Monatshefte für Chemie 137, 163–168 (2006)
DOI 10.1007/s00706-005-0418-4

Monatshefte für Chemie
Chemical Monthly
Printed in Austria

Concerning the Thermal Diastereomerization of the Green Fluorescent Protein Chromophore

Beate Hager, Bettina Schwarzinger, and **Heinz Falk***

Institute of Organic Chemistry, Johannes Kepler University Linz, 4040 Linz, Austria

Received September 28, 2005; accepted October 11, 2005
Published online December 13, 2005

Summary. Two model compounds for the green fluorescent protein chromophore were prepared. One of them incorporates the natural 4-hydroxybenzylidene group of the natural tyrosin derived chromophore, the other one bears a methyl group instead of the hydroxy group. Whereas the photochemically prepared (*E*)-diastereomer of the first compound very effectively reverted thermally (room temperature) to the thermodynamically stable (*Z*)-diastereomer, the (*E*)-diastereomer of the second derivative proved to be stable even at elevated temperatures for more than a day. This finding can be rationalized by constructing the appropriate resonance structures showing that only in the first case an effective delocalization enables partial single bond character of the benzylidene double bond. From the standpoint of chemical etiology, only Nature's choice of the tyrosin derived chromophore of the green fluorescent protein provides an efficient radiationless thermal relaxation channel for the unwanted photo-diastereomerization product formed after excitation besides the dominating fluorescence channel of its chromophore.

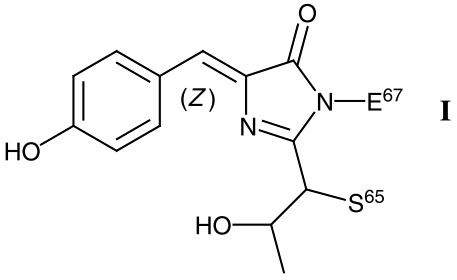
Keywords. Green fluorescent protein; Benzylideneimidazolinones; Thermal barrier; Chemical etiology; Radiationless relaxation.

Introduction

The greenish bioluminescence of the jellyfish *Aequorea victoria* originates from the closely associated calcium(II) binding protein aquorin and the green fluorescent protein (GFP) [1, 2]. Because of its importance as a cloneable reporter of gene expression, GFP has been intensively studied over the past decade and its applications, structural details, and photophysical behavior have been thoroughly reviewed [3]. The fluorescence of GFP is emitted from its (4-hydroxybenzylidene)imidazolinone chromophore **I**, which is autocatalytically generated by a posttranslational intramolecular cyclization of the S^{65}-Y^{66}-E^{67} pre-protein sequence followed by dehydrogenation of the tyrosine (Y) moiety [4]. Interestingly enough, the strong

* Corresponding author. E-mail: heinz.falk@jku.at

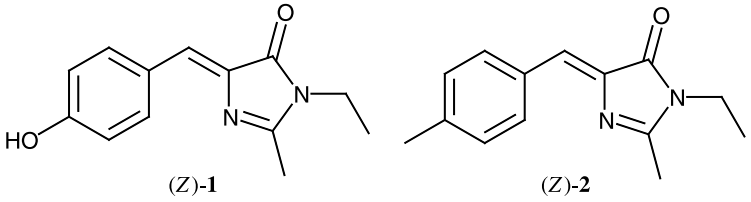
fluorescence of this unit is lost upon denaturation of the protein and is also absent in model compounds of the chromophore, a feature that is attributed to the resulting flexibilization of the chromophore [5].



I

Formula 1

Numerous X-ray studies of the protein and its mutant analogs as well as of model compounds have revealed that the most stable form of the benzylidenic substructure **I** is the (*Z*)-diastereomer [3]. A detailed study [6] has demonstrated that the free energy difference between the (*Z*)-diastereomer and its photoaccessible (*E*)-diastereomer is in the order of $5\,kJ \cdot mol^{-1}$ depending only slightly on the ionization of the chromophore. The thermal activation barrier between the two diastereomers has been estimated to amount to about $50\,kJ \cdot mol^{-1}$, again depending only marginally on the ionization of the chromophore [6]. This latter value is considerably smaller than the values usually found for comparable diastereomerization reactions at exocyclic double bonds of heterocycles, *e.g.* in the dipyrrinones, which are in the order of $100\,kJ \cdot mol^{-1}$ [7]. This rather large difference led us to the question if the low activation barrier encountered is an intrinsic property of the benzylideneimidazolinone chromophoric system or is due to the special choice of Nature of the tyrosine amino acid moiety. Thus, in this study we compare the thermal diastereomerization possibility of the model compound pair (*E*)/(*Z*)-**1**, which is close in constitution to the natural GFP chromophore (*Z*)-**I**, with the model system (*E*)/(*Z*)-**2**, which is devoid of the tyrosine derived 4-hydroxyphenyl group.



(*Z*)-**1** (*Z*)-**2**

Formula 2

Results and Discussions

The (*Z*)-diastereomers of **1** and **2** were prepared by condensation of 4-hydroxy- and 4-methylbenzaldehyde with *N*-acetylglycin followed by an *Erlenmeyer-*

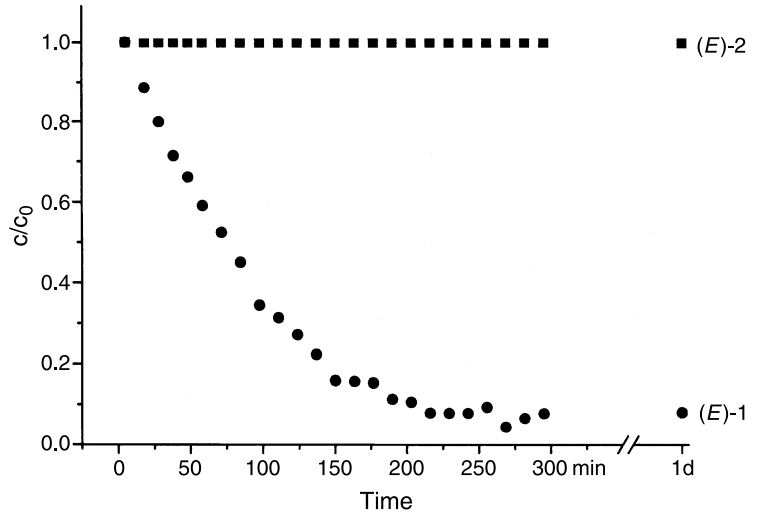


Fig. 1. Thermal relaxation of (*E*)-**1** (25°C) and (*E*)-**2** (25–50°C) in CD_3OD

Plöchl synthesis [8]. Then, the resulting azlactones **3** and **4** were converted to the target azlactames (*Z*)-**1** and (*Z*)-**2** by refluxing with ethylamine according to the procedure of *Kojima et al.* [9]. The (*Z*)-configuration of these compounds was inferred by comparison with similar model compounds of proven configurations [6]. Photodiastereomerizations of the (*Z*)-diastereomers dissolved in methanol into their (*E*)-diastereomers were achieved by irradiation with the light of a high-pressure mercury lamp. The ^1H NMR resonances were assigned straightforwardly and the thermal fate of the initial (*E*)-diastereomers could be easily followed monitoring the nicely separated CH-, CH_3-, and CH_2CH_3-signal intensities in CD_3OD solution.

In these experiments it turned out that the signals of (*E*)-**1** decayed rapidly into those of (*Z*)-**1** even at room temperature following a first order reaction (Fig. 1). In contrast to this finding, (*E*)-**2** remained stable up to the boiling point of methanol for several hours to even a day (Fig. 1). The equilibration between (*Z*)- and (*E*)-**1** led to a free enthalpy difference of about $5\,kJ \cdot mol^{-1}$ consistent with findings on other model systems [6] and calculations [3]. Due to the limited precision of the NMR measurements we refrained from estimating an activation energy. But to determine this value for **1** would, of course, be unnecessary because it had been estimated already for a similar molecule [6]. The important and significant result of the experiment contained in Fig. 1 is that (*E*)-**1** easily reverts to (*Z*)-**1** whereas (*E*)-**2** does not. This result clearly demonstrated that the 4-hydroxy group of **1** is a *conditio sine qua non* for the very easy thermal equilibration providing an efficient thermal relaxation channel of the system.

These findings, unexpected at first sight, might be properly rationalized by constructing the appropriate resonance structures of the conjugated systems of **1** and **2** taking the hydroxy group of **1** (in its different ionization states) into account in the classical manner of organic chemistry as drawn in Fig. 2. Accordingly, resonance structures involving the benzylidene partial structure can be constructed exclusively in the case of the 4-hydroxy group (dissociated or undissociated makes

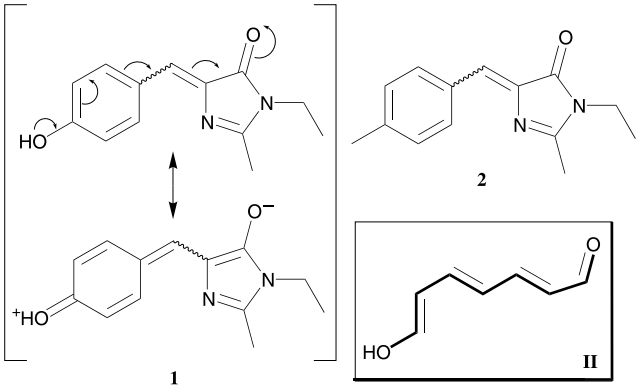


Fig. 2. Resonance structures of (*Z*)-**1** and (*Z*)-**2**

no difference) in **1**. Thus, **1** can be addressed as a vinylogous carboxylic acid or vinylogous carboxylate (**II**) as shown in Fig. 2 yielding partial single bond character and pronounced delocalizability of the benzylidenic double bond. From this result it becomes immediately clear why the ionization state of the 4-hydroxy group is not important [6] for the diasteromerization process: the phenolic –O–H and $-O^{(-)}$ groups behave similarly in providing conjugation in this push–pull electronic system, and thus relay partial single bond character to the benzylidene double bond. This effect is non-existent in the 4-methyl analogue **2**, which behaves as a benzylidene-imidazolidinone, *i.e.* like a ''normal'' heterocyclic compound bearing an exocyclic double bond like the dipyrrinones [7].

The result of the present investigation reminds one vividly of the very smooth thermal diastereomerization encountered at the exocyclic double bond in position 5 of the 2,3-dihydrobilin-1,19-diones (the chromophore of *e.g.* phytochrome) as compared to the normal thermal stability of this double bond in the bilin-1,19-diones (as in *e.g.* biliverdin) [10]. Only the partially saturated system of the 2,3-dihydrobilin-1,19-diones compares to the easily delocalizable and accordingly thermally easily diastereomerizable bond system of an *N*-acylenamine.

In conclusion, in the sense of chemical etiology Nature has been very wise to select the tyrosine unit for constructing the chromophore of GFP instead choosing *e.g.* a phenylalanine moiety. Obviously, as derived above only the tyrosine derived chromophore is amenable to display the properties of a vinylogous carboxylic acid or vinylogous carboxylate partial structure needed for an efficient thermal relaxation channel and the necessary flexibility to adapt to the needs of the functional unit of the native GFP. In particular, the easy thermal diastereomerization allows for an efficient radiationless thermal relaxation of the unwanted photo-diastereomerization product ((*E*)-**I**) that might be formed concomitantly after excitation ((*Z*)-**I** → (*E*)-**I**) besides the dominating fluorescence channel of the chromophore.

Experimental

Solvents were of *p.a.* quality unless stated otherwise. Photodiastereomerizations of (*Z*)-**1** and (*Z*)-**2** were executed by means of a Hanau TQ 150Z2 UV-lamp directly in the NMR-tube in *Me*OH-d_4 solutions for 1 h. NMR spectra were recorded on a Bruker Avance DPX 200 MHz spectrometer and DRX 500 MHz spectrometer using a TXI cryoprobe with *z*-gradient coil. 2D NMR experiments were

performed using standard pulse sequences as provided by the manufacturer. Typical 90° hard pulse durations were 8.2 μs (^{1}H) and 16.6 μs (^{13}C), 90° pulses in decoupling experiments were set to 67 μs. HSQC and HMBC experiments were optimized for coupling constants of 145 Hz for single quantum correlations and 10 Hz for multi-bond correlations. The N*O*ESY mixing time was set to 400 ms. Mass spectra were recorded on a Hewlett Packard 5989 quadrupole instrument.

(Z)-1-Ethyl-4-(4-hydroxybenzylidene)-2-methyl-1H-imidazol-5(4H)-one ((*Z*)-**1**, $C_{13}H_{14}N_2O_2$)
To a solution of 0.3 g **3** (1.2 mmol) dissolved in 15 cm^3 *Et*OH containing 0.3 g K_2CO_3 (2.2 mmol), 0.15 cm^3 $C_2H_5NH_2$ (70% in H_2O) (1.9 mmol) were added and the resulting mixture was heated to reflux for 4 h. After cooling the K_2CO_3 was filtered off, the solution was concentrated *in vacuo*, and the crude product was purified by silica gel chromatography using $CHCl_3$/*Me*OH (10/1, *v*/*v*) as the developing solvent to give 84 mg (Z)-**1** (30% yield). TLC: $R_f = 0.6$ ($CHCl_3$:*Me*OH = 10:1); ^{1}H NMR (500 MHz, *Me*OD-d_3, 30°C): $\delta = 1.23$ (t, $J = 7.14$ Hz, –CH_2–C*H*$_3$), 2.38 (s, –CH_3), 3.68 (q, $J = 7.14$ Hz, –CH_2–), 6.84 (d, $J = 8.51$ Hz, –H3′), 7.00 (s, –CH), 8.00 (d, $J = 8.51$ Hz, –H2′) ppm; N*O*ESY (*Me*OD-d_3): –H2′ → –CH and –H3′, –H3′ → –H2′, –CH_2– → –CH_2–C*H*$_3$, –CH_2–C*H*$_3$ → –CH_2–; ^{13}C NMR (125 MHz, *Me*OD-d_3, 30°C): $\delta = 14.6$ (–CH_2–*C*H$_3$), 15.3 (–CH_3), 36.4 (–CH_2–), 116.8 (C3′), 126.9 (–CH), 129.2 (C1′), 135.6 (C2′), 137.1 (C4), 161.5 (C4′), 163.0 (C2), 172.2 (C5) ppm; HMBC (*Me*OD-d_3): CH_2–C*H*$_3$ → C2, C5, and –CH_2–, –CH_3 → C2, C1′ (weak), and C4 (weak), –CH_2– → C5, C2, and –CH_2–*C*H$_3$, H3′ → C4′, C2′, –CH, and C3′, –CH → C5, C2′, and C3′ (weak), H2′ → C4′, C2′, C1′, –CH, and C3′; HSQC data were according to structure; ESI-MS (*Me*OH + 1 vol-% HCOOH, $\gamma \sim 1$ mg cm^{-3}, positive ion mode): $m/z = 231$ ([M − H]$^+$).

(E)-1-Ethyl-4-(4-hydroxybenzylidene)-2-methyl-1H-imidazol-5(4H)-one ((*E*)-**1**, $C_{13}H_{14}N_2O_2$)
^{1}H NMR (200 MHz, *Me*OD-d_3, 30°C): $\delta = 1.24$ (t, $J = 7.14$ Hz, –CH_2–C*H*$_3$), 2.35 (s, –CH_3), 3.68 (q, $J = 7.14$ Hz, –CH_2–), 6.84 (d, $J = 8.56$ Hz, –H3′), 7.15 (s, –CH), 8.25 (d, $J = 8.56$ Hz, –H2′) ppm.

(Z)-1-Ethyl-4-(4-methylbenzylidene)-2-methyl-1H-imidazol-5(4H)-one ((*Z*)-**2**, $C_{14}H_{16}N_2O$)
Compound (*Z*)-**2** was prepared from **4** according to (*Z*)-**1** in 15% yield. ^{1}H NMR (500 MHz, *Me*OD-d_3, 30°C): $\delta = 1.33$ (t, $J = 7.14$ Hz, –CH_2–C*H*$_3$), 2.11 (s, –CH_3), 2.36 (s, ar–CH_3), 4.26 (q, $J = 7.14$ Hz, –CH_2–), 7.23 (d, $J = 7.41$ Hz, –H3′), 7.39 (s, –CH), 7.49 (d, $J = 7.41$ Hz, –H2′) ppm; N*O*ESY (*Me*OD-d_3): –H3′ → ar–CH_3 and –H2′, –CH_2– → –CH_2–C*H*$_3$; ^{13}C NMR (125 MHz, *Me*OD-d_3, 30°C): $\delta = 14.5$ (–CH_2–*C*H$_3$), 21.4 (ar–CH_3), 22.4 (–CH_3), 62.5 (–CH_2–), 126.1 (C4), 130.4 (C3′), 131.0 (C2′), 132.0 (C1′), 135.4 (–CH), 141.3 (C4′), 166.8 (C5), 173.2 (C2) ppm; HMBC (*Me*OD-d_3): C*H*$_3$ → C2, ar–CH_3 → C4′ and C3′, H3′ → C4′, C1′, and –CH, –CH → C5 and C2′, H2′ → C4′, –CH, C3′, and C1′ (weak); HSQC data were according to structure; ESI-MS (*Me*OH + 1 vol-% HCOOH, $\gamma \sim 1$ mg cm^{-3}, positive ion mode): $m/z = 229$ ([M − H]$^+$). This diastereomer proved to be configurationally stable in H_2O solution up to 100°C.

(E)-1-Ethyl-4-(4-methylbenzylidene)-2-methyl-1H-imidazol-5(4H)-one ((*E*)-**2**, $C_{14}H_{16}N_2O$)
^{1}H NMR characterization was only possible for a mixture of the (*Z*)-**2**:(*E*)-**2** diastereomers (ratio 3:1), which resulted in a few overlaps, in particular in the aromatic region. ^{1}H NMR (500 MHz, *Me*OD-d_3, 30°C): $\delta = 1.51$ (t, $J = 7.14$ Hz, –CH_2–C*H*$_3$), 2.05 (s, –CH_3), 2.33 (s, ar–CH_3), 4.15 (q, $J = 7.14$ Hz, –CH_2–), 7.05–7.30 (m, 5H, no exact assignment possible due to an overlap with the signals of (*Z*)-**2**) ppm. This diastereomer proved to be configurationally stable in H_2O solution up to 100°C, but reverted partially to (*Z*)-**2** under this condition.

(Z)-4-(4-Acetoxybenzylidene)-2-methyloxazol-5(4H)-one (**3**, $C_{13}H_{11}NO_4$)
N-Acetylglycine, 4.6 g (40 mmol), 5.6 g 4-hydroxybenzaldehyde (46 mmol), and 2.6 g anhydrous CH_3COONa (30 mmol) were added to 20 cm^3 $(CH_3CO)_2O$ and the resulting mixture was heated to reflux for 2 h. The solution was cooled with ice and then poured into 100 cm^3 ice-H_2O. The precipitate

was collected by filtration and washed with small amounts of *Et*OH giving 5.8 g **3** (59%). ^{1}H NMR (200 MHz, *Me*OD-d_3, 30°C): $\delta = 2.30$ (s, –CO–CH_3), 2.40 (s, –CH_3), 7.14 (s, –CH), 7.21 (d, $J = 8.56$ Hz, –H3′), 8.19 (d, $J = 8.56$ Hz, –H2′) ppm.

(Z)-4-(4-Methylbenzylidene)-2-methyloxazol-5(4H)-one (**4**, $C_{12}H_{11}NO_2$)
Compound **4** was prepared according to **3** in 20% yield. ^{1}H NMR (500 MHz, *DMSO*-d_6, 30°C): $\delta = 2.36$ (s, –CO–CH_3), 2.39 (s, –CH_3), 7.18 (s, –CH), 7.31 (d, $J = 7.68$ Hz, –H3′), 8.08 (d, $J = 7.68$ Hz, –H2′) ppm.

Acknowledgements

The cryogenic 500 MHz probe used was purchased from FWF project P15380 (project leader: Prof. Dr. *N. Müller*). We are grateful to DI. *M. Waser* for recording the 500 MHz NMR-spectra, to Prof. Dr. *C. Klampfl* for his support in recording of MS and to *V. Romanova* for assistance in preparative work.

References

[1] Shimomura O (1979) FEBS Lett **104**: 220
[2] Charbonneau H, Walsh KA, McCann RO, Prendergast FG, Cormier MJ, Vanaman TC (1985) Biochem **24**: 6762
[3] Zimmer M (2002) Chem Rev **102**: 759
[4] Heim R, Cubitt AB, Tsien RY (1995) Nature **373**: 663; Niwa H, Inouye S, Hirano T, Matsuno T, Kojima S, Kubota M, Ohashi M, Tsuji FI (1996) Proc Natl Acad Sci **93**: 13617
[5] Ward WW, Bokman SH (1982) Biochem **21**: 4535; Webber NM, Litvinenko KL, Meech SR (2001) J Phys Chem B **105**: 8036
[6] He X, Bell AF, Tonge PJ (2003) FEBS Lett **549**: 35
[7] Falk H, Grubmayr K, Hofer O, Neufingerl F, Ribó JM (1976) Monatsh Chem **107**: 831
[8] Plöchl J (1883) Ber Dtsch Chem Ges **16**: 2815
[9] Kojima S, Ohkawa H, Hirano T, Maki S, Niwa H, Ohashi M, Inouye S, Tsuji FI (1998) Tetrahedron Lett **39**: 5239
[10] Falk H (1989) The Chemistry of Linear Oligopyrroles and Bile Pigments, Springer Wien New York, p 521