Monatshefte für Chemie 138, 859–862 (2007)
DOI 10.1007/s00706-007-0738-7
Printed in The Netherlands

Monatshefte für Chemie
Chemical Monthly

The Thermal Diastereomerization of the Tryptophane-Derived Green Fluorescent Protein Chromophore[#]

Karoline Fendler, Beate Hager, and **Heinz Falk***

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

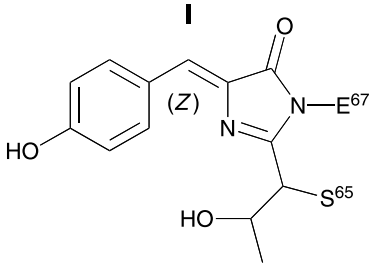
Received May 30, 2007; accepted (revised) May 31, 2007; published online July 20, 2007

Summary. Two model compounds for the tryptophane variant of the green fluorescent protein chromophore containing a 3-indolyl and 2-pyrrolyl moiety were prepared. For the first one the (*Z*)-diastereomer was found to be more stable than the (*E*)-diastereomer by 5.7 kJ mol^{-1}. It could be photo-diastereomerized and its thermal equilibration was studied, whereas the second one underwent photo-destruction. From an *Arrhenius* plot an activation barrier for the (*E*) to (*Z*) diastereomerization of 85.6 kJ mol^{-1} could be determined. Thus, it could be demonstrated that in contrast to the corresponding phenyl derivative studied recently the tyrosine- and tryptophane-derived chromophores of the green fluorescent protein are amenable to fast thermal diastereomerization, which is of fundamental importance for the fluorescence and photoswitching processes in the corresponding proteins.

Keywords. Green fluorescent protein; Indolylideneimidazolinones; Thermal barrier; Photoswitching; Radiationless relaxation.

Introduction

Two model compounds for the green fluorescent protein (GFP) chromophore **I** [1] have been prepared recently [2]. One of them (**1**) comprises the natural 4-hydroxybenzylidene group of the natural tyrosin-derived chromophore, the other one (**2**) bears a methyl group instead of the hydroxy group. Whereas the photochemically prepared (*E*)-diastereomer of **1** reverted spontaneously (room temperature) to the thermodynamically stable (*Z*)-diastereomer, the (*E*)-diastereomer of **2** proved to be stable even at elevated temperatures for an extended period [2].

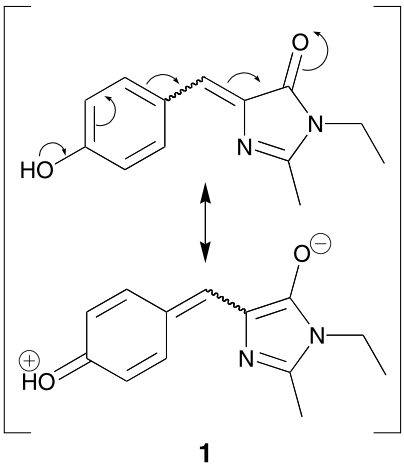


This finding has been rationalized by constructing the appropriate resonance structures as exemplified by those of **1** in Fig. 1 showing that exclusively for **1** an effective delocalization of the oxygen lone pair enables a partial single bond character of the benzylidene double bond, and thus, a significantly reduced activation barrier of the exocyclic double bond at the imidazolinone ring, which is in the order of 50 kJ mol^{-1} [2, 7]. Accordingly, 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. But this behavior is also of high importance for the recently introduced area [3] of the reversible photoswitching of GFP-like proteins, such as the prototype asFP595, isolated from *Anemonia sulcata* [4]. These proteins can be reversibly pho-

[#] Dedicated to the memory of Prof. Dr. *Karl Schlögl*
* Corresponding author. E-mail: heinz.falk@jku.at

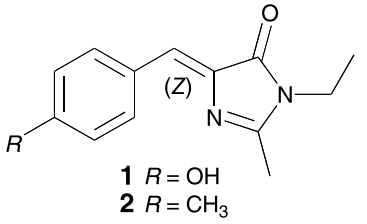
toswitched between a fluorescent and a non-fluorescent state, which make them very promising in several areas of sciences as *e.g.* in nano-scale resolution in far-field fluorescence microscopy or in erasable three dimensional data storage. The thermal relaxation between the photodiastereomers is a fundamental step in the corresponding photo-cycles.

To further prove our hypothesis of lowering the diastereomerization activation barrier by lowering the double bond character of the benzylidenic double bond, we now report an investigation of the thermal diastereomerization behavior of aryl derivatives, which are characterized like **1**, by potential lone pair delocalization involving the exocyclic double bond of the imidazolone partial structure. Moieties with such delocalization potential, like the 3-indolylmethylidene chromophore have been already characterized in GFP mutants [5].

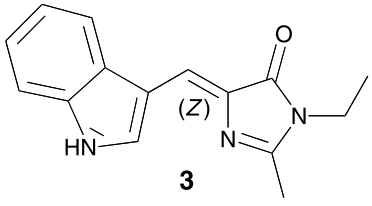


1

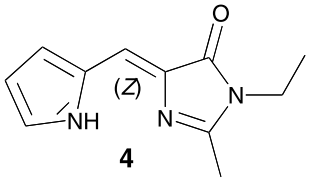
Fig. 1. Mesomeric structures of **1** showing the lone pair delocalization from the phenolic oxygen atom, which leads to a partial double bond character and thus lowered diastereomerization barrier of the exocyclic imidazolinone double bond



1 $R = OH$
2 $R = CH_3$



3



4

Results and Discussion

The (*Z*)-diastereomers of **3** and **4** were prepared by first reacting indolyl-3-carbaldehyde and pyrrol-2-carbaldehyde with *N*-acetylglycine followed by an *Erlenmeyer-Plöchl* synthesis [6]. Then, the resulting azlactones **5** and **6** were converted to the target azlactames (*Z*)-**3** and (*Z*)-**4** by refluxing with ethylamine in analogy to the procedure of *Heim et al.* [5]. It might be mentioned that the synthesis of the corresponding 4-imidazolylidene derivative proved to be unsatisfactory at the azlactame step – only the lactone (*Z*)-**7** could be obtained. The (*Z*)-configuration of these compounds was deduced by comparison with similar model compounds of proven configurations [2, 7]. Photo-diastereomerization of the (*Z*)-diastereomer dissolved in methanol into its (*E*)-diastereomer could be achieved by irradiation with the light of a high-pressure mercury lamp only in the case of (*Z*)-**3**, whereas in the case of the pyrrolyl derivative (*Z*)-**4** complete photo-destruction, as monitored by ^{1}H NMR and mass spectrometry, occurred. The ^{1}H NMR resonances were assigned straightforwardly (see Experimental part) and the thermal fate of the initial (*E*)-diastereomer (*E*)-**3** was easily followed monitoring the nicely separated CH-signal intensities in CD_3OD solution.

From the temperature dependence of the thermal diastereomerization of (*E*)-**3** the thermodynamic stabilities of the two diastereomers resulting from the equilibration and the activation barrier using an *Arrhenius* plot could be deduced. With respect to the thermodynamic stability the less sterically congested diastereomer (*Z*)-**3** is clearly favoured by $5.7 \pm 0.5\,kJ\,mol^{-1}$ – a feature, which is also nicely reproduced by force-field and AM1 calculations, from which values of 5.2 and $3.5\,kJ\,mol^{-1}$ were obtained. The thermal diastereomerization barrier of (*E*)-**3** to (*Z*)-**3** deduced from the *Arrhenius* diagram of Fig. 2 amounts to $85.6 \pm 1\,kJ\,mol^{-1}$, a value, which allows for fast thermal equilibration at room temperature

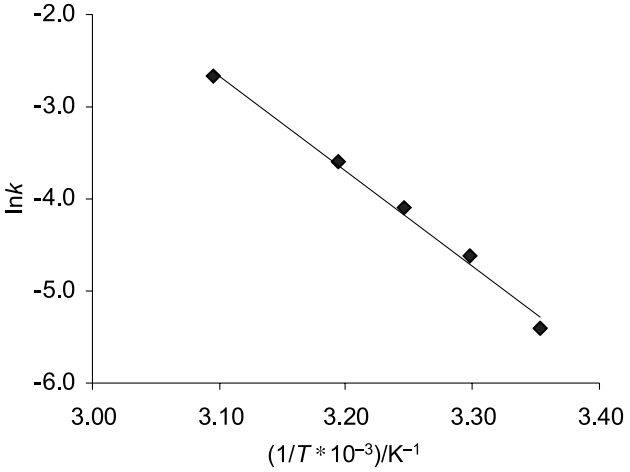
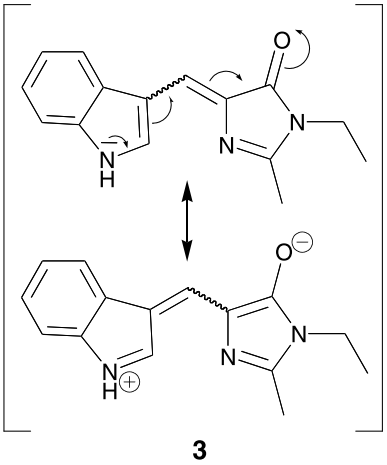


Fig. 2. *Arrhenius* diagram for the thermal diasteromerization of (*E*)-**3** to (*Z*)-**3**

and is similar to the one estimated for (*E*)-**1** (in the order of 50 kJ mol^{-1}). But this value is significantly different from the one encountered for (*E*)-**2**, which is according to the high thermal stability of this diastereomer in the order of $\gg$100 kJ mol^{-1} [2].

From these results it becomes obvious that in addition to the tyrosine-derived GFP chromophore also the tryptophane-derived chromophore is amenable to room temperature thermal equilibration of its photochemically produced (*E*)-configured diastereomer. As in the case of the tryptophane-derived (*E*)-**1** this easy thermal diastereomerization process can be attributed to an effective delocalization of the indole nitrogen lone pair, which enables a partial single bond character of the exocyclic double bond of the imidazolone partial structure as indicated in Fig. 3 in close analogy to Fig. 1. Accordingly, the tryptophane-derived chromophore in GFPs should be also amenable to photoswitching processes [3] as found for the tyrosine-derived one.



3

Fig. 3. Mesomeric structures of **3** showing the lone pair delocalization from the indole nitrogen atom, which leads to a partial double bond character and thus lowered diastereomerization barrier of the exocyclic imidazolinone double bond

In conclusion, we were able to further prove our hypothesis [2] that the exocyclic double bond at the imidazolinone partial structure of GFP chromophores can be easily made amenable to thermal diastereomerization due to a lowered thermal diastereomerization barrier if it can adopt a partial double bond character by means of delocalization of a conjugated lone pair. This lone pair could originate from the phenolic oxygen as in the tyrosine-derived unit or from the indole nitrogen of a tryptophane-derived chromophore.

Experimental

Solvents were of p.a. quality. Photo-diastereomerizations of (*Z*)-**3** and (*Z*)-**4** were executed by means of a Hanau TQ 150Z2 UV-lamp directly in the NMR-tube in *Me*OD-d_3-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 mixing times were set to 400 ms for N*O*ESY and 60 ms for TOCSY experiments with MLEV 17 – spinlock sequence at 5 kHz RF-field strength. Mass spectra were recorded on Hewlett Packard 5987 quadrupole and Thermo Finnigan LCQ Deca XP-plus instruments. IR and UV-Vis spectra were recorded using the Bruker Tensor 27 and Varian Cary 100 Bio UV-Vis spectrometers. Force field and AM1 calculations were executed using the commercial programs pcmodel 9 and AM1.

(Z)-1-Ethyl-4-(3-indolylmethylidene)-2-methyl-1H-imidazol-5(4H)one ((*Z*)-**3**, $C_{15}H_{15}N_3O$)

To a solution of 2.58 g (*Z*)-**5** (9.6 mmol) in 120 cm^3 *Et*OH containing 2.43 g K_2CO_3 (17.6 mmol) 1.2 cm^3 $C_2H_5NH_2$ (70% in H_2O) (15.2 mmol) were added and the resulting mixture was heated to reflux for 4 h. After cooling 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 (50:1, *v*:*v*) as the developing solvent to give 1.16 g (*Z*)-**3** (47% yield). Mp 72°C; TLC: $R_f = 0.4$ ($CHCl_3$:*Me*OH = 15:1); ^{1}H NMR (500 MHz, *Me*OD-d_3, 30°C): $\delta =$ 1.26 (t, $J = 7.32$ Hz, –CH_2–CH_3), 2.42 (s, –CH_3), 3.72 (q, $J = 7.32$ Hz, –CH_2–), 7.21 (m, ar-H5,6), 7.44 (d, $J = 7.02$, ar-H7), 7.54 (s, –CH=), 7.96 (d, $J = 7.63$, ar-H4), 8.38 (s, ar-H2) ppm; N*O*ESY (*Me*OD-d_3): –CH_2– $\rightarrow$ –CH_2–CH_3,

$-CH_2-CH_3 \rightarrow -CH_2-$, ar-H5,6 → ar-H4 and ar-H7, ar-H4 → ar-H2 and ar-H5,6; ^{13}C NMR (125 MHz, *Me*OD-d_3, 30°C): $\delta = 14.7$ ($-CH_2-CH_3$), 15.1 ($-CH_3$), 36.4 ($-CH_2-$), 112.5 (ar-C7a), 113.0 (ar-C7), 119.7 (ar-C4), 122.3 (ar-C5), 123.1 (–CH=), 124.0 (ar-C6), 128.8 (ar-C3), 133.7 (ar-C2), 134.6 (C4), 138.0 (ar-C3a), 160.1 (C2), 171.6 (C=O) ppm; HMBC (*Me*OD-d_3): $-CH_3 \rightarrow$ C2, ar-H5,6 → ar-C3a, ar-C3, and ar-C7a, ar-H7 → ar-C5, ar-C6, ar-C4 (weak), ar-C3 (weak), and ar-C3a (weak), –CH= → ar-C3, ar-C2, and C=O, ar-H4 → ar-C5, ar-C7a, and ar-C3 (weak), ar-H2 → ar-C7a, ar-C3, ar-C3a, and –CH=; HSQC data were according to structure; ESI-MS (*Me*OH + 1 vol.% HCOOH, $\gamma \sim 1\,mg\,cm^{-3}$, positive ion mode): $m/z = 254$ ($[M-H]^+$); IR (KBr): $\bar{\nu} = 3223$, 2928, 1679, 1630, 1552, 1512, 1442, 1401, 1370, 1347, 1252, 1229, 1128, 749 cm^{-1}; UV-Vis (methanol, $c = 2.10^{-5}\,mol\,dm^{-3}$): λ_{max} (ε) = 220 (16350), 281 (7300), 405 (16300) nm ($dm^3\,mol^{-1}\,cm^{-1}$).

(E)-1-Ethyl-4-(3-indolylmethylidene)-2-methyl-1H-imidazol-5(4H)-one ((*E*)-**3**, $C_{15}H_{15}N_3O$)
Prepared by photo-diastereomerization from (*Z*)-**3**. ^1H NMR (200 MHz, *Me*OD-d_3, 30°C): $\delta = 1.29$ (t, $J = 7.32$ Hz, $-CH_2-CH_3$), 2.39 (s, $-CH_3$), 3.76 (q, $J = 7.32$ Hz, $-CH_2-$), 7.00 (m, ar-H5,6), 7.46 (m, ar-H7), 7.67 (s, –CH=), 7.83 (d, J = 7.63, ar-H4), 9.33 (s, ar-H2) ppm; UV-Vis (methanol, $c = 2.10^{-5}\,mol\,dm^{-3}$) a photoequilibration of (*Z*)-**3** indicated a small bathochromic shift of the long wavelength band to $\lambda_{max} = 407$ nm with a slight decrease in its intensity.

(Z)-1-Ethyl-4-(2-pyrrolylmethylidene)-2-methyl-1H-imidazol-5(4H)one ((*Z*)-**4**, $C_{11}H_{13}N_3O$)
Compound (*Z*)-**4** was prepared from (*Z*)-**6** according to (*Z*)-**3** but without silica gel chromatography in 66% yield. Mp 76°C; ^1H NMR (500 MHz, *Me*OD-d_3, 30°C): $\delta = 1.24$ (t, J = 7.02 Hz, $-CH_2-CH_3$), 2.79 (s, $-CH_3$), 3.45 (q, $J = 7.02$ Hz, $-CH_2-$), 6.78 (d, $J = 2.9$ Hz, ar-H5), 7.0 (t, $J = 2.9$ Hz, ar-H4), 7.61 (s, ar-H3), 8.02 (s, –CH=) ppm; N*O*ESY (*Me*OD-d_3): $-CH_2- \rightarrow -CH_2-CH_3$, $-CH_2-CH_3 \rightarrow -CH_2-$, ar-H4 → ar-H3 and ar-H5; ^{13}C NMR (125 MHz, *Me*OD-d_3, 30°C): $\delta = 15.2$ ($-CH_2-CH_3$), 20.8 ($-CH_3$), 35.3 ($-CH_2-$), 105.9 (ar-C5), 112.7 (–CH=), 114.2 (ar-C3), 118.2 (ar-C4), 133.2 (ar-C2), 133.6 (C4), 147.0 (C2), 166.9 (C=O) ppm; HMBC (*Me*OD-d_3): $-CH_2- \rightarrow$ C=O and $-CH_2-CH_3$, ar-H5 → –CH=, ar-C3, ar-C4, ar-C2, and C4, ar-H4 → ar-C3, ar-C5, ar-C2, and C4, ar-H3 → ar-C5, ar-C4, –CH=, ar-C2 and C4; HSQC data were according to structure; TOCSY (*Me*OD-d_3): –CH= → ar-H3; ESI-MS (*Me*OH + 1 vol.% HCOOH, $\gamma \sim 1\,mg\,cm^{-3}$, positive ion mode): $m/z = 204$ ($[M-H]^+$); IR (KBr): $\bar{\nu} = 3390$, 3277, 3131, 2972, 1653, 1555, 1433, 1387, 1362, 1293, 1236, 744 cm^{-1}; UV-Vis (methanol, $c = 2.10^{-5}\,mol\,dm^{-3}$): λ_{max} (ε) = 195 (2750), 224 (4850), 249 (5250), 295 (3225), 305 (2850), 344 (850) nm ($dm^3\,mol^{-1}\,cm^{-1}$).

(Z)-4-(3-Acetindolylmethylidene)-2-methyloxazol-5(4H)one ((*Z*)-**5**, $C_{15}H_{12}N_2O_3$)
Indol-3-carbaldehyde 3.34 g (23.0 mmol), 2.34 g *N*-acetylglycine (20.1 mmol), and 1.23 g anhydrous CH_3COONa (15.0 mmol) were added to 12 cm^3 $(CH_3CO)_2O$ and the resulting mixture was heated to reflux for 2 h. The solution was cooled with ice and poured into 100 cm^3 ice-H_2O. The precipitate was collected by filtration and washed with small amounts of *Et*OH giving 2.84 g (*Z*)-**5** (53% yield). Mp 198°C; ^1H NMR (200 MHz, *DMSO*-d_6, 30°C): $\delta = 2.43$ (s, $-CH_3$), 2.74 (s, $-(CO)-CH_3$), 7.42 (m, ar-H5,6), 7.51 (s, –CH=), 8.36 (m, ar-H4,7), 8.72 (s, ar-H2) ppm.

(Z)-4-(3-Acetpyrrolylmethylidene)-2-methyloxazol-5(4H)one ((*Z*)-**6**, $C_{11}H_{10}N_2O_3$)
Compound (*Z*)-**6** was prepared from pyrrol-2-carbaldehyde analogously to (*Z*)-**5** and purified by silica gel chromatography using $CHCl_3$:ethylacetate (3:1) in 15% yield. Mp 206°C; ^1H NMR (200 MHz, *DMSO*-d_6, 30°C): $\delta = 2.36$ (s, $-CH_3$), 2.65 (s, $-(CO)-CH_3$), 6.52 (t, $J = 3.09$ Hz, ar-H4), 7.63 (d, J = 3.09 Hz, ar-H5), 7.79 (d, ar-H3), 8.00 (s, –CH=) ppm.

(Z)-4-(2-Imidazolylmethylidene)-2-methyloxazol-5(4H)-one ((*Z*)-**7**, $C_8H_7N_3O_2$)
Compound (*Z*)-**7** was prepared from imidazol-2-carbaldehyde according to (*Z*)-**5** and purified by silica gel chromatography using $CHCl_3$:*Me*OH (2:1) in 15% yield. Mp 265°C; ^1H NMR (200 MHz, *DMSO*-d_6, 30°C): $\delta = 2.83$ (s, $-CH_3$), 7.52 (s, ar-H5), 7.94 (s, ar-H3), 8.53 (s, –CH=) ppm.

Acknowledgements

The cryogenic 500 MHz probe used was purchased from FWF project P15380 (project leader: *N. Müller*). We are grateful to Prof. Dr. *C. Klampfl*, and D. I. *U. Karner* for recording of mass spectra. Advice of Dr. *C. Etzlstorfer* with respect to the force field and AM1 calculations is gratefully acknowledged.

References

[1] For a review see: Zimmer M (2002) Chem Rev **102**: 759
[2] Hager B, Schwarzinger B, Falk H (2006) Monatsh Chem **137**: 163
[3] Schäfer LV, Groenhof G, Klingen AR, Ullmann GM, Boggio-Pasqua M, Robb MA, Grubmüller H (2007) Angew Chem **119**: 536
[4] Andresen M, Wahl MC, Stiel AC, Gräter F, Schäfer LV, Trowitzsch S, Weber G, Eggeling C, Grubmüller H, Hell SW, Jakobs S (2005) Proc Natl Acad Sci USA **102**: 13070
[5] a) Heim R, Prasher DC, Tsien RY (1994) Proc Natl Acad Sci USA **91**: 12501; b) Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY (1995) TIBS **20**: 448; c) Heim R, Tsien RY (1996) Curr Biol **6**: 178; d) Kojima S, Ohkawa H, Hirano T, Maki S, Niwa H, Ohashi M, Inouye S, Tsuji FI (1998) Tetrahedron Lett **39**: 5239
[6] Plöchl J (1883) Ber Dtsch Chem Ges **16**: 2815
[7] He X, Bell AF, Tonge PJ (2003) FEBS Lett **549**: 35