l	1.68 s	13.5	C-12, C-13, C-14	
14'	3.94 s	68.6	C-12, C-13, C-13 ^l	H-12

Footnotes: ^a This is 18^l in the tetrapyrrole unit.

Figures S1 – S5: 2D NMR spectra of CV-bilin.

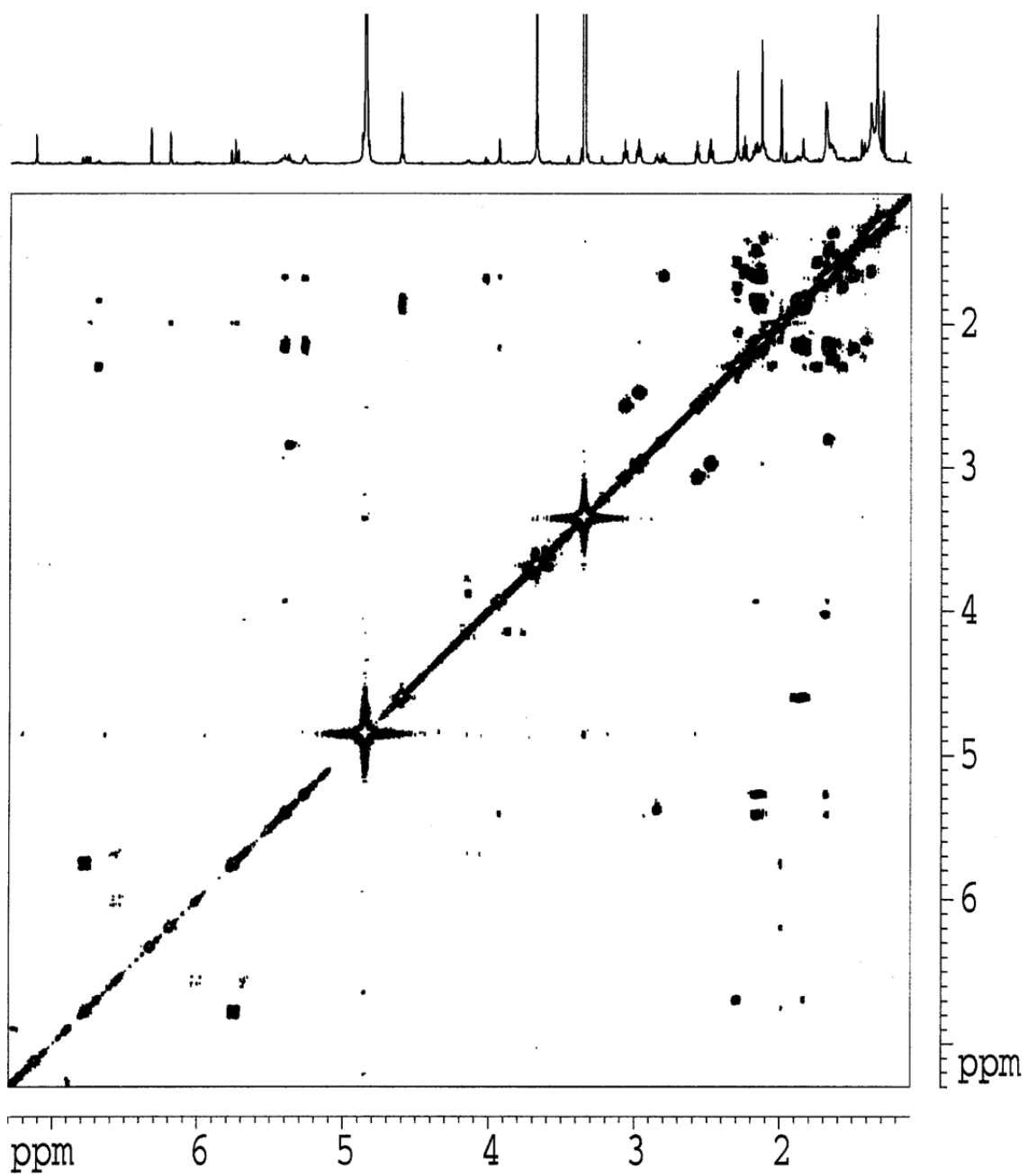


Figure S1. 2D COSY spectrum of CV-bilin in CD₃OD.

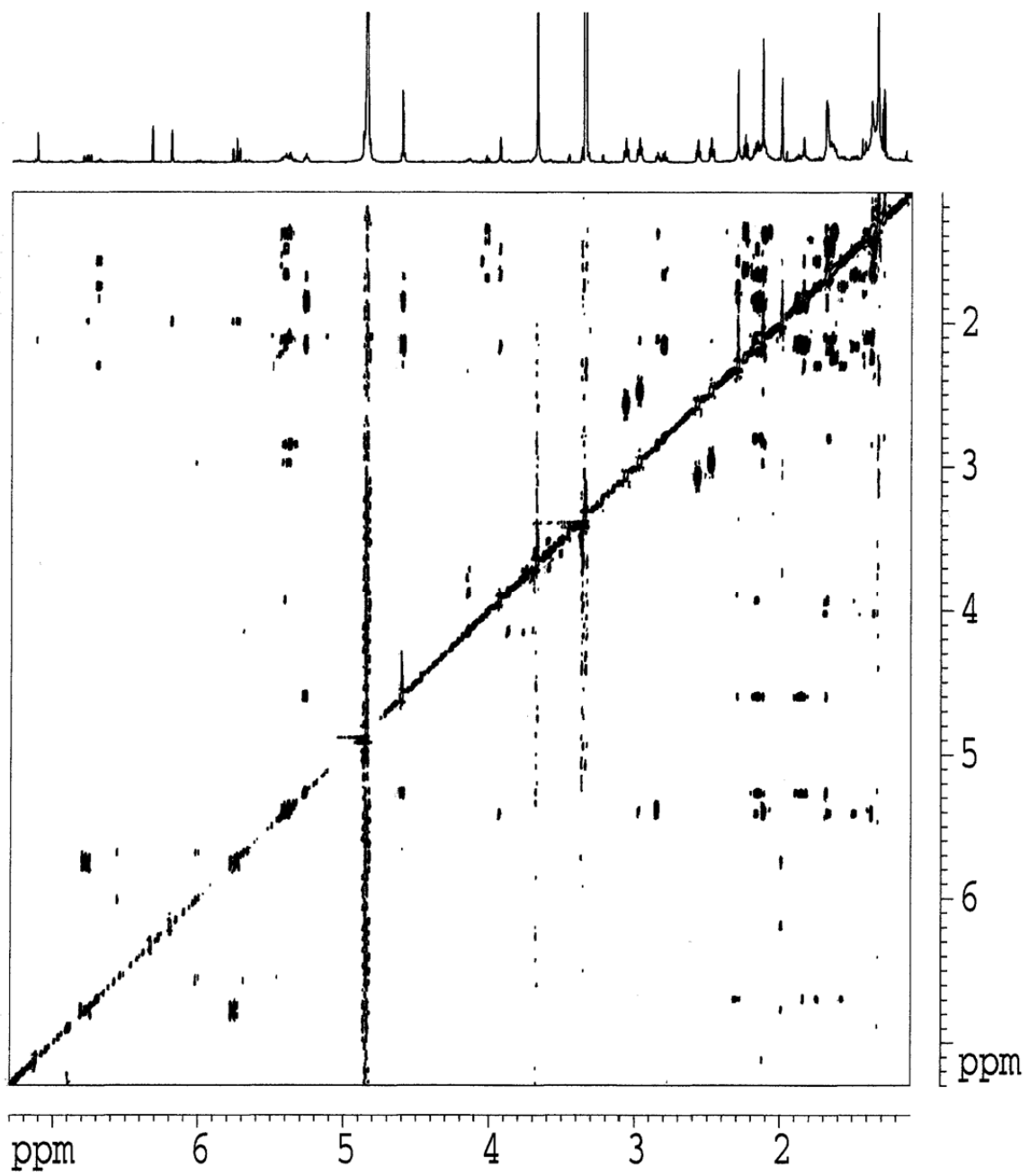


Figure S2. 2D TOCSY spectrum of CV-bilin in CD₃OD.

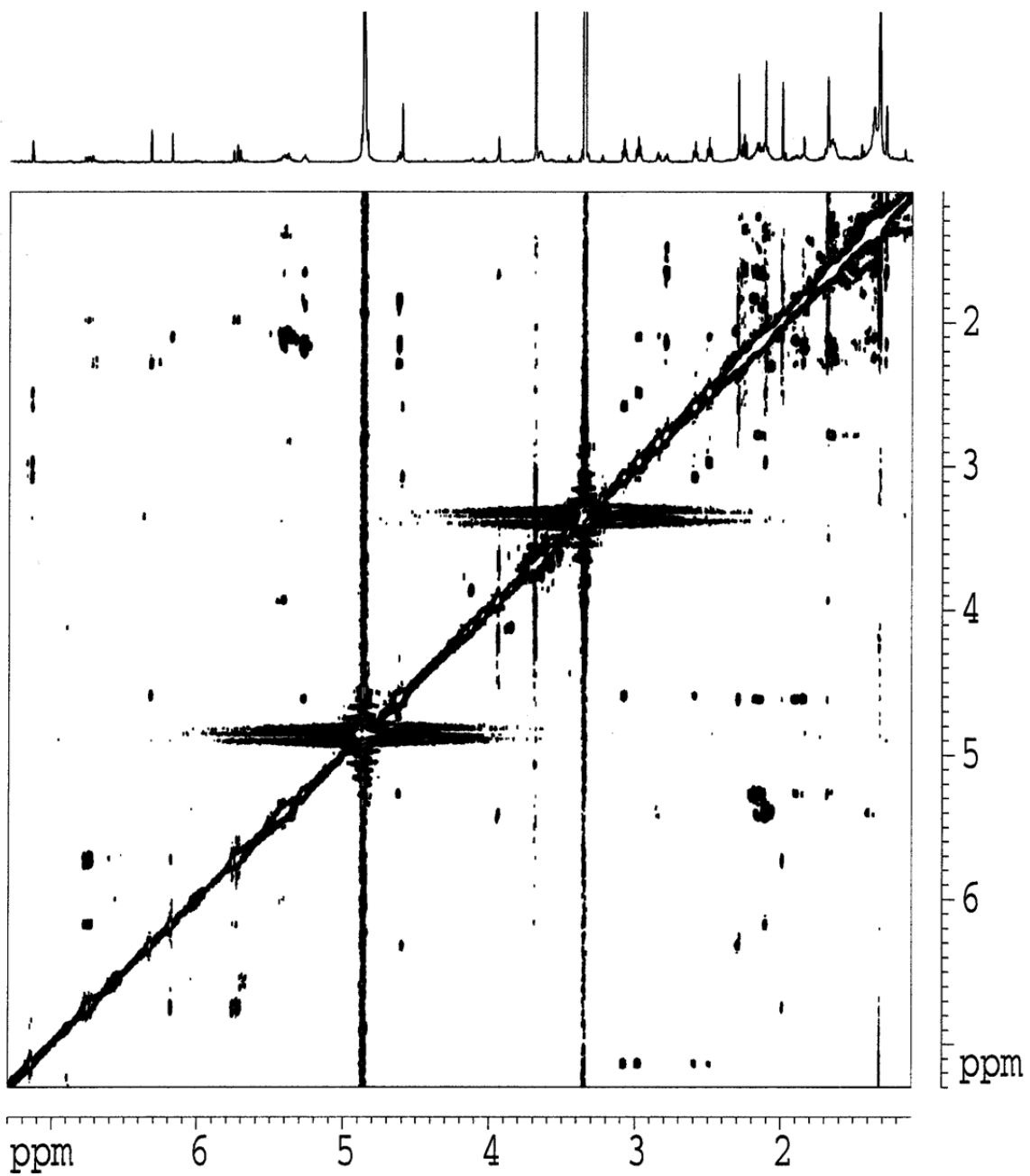


Figure S3. 2D ROESY spectrum of CV-bilin in CD₃OD with a small amount of C₆D₆.

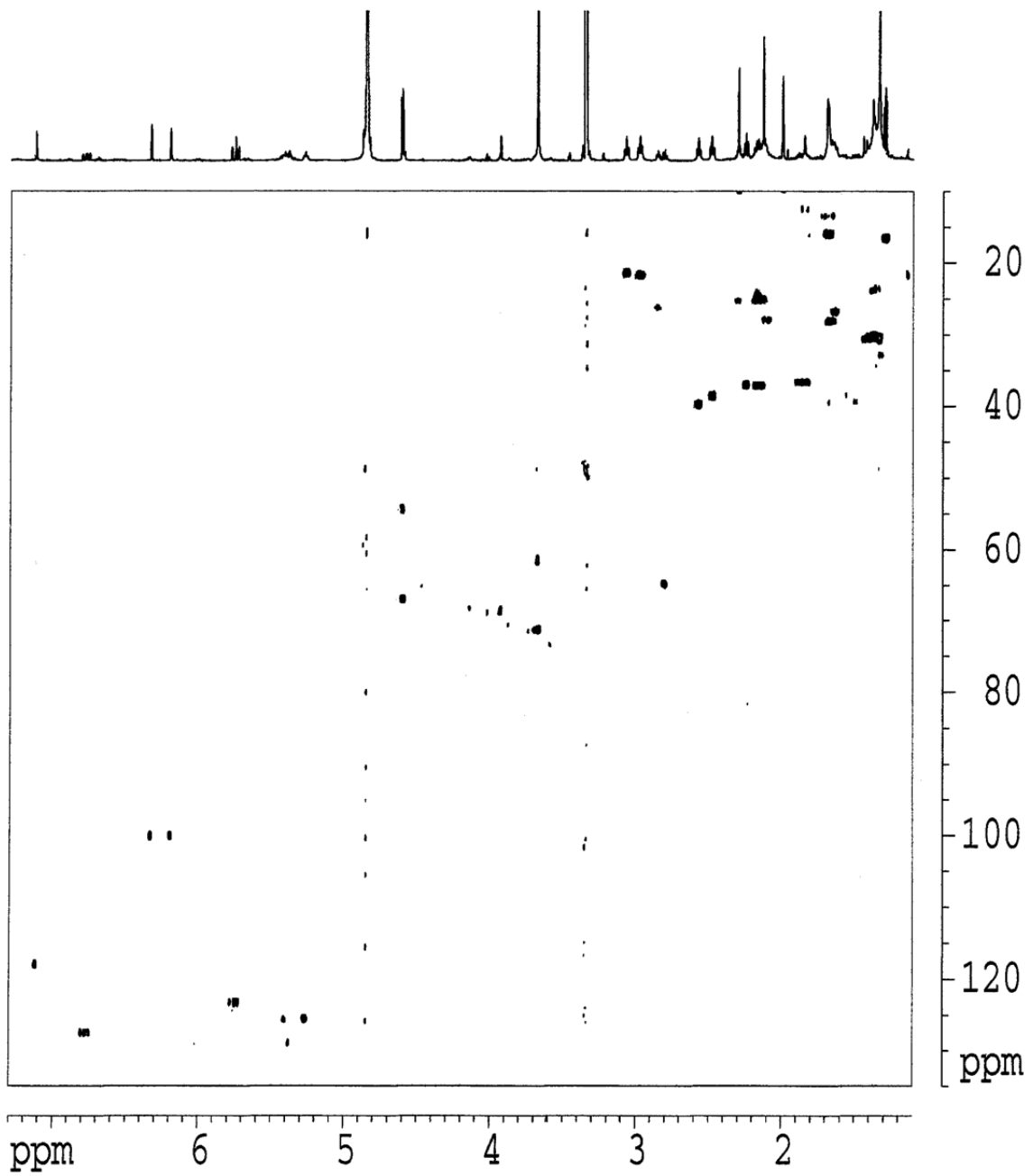


Figure S4. 2D C-H direct correlation (HMQC) of CV-bilin in CD_3OD .

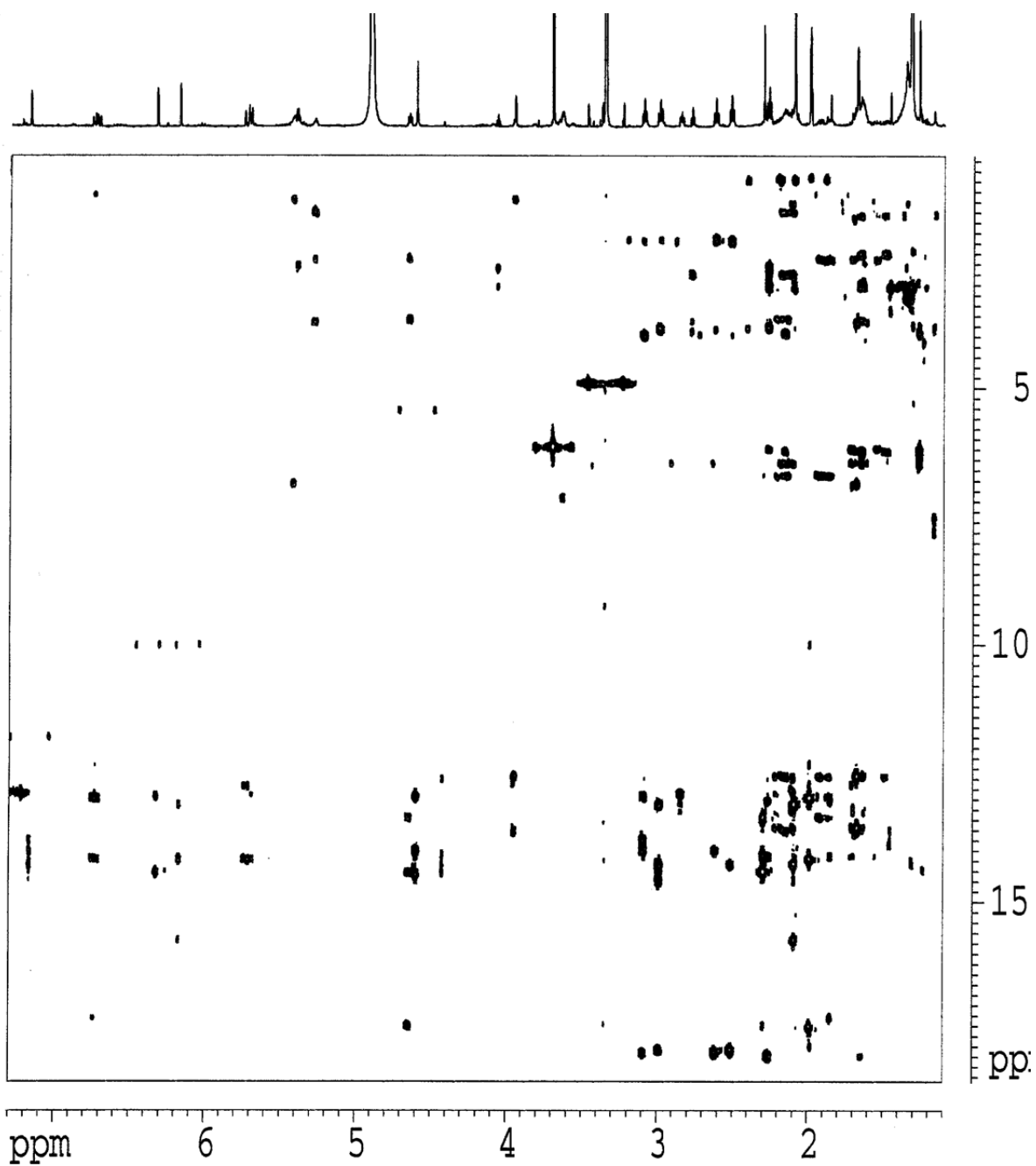


Figure S5. 2D C-H long-range correlation (HMBC) of CV-bilin in CD_3OD with a small amount of C_6D_6 .