New Insights into the Conversion of Versicolorin A in the Biosynthesis of Aflatoxin B1

David Conradt, Michael A. Schätzle, Julian Haas, Craig A. Townsend and Michael Müller

Supporting Information

I.	General remarks	S2
II.	Molecular cloning, bacterial expression and activity measurements	S3
III.	Chemoenzymatic reductions (experimental data and analytical methods)	S5
IV.	^{1}H and ^{13}C NMR spectra	S8
V.	CD and VCD spectra	S10
VI.	Nucleotide and protein sequences	S12
VII.	References	S14

I. General remarks

Nuclear magnetic resonance (NMR) spectra were recorded at 24 °C on a DRX 400 spectrometer (Bruker) operating at 400 and 100 MHz for ^{1}H and ^{13}C acquisitions, respectively. Chemical shifts (δ) of the ^{1}H NMR and ^{13}C NMR spectra are reported in ppm with a solvent resonance as internal standard (^{1}H NMR: acetone-d_5 2.05, ^{13}C NMR: acetone-d_6 30.0). CD spectroscopy was carried out on a Jasco J-810 spectrophotometer (Jasco International Co.). Activity measurements were performed on a UV Mini 1240 UV/Vis spectrophotometer from Shimadzu. All chemical reagents and solvents were obtained from Sigma-Aldrich, Applichem or Roth. Versicolorin A was provided by C. A. Townsend, John Hopkins University Baltimore, USA. Glucose dehydrogenase (GDH) was obtained from evocatal (Düsseldorf, Germany). Yields refer to chromatographically pure materials; conversions were calculated from the reactant-product ratios in crude ^{1}H NMR spectra.

FT-VCD and FT-IR spectra were recorded on a Bruker Tensor 27 FT-IR spectrometer equipped with a Bruker PMA 50 VCD module (Bruker Optik GmbH, Ettlingen) with a resolution of 4 cm^{-1} in a rotating BaF_2 cell (path length 110 μm). Experimental spectra represent the average of 4 h 40 min measurement. VCD and IR-spectra were corrected by subtraction of the solvent spectrum. Opus 7.0 software (Bruker Corporation) was used to analyze the spectra. The conformer search was carried out at the molecular mechanics level (MMFF) using Spartan 08 (Wavefunction Inc., Irvine, CA, USA). Geometry optimization and calculation of energies as well as IR and VCD frequencies and intensities of 10 conformers at DFT level (B3LYP/6-31+G(d,p), SMD solvation model for dmso) was carried out with Gaussian 09 Revision B.01.[1,2] The calculated vibrational frequencies were scaled by 0.98. Shown spectra were obtained as Boltzmann-weighted average of the calculated spectra of each single conformer with Lorentzian lineshapes of 6 cm^{-1} width around calculated intensities.

II. Molecular cloning, bacterial expression and activity measurements

Bacteria, DNA preparation and expression vector

E. coli TG1 and BL21(DE3) (Stratagene) were used for cloning and expression of AflM-his. Codon-optimized, His-tagged *aflM-his* from *Aspergillus parasiticus* including an N-terminal T7 10 feeder and the restriction sites was synthesized by MWG Eurofins Operon. The upstream construct was identical to the one described elsewhere[3,4] (see also **VI.**). Ligation into pET19b by XbaI and BamHI restriction sites was also performed by MWG Eurofins Operon (see **VI.**).

Cloning and sequence analysis

Plasmid DNA isolation was performed using the Bioneer AccuPrep Plasmid Mini Extraction Kit. Transformation of competent *E. coli* cells was performed by applying a heatshock at 42 °C for 90 s. Correct sequences were confirmed by sequence analysis through GATC-Biotech.

Media and growth conditions

One clone was picked and dispersed in 5 mL of LB-media (Lennox), followed by incubation overnight (37 °C, 160 rpm). Ampicillin (100 $\mu g \cdot mL^{-1}$) was added as required.

Cultivation and expression

The overnight cultures were diluted to 1 L of medium each (ampicillin 100 $\mu g \cdot mL^{-1}$). IPTG (0.1 mM) was added after reaching the midlog phase (OD_{600}= 0.6). The cultures were incubated for 16 h at 25 °C and 160 rpm.

Workup and storage

The harvested *E. coli* cells were resuspended in KPi buffer (50 mM, pH = 7.0; 5 mL per harvested cells of 1 L medium). The cells were disrupted by sonication (6 times 15 sec, Branson Sonifier II "Modell W-250", Heinemann), followed by centrifugation (30 min, 12000 g, 4 °C). Glycerol (20% v/v) was added and the crude enzyme preparation was frozen at –20 °C.

Enzyme purification

AflM-his was purified by Ni-NTA affinity chromatography. Non-specifically bound proteins were washed off with 200 mM imidazole in KPi buffer (50 mM, pH 7.0). Elution was performed with a 500 mM imidazole in KPi buffer (50 mM, pH 7.0). The eluted solution was desalted by gel filtration (Sephadex™ G-25M, GE Healthcare). Concentration of the protein was performed by ultrafiltration (Amicon Ultra-15 Centrifugal Filter Units, 10000 nominal molecular weight limit, Merck Millipore). Concentration of the protein was determined by measuring the UV absorption at 280 nm (Nanodrop 2000, Thermo Scientific, extinction coefficient 18700 $M^{-1} \cdot cm^{-1}$, molecular weight 30.7 kDa).

Activity measurements

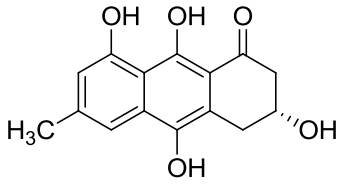
GDH was assayed as described elsewhere.[3]

Blank experiments

Blank experiments were performed to exclude background reactions ensuing from GDH and crude enzyme preparation. For this purpose, the host vector pET19b was transformed in *E. coli* BL21 and expressed in the same manner like *AflM*_pET19b in *E. coli* BL21. The obtained crude preparation was used in black experiments instead of purified AflM-his using the respective conditions. In all cases, no background reaction was observed.

III. Chemoenzymatic reductions (experimental data and analytical methods)

(*R*)-3,8,9,10-Tetrahydroxy-6-methyl-3,4-dihydroanthracen-1(2*H*)-one (14)



14

A stream of nitrogen was bubbled through the buffer solution for 30 min, followed by degassing under reduced pressure before use. In a total volume of 50 mL of KPi buffer (50 mM, pH 7.0), D-glucose (166 mg, 926 µmol, 5 eq.), $NADP^+$-Na (15.5 mg, 18.5 µmol, 10 mol%), $Na_2S_2O_4$ (322 mg, 1.8 mmol, 10 eq.), 50 mg (185 µmol) of emodin (**9**), 215 U of GDH and 4.8 mg of AflM-his were stirred under a nitrogen atmosphere for 24 hours. The solution was extracted three times with ethyl acetate (50 mL), dried over $MgSO_4$, and the solvent was removed under reduced pressure. The conversion (50%) was immediately determined by ^{1}H NMR analysis of the crude product. 10 mg (37 µmol, 20%) of the title compound was obtained after automated flash chromatography (Isolera Prime, Biotage, SNAP Cartridge: KP-Sil 25g, solvent gradient: cyclohexane/ethyl acetate 22:78 to 0:100, flow rate: 25 mL·min^{-1}). The analytical data were consistent with the one described elsewhere.[4]

TLC (cyclohexane/ethyl acetate, 1:9 v/v): R_f = 0.22

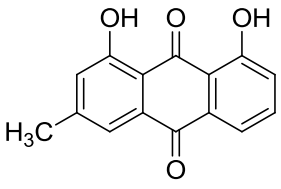
^{1}H NMR: (400 MHz, acetone-d_6), δ 2.44 (s, 3 H, CH_3), 2.80 (ddd, ^2J = 17.1 Hz, ^3J = 7.2 Hz, ^4J = 1.2 Hz, 1 H, H-2), 2.98 (ddd, ^2J = 17.1 Hz, ^3J = 3.6 Hz, ^4J = 1.1 Hz, 1 H, H-2), 3.06 (ddd, ^2J = 16.4 Hz, ^3J = 6.8 Hz, ^4J = 1.1 Hz, 1 H, H-4), 3.26 (ddd, ^2J = 16.4 Hz, ^3J = 3.8 Hz, ^4J =1.1 Hz, 1 H, H-4), 4.37 (d, ^3J = 4.3 Hz, 1 H, OH-3), 4.41–4.49 (m, 1 H, H-3), 6.68 (s, 1 H, H-7), 7.46 (s, 1 H, H-5), 7.65 (s, 1 H, OH-10), 9.78 (s, 1 H, OH-8), 15.94 (s, OH-9).

^{13}C NMR: (100 MHz, acetone-d_6), δ 22.5 (CH_3), 32.6 (C-4), 46.7 (C-2), 65.9 (C-3), 110.0 und 111.6 (C-8a und C-9a), 113.2 und 113.4 (C-5 und C-7), 117.7 (C-4a), 133.9 (C-10a), 141.6 (C-10), 143.7 (C-6), 158.9 und 160.2 (C-8 und C-9), 204.9 (C-1).

CD: (c = 182 μmol·L^{-1}, l = 0.5 cm, acetonitrile), λ [nm] (Mol. CD) = 197 (0.98), 231 (–1.08), 243 (–0.61), 262 (–1.71), 280 (–0.13), 291 (–0.44), 320 (0.22), 354 (0.02), 414 (0.43); see also **VII.**

$C_{15}H_{14}O_5$: 274.27 g·mol^{-1}.

Chrysophanol (11)



11

In the above-mentioned conversion of emodine (**9**) to **14**, the formation of the titled compound (8%) was observed and determined by ^{1}H NMR experiments.

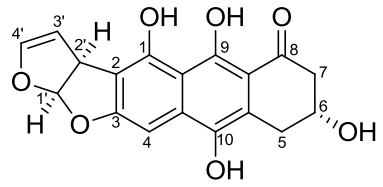
TLC (cyclohexane/ethyl acetate, 1:9 v/v): R_f = 0.91

^{1}H NMR: (400 MHz, acetone-d_6), δ 2.50 (CH_3), 7.18–7.19 (m, 1 H, H-2), 7.36 (dd, ^3J = 8.1 Hz, ^4J = 1.5 Hz, 1 H, H-7), 7.62 (d, ^4J = 1.5 Hz, 1 H, H-4), 7.79 (dd, ^3J = 7.5 Hz, ^4J = 1.5 Hz, 1 H, H-5), 7.83 (dd, ^3J = 7.5 Hz, ^3J = 8.1 Hz, 1 H, H-6), 11.95 (s, 1 H, OH-1), 12.05 (s, 1 H, OH-8).

^{13}C NMR: (100 MHz, acetone-d_6), δ 22.1 (CH_3), 113.9 (C-9a), 115.8 (C-8a), 120 (C-5), 121.7 (C-4), 124.8 (C-2), 125.2 (C-7), 133.4 (C-4a), 133.8 (C-10a), 138.3 (C-6), 150 (C-3), 163 (C-8), 164 (C-1), 182 (C-10), 192 (C-9).

$C_{15}H_{10}O_4$: 254.24 g·mol^{-1}.

(1'*R*,2'*S*,6*R*)-1,6,9,10-tetrahydroxy-2',5,6,7-tetrahydroanthra[3,2-b]furo[2',1'-*d*]furan-8(1'*H*)-one (17)



17

The reduction was performed as described above using a total volume of 5 mL of KPi buffer (50 mM, pH 7.0), 13 mg of D-glucose (72 µmol, 5 eq.), 1.2 mg of $NADP^+$-Na (1.5 µmol, 10 mol%), 25 mg of $Na_2S_2O_4$ (145 µmol, 10 eq.), 5 mg (14.6 µmol) of versicolorin A (**3**), 6 U of GDH and 5.4 mg of AflM-his. After 24 hours, the solution was extracted three times with ethyl acetate (5 mL), dried over $MgSO_4$, and the solvent was removed under reduced pressure. The conversion (25%) was determined from the ^{1}H NMR spectrum of the crude product. The absolute configuration (6*R*) was assumed with regard to compound **14**.

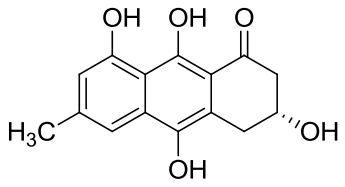
TLC (n-hexane/ethyl acetate/acetic acid, 8:1.8:0.2 v/v/v): R_f = 0.43

^{1}H NMR: (400 MHz, acetone-d_6), δ 2.80 (dd, ^2J = 17.3 Hz, ^3J = 7.2 Hz, 1 H, H-7), 2.97 (dd, ^2J = 17.3 Hz, ^3J = 3.8 Hz, 1 H, H-7), 3.05 (dd, ^2J = 16.3 Hz, ^3J = 9.9 Hz, 1 H, H-5), 3.24 (dd, ^2J = 16.3 Hz, ^3J = 3.3 Hz, 1 H, H-5), 4.38–4.46 (m, 1 H, H-6 und OH-6), 4.75 (dt, ^3J = 2.5 Hz, ^3J = 7.1 Hz, 1 H, H-2'), 5.42 (t, ^3J = 2.5 Hz, 1 H, H-3'), 6.59 (t, ^3J = 2.5 Hz, 1 H, H-4'), 6.83 (d, ^3J = 7.1 Hz, 1 H, H-1'), 7.04 (s, 1 H, H-4), 7.61 (s, 1 H, OH-10), 10.14 (s, 1 H, OH-1), 16.1 (s, 1 H, OH-9).

$C_{18}H_{14}O_7$: 342.30 g·mol^{-1}.

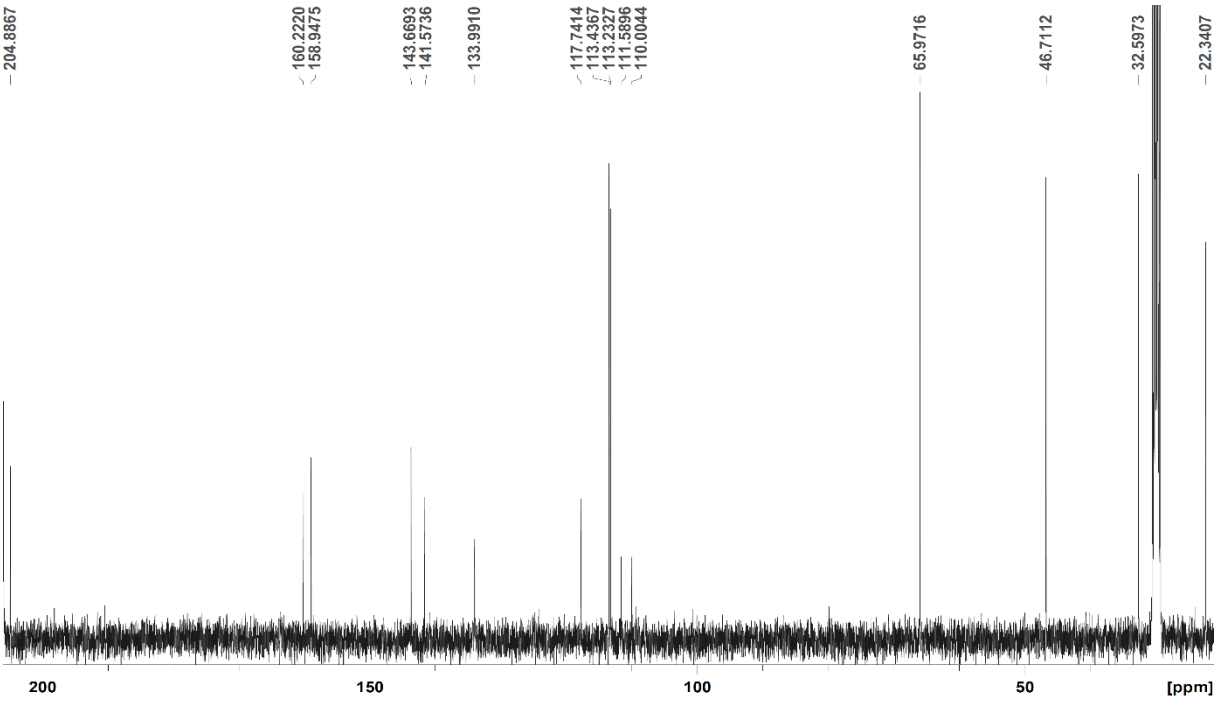
IV. ^{1}H and ^{13}C NMR spectra

^{1}H NMR (acetone-d_6)



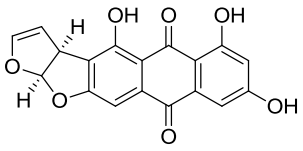
14

^{13}C NMR (acetone-d_6)

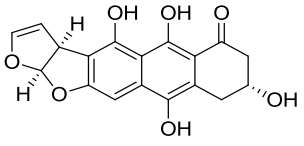


^{1}H NMR (acetone-d_6)

3/**17** = 4:1

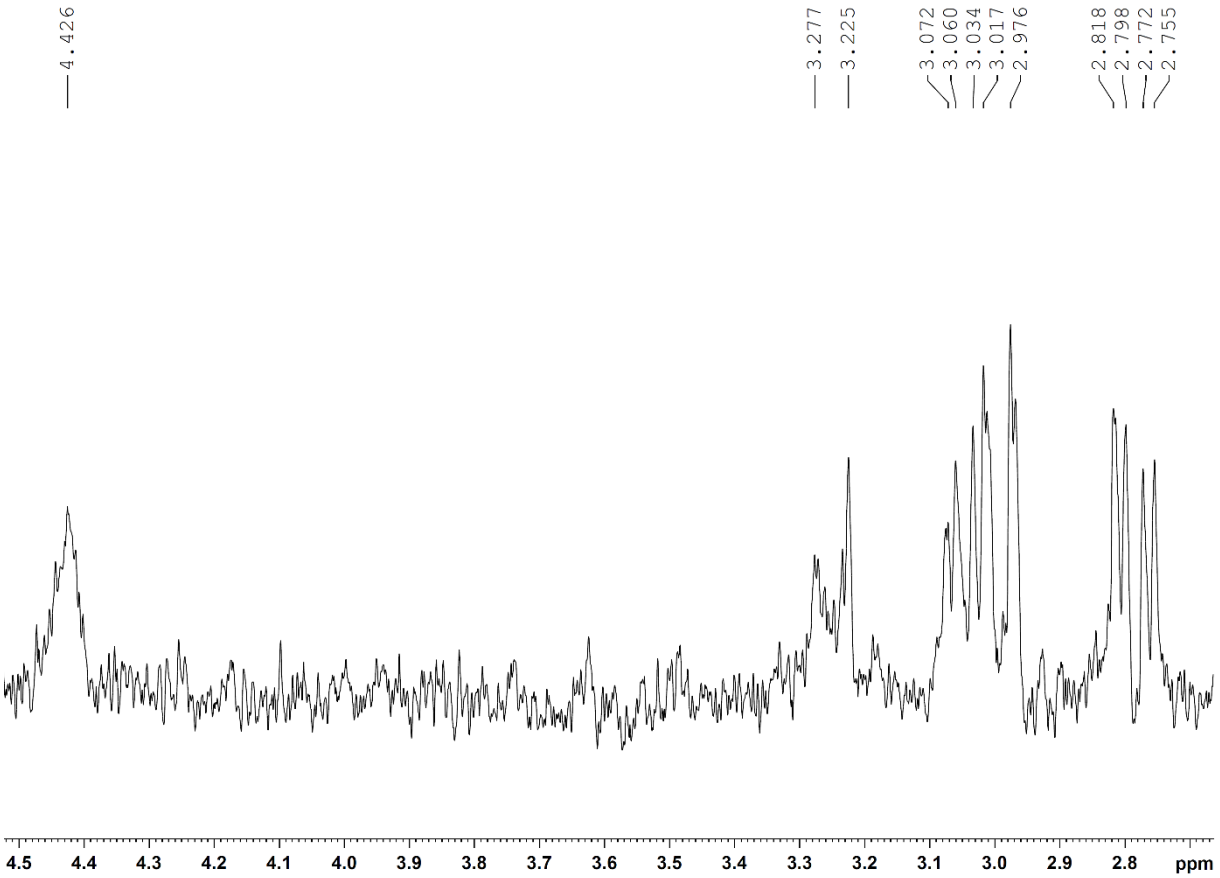


3

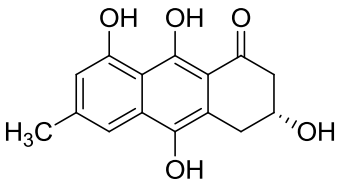


17

TOCSY NMR (3.24 ppm, acetone-d_6)



V. CD and VCD spectra



14

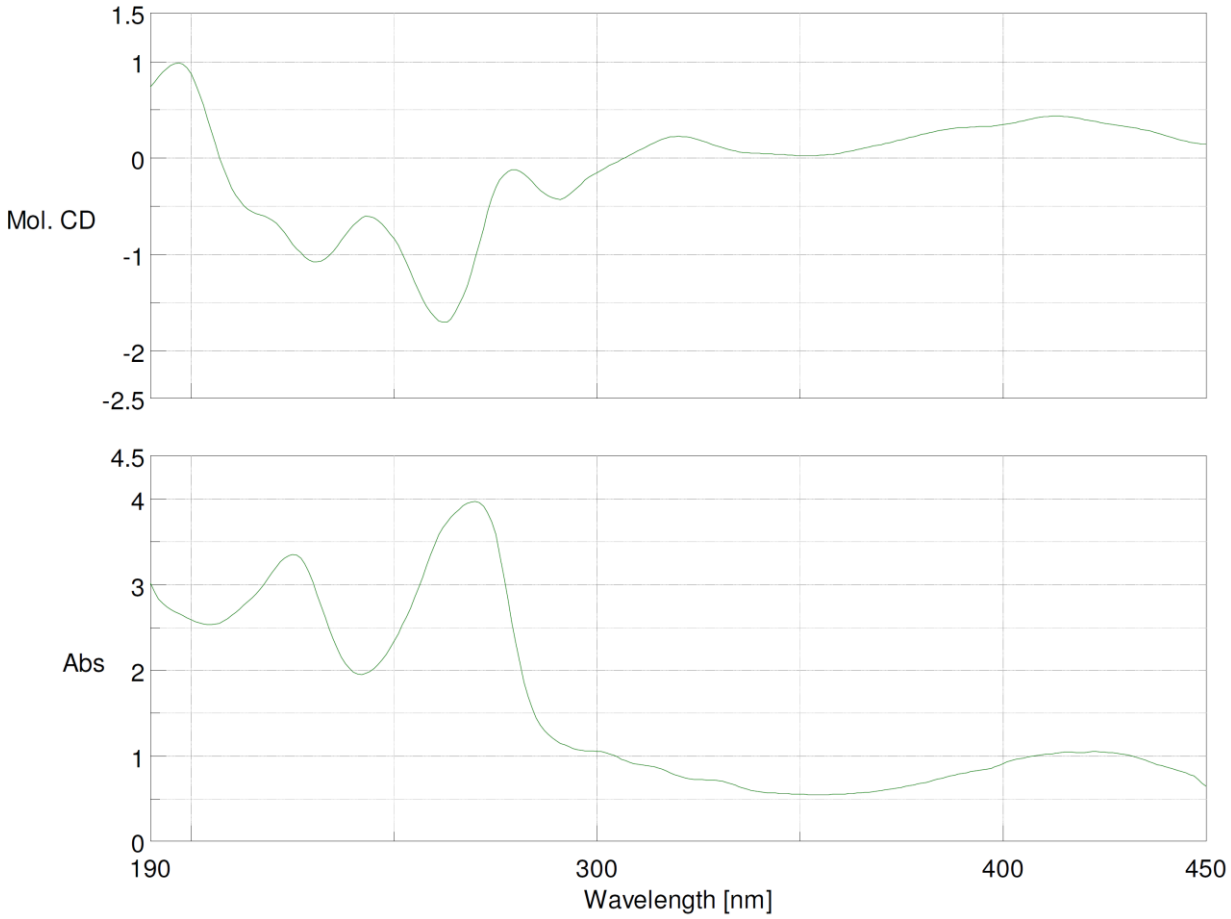
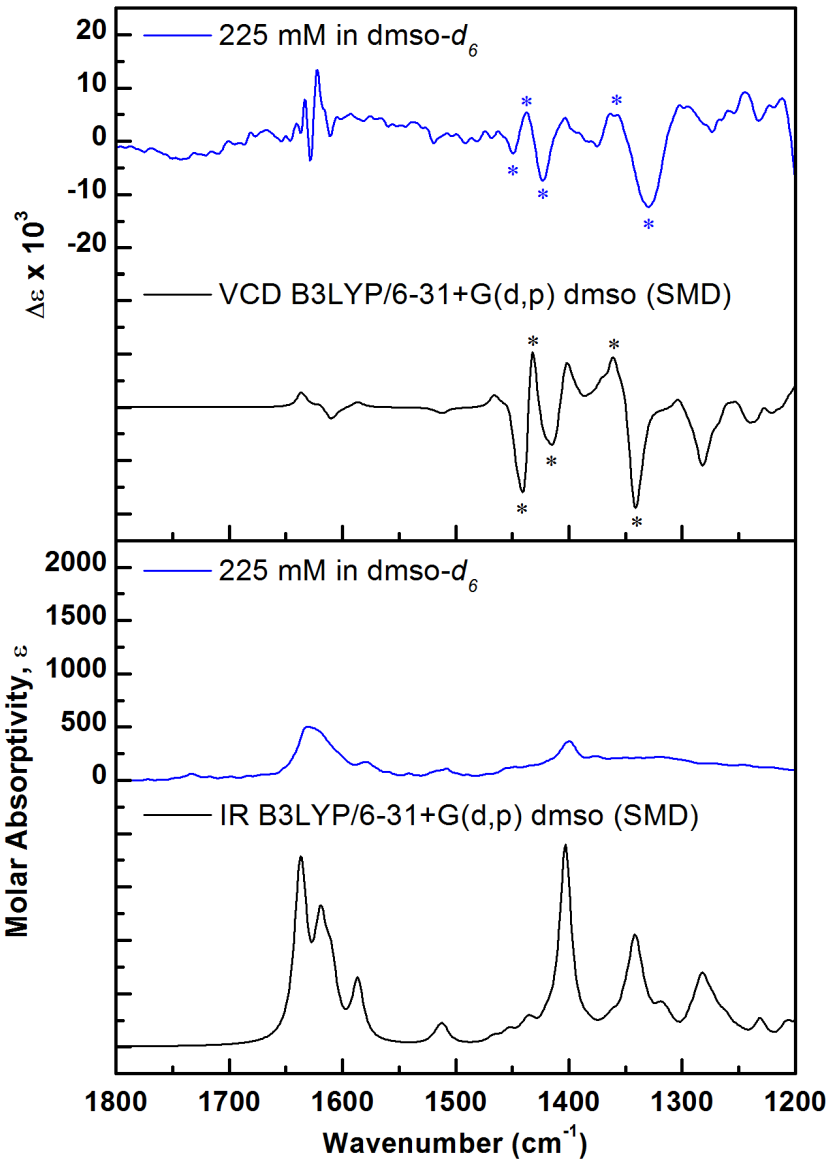
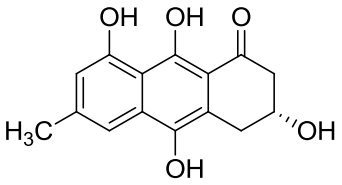


Figure S1. UV and CD spectrum of **14**.





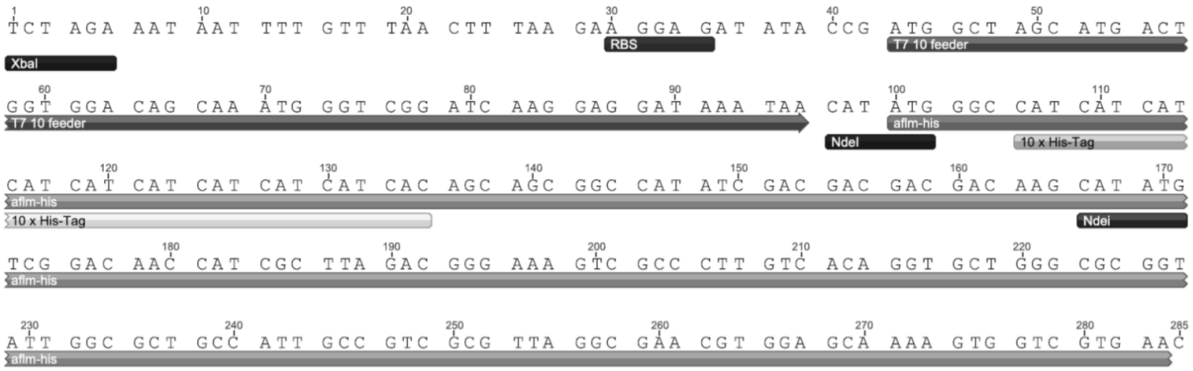
14

Figure S2. Experimental IR and VCD spectra of **14** (225 mM in dmso-d_6) in comparison to theoretical IR and VCD spectra. The spectra were calculated at the B3LYP/6-31+G(d,p) level with the SMD solvation model for dmso for 10 conformers of **14** and Boltzmann-averaged in respect to relative energies.[1,2] There is a good agreement of calculated VCD bands showing the same sign for the majority of conformers, in particular those with highest Boltzmann weight (marked with an asterisk), with corresponding bands in the experimental spectrum.

VI. Nucleotide and protein sequences

Nucleotide sequence

Upstream construct of *aflM-his*_pET19b



***AflM-his* (codon optimized)**

ATG GGC CAT CAT CAT CAT CAT CAT CAT CAT CAT CAC AGC
AGC GGC CAT ATC GAC GAC GAC GAC AAG CAT ATG TCG GAC
AAC CAT CGC TTA GAC GGG AAA GTC GCC CTT GTC ACA GGT
GCT GGG CGC GGT ATT GGC GCT GCC ATT GCC GTC GCG TTA
GGC GAA CGT GGA GCA AAA GTG GTC GTG AAC TAC GCT CAC
AGT CGT GAG GCA GCA GAG AAA GTT GTC GAG CAG ATC AAA
GCG AAC GGT ACC GAT GCC ATT GCC ATC CAA GCG GAT GTG
GGC GAT CCT GAA GCG ACG GCG AAA CTG ATG GCG GAA ACC
GTA CGC CAT TTC GGC TAT CTC GAT ATC GTG AGC TCA AAT
GCG GGC ATT GTT TCC TTT GGT CAC CTG AAG GAC GTT ACT
CCG GAA GAG TTT GAC CGG GTT TTC CGT GTG AAC ACT CGT
GGT CAG TTC TTT GTG GCA CGT GAA GCG TAT CGC CAC ATG
CGT GAA GGA GGT CGC ATC ATT CTG ACC AGC TCG AAC ACA
GCA TGC GTA AAA GGC GTT CCG AAA CAT GCG GTG TAT TCA
GGC AGC AAA GGG GCC ATT GAC ACG TTT GTG CGC TGT ATG
GCG ATC GAT TGC GGC GAT AAG AAG ATT ACG GTC AAT GCT
GTT GCC CCA GGT GCA ATC AAA ACC GAC ATG TTT CTG GCC
GTG TCT CGC GAA TAC ATT CCC AAT GGG GAA ACG TTC ACC
GAT GAA CAG GTA GAT GAG TGT GCT GCG TGG TTG AGT CCG
CTG AAT CGG GTA GGA TTG CCG GTA GAT GTT GCT CGC GTT
GTG TCC TTT CTG GCG TCT GAT ACC GCC GAA TGG GTG AGC
GGC AAA ATT ATC GGC GTC GAT GGT GGT GCA TTC CGC TAA

Protein sequence of AflM-his

MGHHHHHHHHHHSSGHIDDDDKHMSDNHRLDGKVALVTGAGRGIGAAIAVALGERGA
KVVVNYAHSREAAEKVVEQIKANGTDAIAIQADVGDPEATAKLMAETVRHFGYLDIVSSNA
GIVSFGHLKDVTPEEFDRVFRVNTRGQFFVAREAYRHMREGGRIILTSSNTACVKGVPKHA
VYSGSKGAIDTFVRCMAIDCGDKKITVNAVAPGAIKTDMFLAVSREYIPNGETFTDEQVDEC
AAWLSPLNRVGLPVDVARVVSFLASDTAEWVSGKIIGVDGGAFR

VII. References

[1] Gaussian 09, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian, Inc.*, Wallingford CT, **2009**.

[2] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

[3] M. A. Schätzle, S. Flemming, S. M. Husain, M. Richter, S. Günther, M. Müller, *Angew. Chem.* **2012**, *124*, 2697–2700.

[4] M. A. Schätzle, S. M. Husain, S. Ferlaino, M. Müller, *J. Am. Chem. Soc.* **2012**, *134*, 14742–14745.