Mutagenic Stabilization of the Photocycle Intermediate of Green Fluorescent Protein (GFP)

Jens Wiehler,[b] Gregor Jung,[a] Christian Seebacher,[a] Andreas Zumbusch,*[a] and Boris Steipe[b]

The optical spectra of the Aequorea victoria *green fluorescent protein (GFP) are governed by an equilibrium between three different chromophore states. Mutants that predominantly show either the protonated (A) or the deprotonated (B) form of the chromophore have previously been described. In contrast, the I form, which is formed by rapid excited-state deprotonation of the A form of the chromophore, has only been described as an obligatory photochemical intermediate. We report the design of a new GFP mutant with a stabilized I form. For this purpose, we introduced two isosteric point mutations, Thr203Val and Glu222Gln, that selectively raise the potential energy of both the A and the B form. Knowledge of the absorption spectrum of the I form at room temperature allows the detailed analysis of concentration dependent changes in bulk wild-type(wt)-GFP spectra, as well as the determination of the dimerization constant of GFP. This information expands the use of GFP to that of a spectral probe for protein concentration. We determined energy differences between the chromophore ground states in the monomer and the dimer and reconstructed part of the potential energy surface.*

Introduction

Green fluorescent protein (GFP) has been established as one of the most important in vivo markers for gene expression and protein localization.[1] This protein is unique in that its fluorophore, *p*-hydroxybenzylideneimidazolinone, is formed from the tripeptide Ser65-Tyr66-Gly67, so that the protein fluoresces without the addition of an external cofactor.[2, 3] Crystallographic studies have revealed that the chromophore is located in the middle of an α helix that runs coaxially through the core of a rigid β barrel scaffold structure formed from 11 β strands.[4, 5]

Wild-type(wt)-GFP, which is found in high concentrations in organelles of the Pacific jellyfish *Aequorea victoria*,[6] exhibits two absorption bands in the visible spectral region. Based on X-ray structural data, an interconversion between three different forms of the chromophore was postulated (Scheme 1).[7, 8] A band at $\lambda_{abs} = 397$ nm characterizes the A form of the chromophore and was attributed to a neutral phenol group (RH), while the B form, which has an anionic phenolate group (R^-), absorbs at $\lambda_{abs} = 476$ nm.[2, 3] The third form, the so-called I form, is an anionic form, produced after rapid excited-state deprotonation of the neutral A form.[9, 10] Accordingly, excitation of either the A or the B form results in emission from R^- centered at $\lambda_{em} = 508$ nm and $\lambda_{em} = 503$ nm, respectively. In the I form, the immediate protein environment and its hydrogen bonding network are still adapted to the protonated chromophore. The excited-state A* form quantitatively deprotonates into the I* form, which, after emission, becomes the I form. As the subsequent reprotonation of the chromophore is fast, the I form was only considered to be a transient state in the photocycle of wt-GFP at room temperature. However, this form of the chromophore has been detected by optical spectroscopy after being trapped at cryogenic temperatures, and by time-resolved spectroscopy at room temperature.[11–14] Although the I form was found to only persist below 100 K, it is of central importance for the understanding of the potential energy landscape of GFP.

Herein we show how site-directed mutagenesis leading to the double mutant Thr203Val/Glu222Gln can be used to stabilize the I form of GFP. The mutagenesis strategy was based on results from investigation of the two single mutants Thr203Val and Glu222Gln.[20, 28] While many mutants that predominantly show either the A or the B form of the chromophore have been produced for the application of GFP as a fluorescent marker, no mutant with an I form stabilized at room temperature has been reported to date.[1] By using the new mutant, we obtained room-temperature spectra of the I form. Surprisingly, the spectra of the I form perfectly coincide with a spectral component of wt-GFP observed at high protein concentrations. In fact, the wt-GFP spectrum can be quantitatively explained as the sum of the

[a] *Dr. A. Zumbusch, Dr. G. Jung,[+] Dr. C. Seebacher*
Department Chemie, Ludwig-Maximilians-Universität
Butenandtstraße 11, 81377 München (Germany)
Fax: (+49) 89-2180-77545
E-mail: andreas.zumbusch@cup.uni-muenchen.de

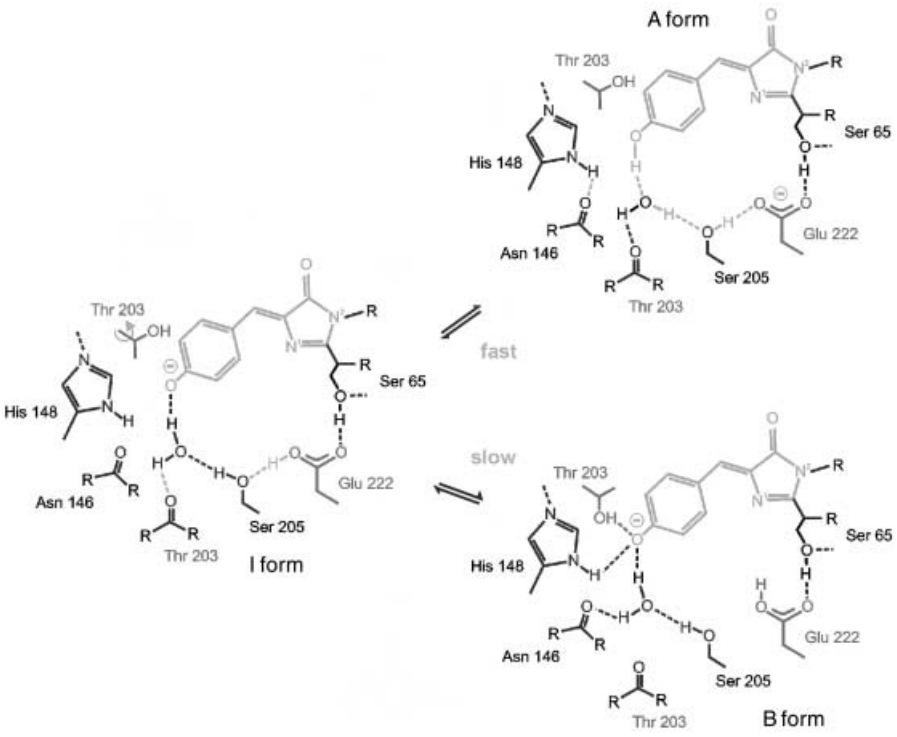
[+] *Current address:*
Department of Chemistry, University of California
B77 Hildebrand Hall, Berkeley, CA 94720 (USA)

[b] *Dr. J. Wiehler,[++] Prof. Dr. B. Steipe[+++]*
Genzentrum der Ludwig-Maximilians-Universität
Feodor-Lynen Straße 25, 81377 München (Germany)

[++] *Current address:*
Biomax Informatics AG
Lochhamer Straße 11, 82152 Martinsried (Germany)

[+++] *Current address:*
Department of Biochemistry, University of Toronto
1 King's College Circle, Toronto, Ontario M5S 1A8 (Canada)

 © 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim DOI: 10.1002/cbic.200300595



Scheme 1. Scheme showing the A (λ_{abs} = 397 nm), B (λ_{abs} = 476 nm, λ_{em} = 503 nm), and I (λ_{abs} = 499 nm, λ_{em} = 508 nm) forms of GP and mechanisms of conversion and photoisomerization based on refs. [7, 8]. The mutated side chains Thr203 and Glu222 are depicted in gray. In the B form, the Thr203 hydroxy group is hydrogen bonded to the anionic phenolate, thereby stabilizing its negative charge. In the A and the I form, the Thr203 side chain is turned away from the chromophore and Glu222 forms a hydrogen bond to Ser205.

spectra of the mutants that selectively populate the A, B, and I states of the protein, respectively. Concentration-dependent changes in the wt-GFP spectrum caused by dimerization have been reported and exploited for proximity imaging.[15] Our results allow the quantitative analysis of such concentration-dependent changes of the wt-GFP absorption in terms of ground-state population changes. We used such an analysis to determine the as yet unknown dimerization constant of wt-GFP. The observed spectral changes were attributed to shifts in the energies of the chromophore forms upon dimerization. From our data, a part of the potential energy surface of the protein involved in the interconversion of the A and the B forms could be constructed. Our approach also allows a qualitative description of the influence of point mutations on the energies of the different ground state forms. The implications of our findings for the biological function of GFP and for its use as a fluorescent probe are discussed below.

Results and Discussion

Concentration-dependent spectra of wt-GFP

It has been pointed out previously that the equilibrium between the A and the B forms of the chromophore in wt-GFP is affected by temperature, pH, ionic strength, cryoprotectors, and protein concentration.[10, 13, 14] The absorption spectrum of a wt-GFP solution changes considerably if the protein concentration is changed from 2.2 μM to 1.6 mM. The peak intensity of the A form at 397 nm increases at the expense of the long-wavelength B-form peak intensity (Figure 1 a). By performing absorption measurements at different concentrations of added Ni^{2+}, we verified that the spectral changes were not due to complex formation of the His_6-tag of the proteins with traces of Ni^{2+} (data not shown). The apparent isobestic point at 428 nm has been interpreted before as resulting from the dimerization of wt-GFP.[16] Close inspection of the spectra shows that three absorption bands are present. Normalization of the long wavelength part of the spectra to the respective maxima shows a change in the band shape caused by an absorption increase at λ_{abs} = 496 nm. This occurs simultaneously with the increase of the A form absorption (Figure 1 b). Differential absorptions ΔAbs were calculated by subtraction of the absorption spectra obtained for the limiting concentrations. The shape of a difference spectrum is not concentration dependent (data not shown). Low-temperature spectroscopy has previously been used to identify a fluorescence excitation band near 500 nm as a signature of the I form of the chromophore.[11, 12] We therefore attribute the differential absorption in the 500 nm region to the population of the I form of the chromophore. While low-temperature experiments are instructive for qualitative interpretation of the room-temperature data, no quantitative conclusions can be made because of the differences in spectral shapes and intensities. Quantitative analysis of the observed concentration-dependent spectral changes requires the knowledge of the room-temperature spectra of the contributing species. Ideally, such spectra are obtained from a GFP mutant with a chromophore stabilized in the I form, as well as from mutants that stabilize the A and B forms. However, the necessary mutations must not shift the respective electronic spectra.

Strategy for site-directed mutagenesis

The interconversion model put forward by Brejc and co-workers was used as a basis to plan specific mutations (Scheme 1).[7] From the model, two important factors determining the equilibrium between the three ground-state forms of the chromophore can be identified: The first of these factors is the acidity of the amino acid at position 222, which is of pivotal importance for the presence of the neutral A form of the chromophore in wt-GFP. The chromophore will only be neutral if in the electronic ground state the acidity at position 222 is higher than that of the phenol group of Tyr66. In the excited state of wt-GFP, the phenol group

 © 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

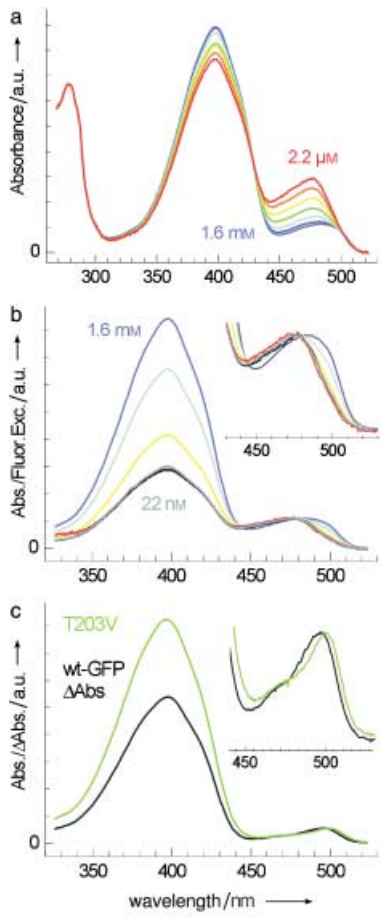


Figure 1. Dependence of spectral properties of wt-GFP on the protein concentration. a) Absorption spectra of serially diluted samples (1.6 mM – 2.2 μM, rainbow colored). The spectra were normalized to the absorption at 280 nm. b) Absorption spectra normalized to the long-wavelength maximum. At the lowest concentrations (1.8 μM (black) and 22 nM (gray)) fluorescence excitation data are shown. The absorption and the fluorescence excitation spectrum at 1.8 μM coincide. c) Differential absorption of two protein concentrations (1.6 mM and 2.2 μM) in comparison to the absorption spectrum of the Thr203Val variant. The long-wavelength region is magnified in the inset. The spectra were normalized to the 470 nm absorption.

is more acidic than Glu222, which in this case acts as a proton acceptor.[17] The presence of polar side chains in the vicinity of the chromophore inside the folded protein makes it difficult to directly compare pK_a values measured in solution. However, these values can be used as a good guideline. Here, the relevant values are p$K_a = 4.5$ for Glu, p$K_a = 10.0$ for phenol in its electronic ground state dissolved in H_2O, and p$K_a = 4.0$ for electronically excited phenol in H_2O.[18, 19] Electronic excitation of the A form consequently leads to the deprotonation of the phenol group and the population of the I form. Because of the rapidity of the excited-state proton transfer, which occurs on a picosecond time scale,[9, 10] the environment of the chromophore does not immediately adopt the optimum configuration for the accommodation of the anionic chromophore. The I form therefore possesses the same protein scaffold and hydrogen-bonding geometry as the A form.

The second important factor determining the protonation state of the chromophore comprises the orientation and the bonding properties of the amino acid at position 203. In the B form of wt-GFP, Thr203 forms a hydrogen bond to Tyr66 of the chromophore that is necessary to stabilize the anionic state. In the A and the I forms, the same hydroxy group of Thr203 is turned away from the phenolic hydroxy group of Tyr66. The interconversion between the different forms is also accompanied by a structural rearrangement of His148.

The key residues in this process, Thr203 and Glu222, can be independently targeted by mutation to specifically lower or raise the potential energy of either the A or the B form. With respect to a potential energy diagram, we call a mutation orthogonal if it only influences the potential energy of one of the three ground state forms of GFP. We used this effect as a basis for a rational approach to the stabilization of the I form in the ground state of a GFP mutant. Reduction of the acidity at position 222 by the Glu222Gln substitution maintains the hydrogen-bonding network in the same geometry as in the A form of wt-GFP.[8] An additional Thr203Val substitution, to give the double mutant Thr203Val/Glu222Gln, in turn prevents hydrogen bonding of the amino acid at position 203, which is needed for the stabilization of the B form in wt-GFP. Since they are isosteric, both mutations can be assumed to leave the protein scaffold unchanged. As will be demonstrated below, this double mutation indeed leads to the desired mutant with a stabilized I form in its ground state.

Spectroscopic characterization

Absorption and fluorescence spectroscopy were used to characterize the single point mutants Thr203Val and Glu222Gln, as well as the double mutant. Experiments were performed at cryogenic temperatures leading to sharper spectra that allowed us to unambiguously assign the different forms of the chromophore. The resulting room- and low-temperature spectra are shown in Figure 2, while the absolute band positions are summarized in Table 1.

The absorption spectrum of the Glu222Gln variant lacks any absorption at 400 nm but is remarkably similar to that of the long-wavelength part of the wt-GFP spectrum (Figure 2a). Close similarities are also found for the low-temperature excitation spectra and the emission spectra obtained with an excitation wavelength of $\lambda_{ex} \sim 470$ nm at low and at room temperature (Figure 2b and 2c). This means that the A form is efficiently suppressed in the ground state of this mutant. Further inspection of the low-temperature excitation spectra shows a small peak at $\lambda_{ex} \sim 500$ nm. Excitation at this spectral position yields an emission spectrum similar to that of the I* form in wt-GFP (Figure 2c). We also observed this behavior at room temperature (Figure 2b). In the red wing of the absorption band we detected a slight dependence of the emission spectra on the excitation wavelength (Figure 2b). This effect was previously reported for wt-GFP by Creemers and co-workers, who interpreted this spectral feature as being caused by the population of the I form.[12] These findings allude to the presence of a small population of the I form in Glu222Gln at cryogenic temperatures

 © 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

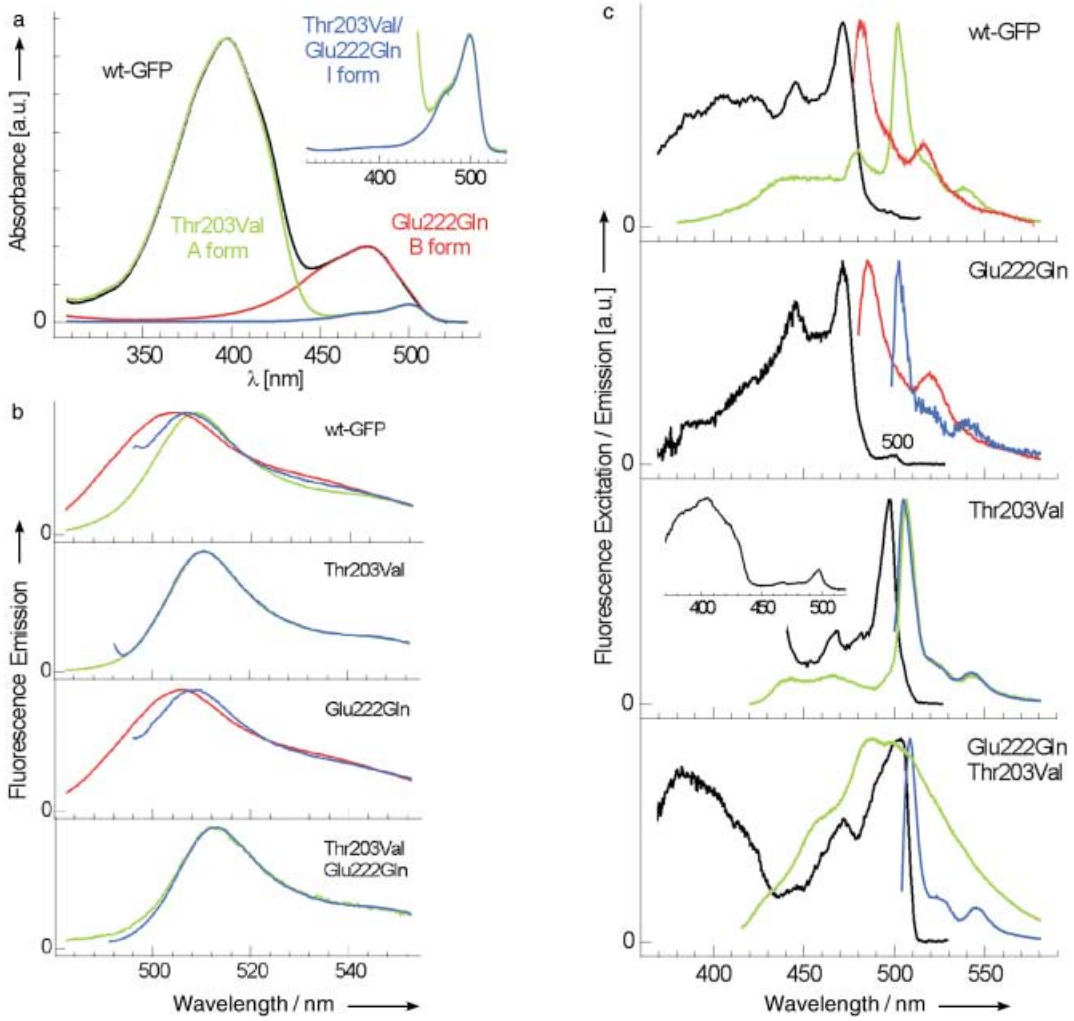


Figure 2. Spectroscopic properties of wild-type and mutant GFP. a) Absorption spectrum of wt-GFP in comparison to those of the mutated proteins representing the A, B, and I forms. The spectra were recorded at protein concentrations of approximately 15 μM. The long-wavelength absorption is enlarged in the inset. b) Fluorescence emission recorded at a protein concentration of around 0.1 μM. Fluorescence was excited at the absorption wavelengths of the A form (green, $\lambda_{ex} = 380$ nm), the B form (red, $\lambda_{ex} = 470$ nm), and the I form (blue, $\lambda_{ex} = 490 - 495$ nm). c) Fluorescence excitation (black) and emission spectra at 3 K in wt-GFP and its mutated variants (colored as in b). The complete excitation spectrum of the Thr203Val variant is depicted in the inset. Excitation wavelengths were 364 nm and 473 nm for wt-GFP, 473 nm and 496 nm for Glu222Gln, 407 nm and 496 nm for Thr203Val, and 407 nm and 502 nm for Glu222Gln/Thr203Val.

Table 1. *Summarized absorption, fluorescence excitation, and fluorescence emission maxima of the three different mutants at room temperature (RT) and at 3 K (LT).*[a]

Mutant	Absorption RT [nm]	Excitation RT [nm]	Emission RT [nm]	Excitation LT [nm]	Emission LT [nm]
wt-GFP	397(A)	397(A)	509(A)	400(A)	503(A)
	476(B)	476(B)	505(B)	472(B)	482(B)
Thr203Val	397(A)	397(A)	511(A)	400(A)	510(A)
	499(I)	499(I)	511(I)	497(I)	505(I)
Glu222Gln	477(B)	477(B)	506(B)	471(B)	484(B)
Thr203Val/Glu222Gln				~385(A)	~500(A)
	499(I)	503(I)	513(I)	503(I)	508(I)

[a] The identity of the respective chromophore state is given in parentheses. Fluorescence excitation and emission detection was carried out at slightly blue-shifted and red-shifted wavelengths with regard to the respective band maxima.

as well as at room temperature. All our findings demonstrate that the Glu222Gln mutation indeed destabilizes the A form of the chromophore without changing the electronic properties of the B form. It appears that the only effect of the isoelectronic substitution of the carboxy group by the amide group is the reduction of the acidity at position 222.

The room-temperature absorption spectrum of the Thr203Val variant is dominated by a broad band centered at 397 nm while

 © 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

only a minor absorption at 499 nm is detected. This confirms the expected prevalence of the A form (Figure 2a). The small peak at $\lambda_{abs} = 499$ nm is again attributed to a minor population of the I form.[20] Excitation of either the A form or the proposed I form at room and cryogenic temperatures leads to indistinguishable emission spectra, which further supports this assignment (Figure 2b and 2c). As expected from the interconversion model, the Thr203Val mutation suppresses the population of the B form.

If the mutations Thr203Val and Glu222Gln are orthogonal as defined above, their combination must lead to a stabilization of the I form. Indeed, none of the dominating absorption bands of wt-GFP are observed in the room-temperature spectrum of the Thr203Val/Glu222Gln variant. Instead, only one major band at $\lambda_{abs} = 499$ nm is detected (inset in Figure 2a). The Thr203Val/Glu222Gln mutant therefore represents a new GFP spectral phenotype. With the exception of a red shift of 2 nm, which lies within the resolution limits of the spectrometer, the room-temperature emission spectra of the double mutant are identical to those obtained for the Thr203Val variant (Figure 2b). In contrast to the absorption spectrum at room temperature, low-temperature excitation spectroscopy reveals a broad, unstructured band centered at 385 nm, which is similar to that of the A form in the Thr203Val mutant (Figure 2c). The I form excitation band of the double mutant at 503 nm is slightly red shifted at cryogenic temperatures. Apart from the red shift, the emission spectrum obtained upon excitation at 502 nm coincides with that of the Thr203Val mutant. The data obtained with optical spectroscopy impressively confirm the effect of the mutations predicted by the interconversion model.[7] The described double mutant is the prototype of the ground state I form of wt-GFP. This mutant allows access to the spectral properties of the I form and demonstrates that large red shifts can be obtained in the electronic spectra of GFP mutants without an aromatic amino acid substitution.

Aggregation behavior of wt-GFP

The absorption bands of the B and the I forms overlap spectrally. Knowledge of the absorption spectrum and the extinction coefficient (ε) of the I form is therefore a prerequisite for a quantitative analysis of the concentration-dependent changes in the wt-GFP spectra. The absorption maxima of both forms at 476 nm and 496 nm were selected as references.

If quantitative chromophore formation is assumed,[21] the extinction coefficients of the A form and the I form can be derived by determining the protein concentration based on the absorption at 280 nm.[22] For the mutants Thr203Val and Thr203Val/Glu222Gln we directly obtained values of $\varepsilon(A) = 25\,500\ mol^{-1}\,cm^{-1}$ at $\lambda_{abs} = 397$ nm and $\varepsilon(I) = 55\,400\ mol^{-1}\,cm^{-1}$ at $\lambda_{abs} = 496$ nm, respectively. Determination of these values is possible because both mutants only contain negligible populations of other chromophore forms (see Figure 2a). The extinction coefficient of the B form, $\varepsilon(B) = 36\,500\ mol^{-1}\,cm^{-1}$ at $\lambda_{abs} = 476$ nm, was derived from the long-wavelength part of the wt-GFP spectrum of the B form by subtracting the I form spectrum. As an independent control for this calculation of the extinction coefficients, the total protein concentration, [A] + [I] + [B], determined by the 280 nm absorption was compared to the sum $A_{397}/\varepsilon(A) + A_{476}(B)/\varepsilon(B) + A_{496}(I)/\varepsilon(I)$ for the concentration-dependent wt-GFP spectra (see Figure 1). Over the whole concentration range of the absorption spectra, the deviations between both measurements are smaller than 3.5%.

At 476 nm the absorption of the I form is 54% of its maximum value, while at 496 nm the B form has an absorption of 20% compared to its maximum value. From these data, the general equations (1) and (2) for the calculation of the respective absolute concentrations of the B and the I form, $A_{476}(B)$ and $A_{496}(I)$, can be derived from the measured total absorptions A_{476} and A_{496} at both absorption wavelengths [Eq. (1), Eq. (2)]:

$$A_{476}(B) = 1.12A_{476} - 0.61A_{496} \quad (1)$$

$$A_{496}(I) = 1.12A_{496} - 0.22A_{476} \quad (2)$$

The concentration-dependent relative populations can now be obtained by using Lambert-Beer's law. These popultions are shown for the A and the B form as well as for the ratio of A form to I form in wt-GFP in Figure 3. These ratios represent the equilibrium constants. With increasing protein concentrations, the populations of the A and I forms simultaneously increase more than sevenfold in comparison to that of the B form, while the ratio of the population of the A form to that of the I form remains constant. At the highest protein concentration, the population of the B form equals that of the I form. The

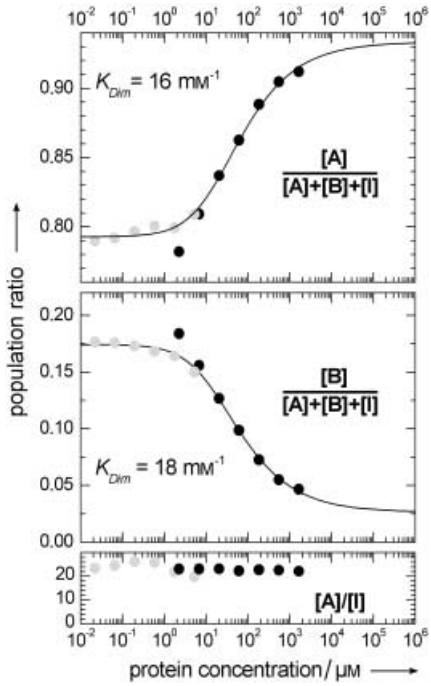


Figure 3. Relative populations of the A and the B forms, as well as the A/I form population ratio at different concentrations of wt-GFP. The dimerization constant K_{Dim} is obtained from a fit (black lines) to the absorption (black dots) and fluorescence excitation data (gray dots). The absorption value at the lowest concentration was not included in the fit.

 © 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

measurement range was extended to low protein concentrations by using fluorescence excitation spectroscopy.

These concentration-dependent changes between the populations of the different ground state forms of wt-GFP are explained by its oligomerization. Since it is known that the equilibrium constants of the A, B, and I forms of the chromophore are sensitive to the protein environment,[9, 11] it is reasonable to expect that these constants depend on the oligomerization state of the protein. Calculations comparing dimerization and trimerization show that the curves predicted for the formation of trimers are much steeper than those observed in Figure 3. Thus, our data do not support the formation of higher oligomers but suggest an equilibrium between the monomeric (M) and the dimeric (D) forms of the protein [Eq. (3)]. It is interesting to note that X-ray structures of wt-GFP show that the unit cells are built from dimers.[4] The ratio of the protein molecules in the dimeric state 2[D] to the total number of molecules [P] can be expressed as in Equation (4). One thus obtains relative populations of, for example, $([A]/([A]+[B]+[I]))_{mon}$ and $([B]/([A]+[B]+[I]))_{dim}$, which lead to Equation (5).

$$2M \rightleftarrows D;\ K_{dim} = \frac{[D]}{[M]^2};\ [P] = [M] + 2[D] \tag{3}$$

$$\frac{2[D]}{[P]} = 1 + \frac{1 - \sqrt{1 + 8K_{dim}[P]}}{4K_{dim}[P]} \tag{4}$$

$$\frac{[A]}{[A]+[B]+I} = \left(\frac{[A]}{[A]+[B]+[I]}\right)_{mon} + \frac{2[D]}{[P]}\left[\left(\frac{[A]}{[A]+[B]+[I]}\right)_{dim} - \left(\frac{[A]}{[A]+[B]+[I]}\right)_{mon}\right] \tag{5}$$

The values of $[A]/([A]+[B]+[I])$ and $[B]/([A]+[B]+[I])$ can be used independently to determine the dimerization constant. Both values agree excellently and we obtain a K_{Dim} value of 17 ± 1 mM^{-1} (Figure 3). Our experiments show that, at protein concentrations higher than 60 μM, more molecules of wt-GFP are in the dimeric form than in the monomeric form. The physiological concentration of wt-GFP in *Aequorea victoria* is in the range of 1 mM.[23] The results presented here agree well with analytical ultracentrifugation experiments in which only monomers and dimers were detected.[24] Note that the saturation concentration for the dimer is not experimentally accessible because at protein concentrations of around 50 mM there is already no water left in the sample. To calculate the water content of the sample, we assumed a specific density of 1.4 g cm^{-3} for the protein.

Potential energy surface of wt-GFP

From the equilibrium constants of the different protein forms, the ground-state energy differences ΔG between them can be calculated by using the relationship $\Delta G = -RT\ln K$. We obtain $\Delta G_{A,B} = -3.7$ kJ mol^{-1} and $\Delta G_{I,B} = 3.8$ kJ mol^{-1} in the monomeric protein and $\Delta G_{A,B} = -8.8$ kJ mol^{-1} and $\Delta G_{I,B} = -1.2$ kJ mol^{-1} in the dimer. The equilibrium between the A and the I form remains unchanged upon dimerization, with $\Delta G_{A,I} = -7.6$ kJ mol^{-1}. The thermodynamic results can be summarized in a potential diagram, as depicted in Figure 4a. The optical spectra demonstrate that the energy levels of the A and B forms can be shifted independently by the orthogonal point mutations Glu222Gln and Thr203Val, respectively (Figure 4b). It is apparent that the Thr203Val mutation mimics the influence of the dimerization on the stability of the B form. X-ray structural investigations show that Thr203 is located centrally on the dimer interface.[4] This spatial organization suggests that structural or electrostatic changes may mediate the observed interdependence of the dimerization and ground-state deprotonation equilibria.

Implications for the biological function of wt-GFP and its use as a label

It has been proposed that the natural function of GFP is the conversion of the blue chemiluminescence of aequorin, a protein produced by the jellyfish *Aequorea victoria*, into green light. The emission spectrum of free aequorin however has an emission maximum at around 470 nm.[6] If it is assumed that the emission spectra of free aequorin and of aequorin bound to GFP are the same, the wt-GFP would have to be in its B form for optimum energy transfer. This could be achieved by single point mutations like Glu222Gln or Ser65Thr.[1] In contrast, single point mutations like Thr203Val have been shown to stabilize the chromophore in its A form. While this demonstrates that either form of the chromophore could easily be stabilized, wt-GFP seems to be a compromise offering on the one hand efficient excited-state proton transfer after excitation at 400 nm, and on the other hand absorption at 470 nm for efficient energy transfer from aequorin emission. These findings support the assumption that wt-GFP serves the purpose of attracting zooplankton on which *Aequorea victoria* is known feed.[25] Hardly any knowledge is available on the wavelength dependence of the phototaxis of zooplankton. One study on the plankton *Daphnia magna*, however, shows a maximum positive phototaxis at 520 nm in good accordance with the fluorescence maximum of wt-GFP.[26] The broad absorption of wt-GFP in the spectral region between 400 and 500 nm, which is caused by the simultaneous presence of the A and the B form, offers the advantage of efficient collection of blue light from residual sunlight, which is then converted into dim green fluorescence. In contrast, the energy transfer from aequorin emission, which is rarely observed in nature,[25] mainly seems to be of importance as a defense mechanism.

The concentration-dependent spectral changes of wt-GFP can be employed for in vivo aggregation studies. Specifically, the quantitative analysis we have presented here could be employed to determine effective local GFP concentrations by microspectroscopy. A method for the detection of homodimerization of fusion proteins by calculating the ratio of the wt-GFP emission intensities observed at 397 nm and 476 nm has been described previously.[15] Here, we have shown that the dimerization of unfused wt-GFP leads to a sevenfold increase in the

 © 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

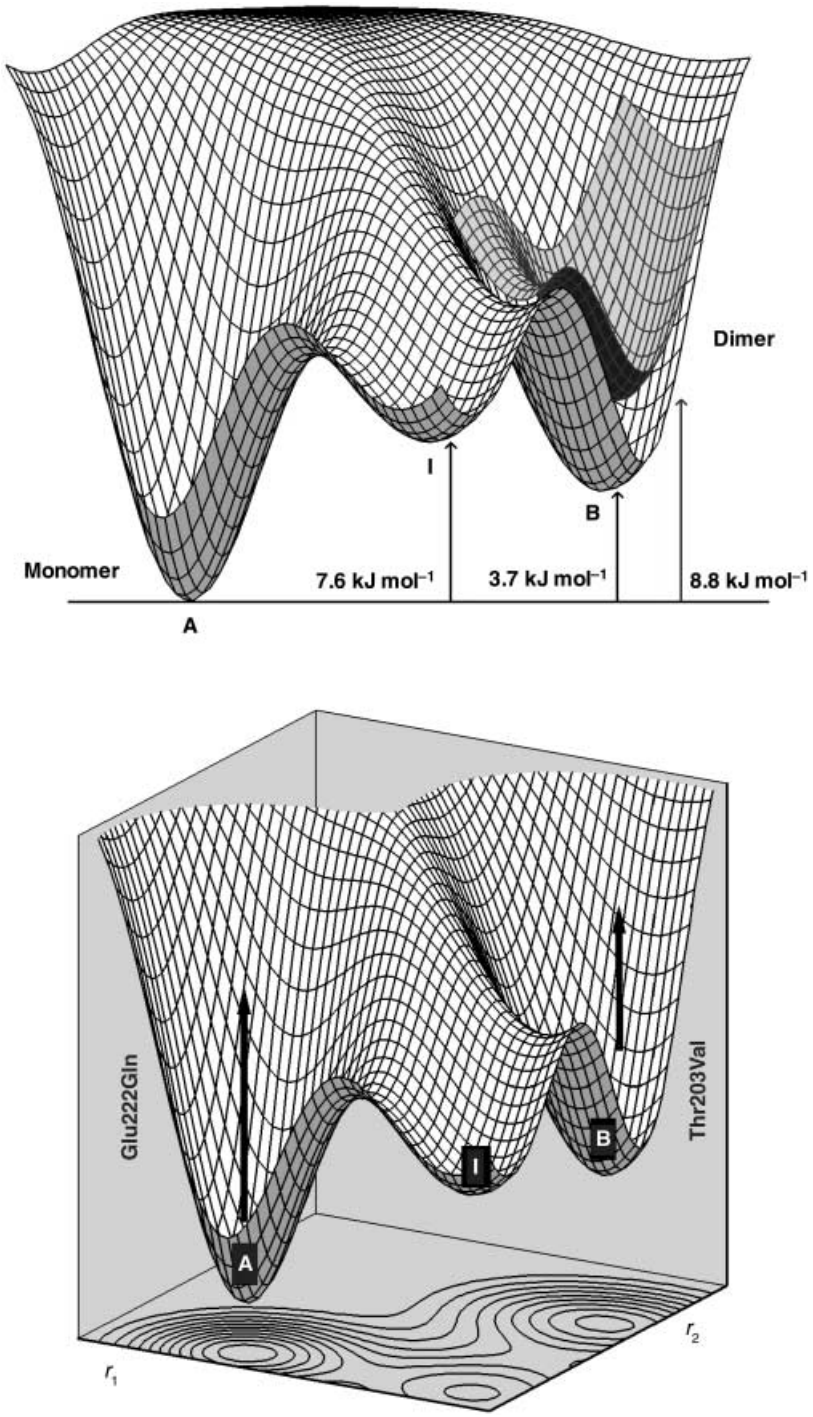


Figure 4. Potential energy surface diagram for the three ground state forms A, B, and I of wt-GFP. a) Change of the potential energy surface upon dimerization. The free energy differences derived from the dimerization constants are indicated. b) Principle of potential energy surface engineering. The orthogonal character of the mutation is indicated by the two coordinates r_1 *and* r_2. r_1 *describes the influence of hydrogen bonding on the potential energy of the A form of the chromophore, while* r_2 *describes the influence of the side chain acidity at position 222 on the hydrogen bonding network and on the potential energy of the B form of the chromophore.*

population ratio of the A form relative to the B form, as seen in Figure 3. Similar behavior is also seen for the dimerization of a fusion protein of wt-GFP with glutathione-S-transferase and for a crosslinked wt-GFP.[15] The opposite effect, an increase of the B form population at the expense of the A form population, was observed for a second fusion protein of wt-GFP with the FK506-binding protein (FKBP).[15] Based on our results and excluding fluorescence quenching by FKBP, we can conclude that in this case the protein structure of wt-GFP must be significantly influenced by the fusion, so that the potential energy of the ground-state forms of the GFP chromophore are shifted into relative positions opposite to those in free wt-GFP.

Conclusion

The results of the present study show that suitable mutations of GFP at positions 203 and 222 can be used to selectively influence the potential energy surface of its chromophore states. The combination of the orthogonal mutations Glu222Gln and Thr203Val selectively stabilizes the I form and thereby makes this transient intermediate of the wt-GFP photocycle available for research. Neither mutation influences the electronic spectra of the chromophore. Knowledge of the I form spectrum at room temperature can be used to quantitatively analyze the concentration-dependent spectroscopic changes of wt-GFP. We demonstrated that the B form is destabilized upon dimerization, which results in an increased spectral contribution by the A and I forms. Concentration-dependent spectroscopy is subsequently employed to determine the dimerization constant of wt-GFP. An empirical potential energy surface diagram representing the relative stability of the three forms was derived. A detailed understanding of the chromophore equilibria and the dimerization process in GFP as presented here sheds light on the biological function of wt-GFP and provides a spectroscopic basis for its application to proximity imaging.[15] Our work shows that potential energy surfaces can be manipulated in a controlled manner, even in the complex environment provided by the protein.

Methods

Samples: The *A. victoria* GFP gene[27] was used as described previously.[28] For cloning purposes, an *Nco*I restriction site was introduced at the 5′-end of the gene. The *Nco*I recognition sequence CCATGG includes the start codon and thus use of this sequence results in a change of the second amino acid from serine to glycine. At the C terminus of the protein, an additional glycine residue was introduced, followed by a His_6 tag. The mutants were constructed by site-directed mutagenesis [29] and expressed in *Escherichia coli* BL21 DE3. The soluble fraction of the bacterial lysate was purified on Ni-NTA

 © 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

agarose columns as published before.[28] The samples were solubilized in phosphate-buffered saline (pH 7.4, 4 mM KH_2PO_4, 16 mM Na_2HPO_4, 115 mM NaCl). The protein concentrations were calculated from the absorption at 280 nm.[29]

Spectroscopic measurements: Absorption spectra were determined on a Uvikon 943 spectrophotometer (Kontron Instruments) with a resolution of 2.5 nm. Excitation and emission spectra were measured on a Hitachi 4500 fluorescence spectrometer with 2.5 nm resolution. For concentration-dependent measurements, the solutions were concentrated and serially diluted threefold. Absorption spectra were measured with path lengths of 0.2, 2, and 10 mm, respectively.

Low-temperature experiments were done in a 90° geometry at 3 K as described previously.[12] Briefly, the proteins were buffered in a 70% glycerol/30% aqueous phosphate buffer mixture (pH 7.0). The sample was diluted in buffer solution to a concentration of approximately 10^{-6} mol L^{-1} and used to fill quartz sample cells. The sample was cooled in a flow cryosystem (Cryovac) to 3 K. The emission of a Xe arc lamp (Oriel 66002), dispersed by a double monochromator (Spex 1402) with a resolution of 0.03 nm, was used for the fluorescence excitation spectra. The fluorescence was passed through a suitable long-pass optical filter and detected with a GaAs photomultiplier (C30034, Hamamatsu) and a photon counting system (SR400/440, Stanford Research). To correct the spectra for the characteristics of the arc lamp, all spectra were divided by the lamp spectrum recorded with the same system. The fluorescence spectra were obtained by illumination with various selected lines from an Ar or Kr ion laser source (Coherent) and detected with the double monochromator and the photon counting system.

Acknowledgements

We thank Christoph Bräuchle for continuous support. Financial support was provided by the Deutsche Forschungsgemeinschaft, SFB 533.

Keywords: chromophores · green fluorescent protein · fluorescence spectroscopy · mutagenesis · photochemistry

[1] R. Y. Tsien, *Annu. Rev. Biochem.* **1998**, *67*, 509 – 544.
[2] A. B. Cubitt, R. Heim, S. R. Adams, A. E. Boyd, L. A. Gross, R. Y. Tsien, *TIBS* **1995**, *20*, 448 – 455.
[3] R. Heim, D. C. Prasher, R. Y. Tsien, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12 501 – 12 504.
[4] F. Yang, L. G. Moss, G. N. J. Phillips, *Nat. Biotechnol.* **1996**, *14*, 1246—1251.
[5] M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, S. J. Remington, *Science* **1996**, *273*, 1392 – 1395.
[6] H. Morise, O. Shimomura, F. H. Johnson, J. Winant, *Biochemistry* **1974**, *13*, 2656 – 2662.
[7] K. Brejc, T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö, S. J. Remington, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 2306 – 2311.
[8] G. J. Palm, A. Zdanov, G. A. Gaitanaris, R. Stauber, G. N. Pavlaksi, A. Wlodawer, *Nat. Struct. Biol.* **1997**, *4*, 361 – 365.
[9] M. Chattoraj, B. A. King, G. U. Bublitz, S. G. Boxer, *Proc. Natl. Sci. USA* **1996**, *93*, 8362 – 8367.
[10] H. Lossau, A. Kummer, R. Heinecke, F. Pöllinger-Dammer, C. Kompa, G. Bieser, T. Jonsson, C. M. Sylva, M. M. Yang, D. C. Youvan, M. E. Michel-Beyerle, *Chem. Phys.* **1996**, *213*, 1 – 16.
[11] C. Seebacher, F. W. Deeg, C. Bräuchle, J. Wiehler, B. Steipe, *J. Phys. Chem. B* **1999**, *103*, 7728 – 7732.
[12] T. M. H. Creemers, A. J. Lock, V. Subramaniam, T. M. Jovin, S. Völker, *Nat. Struct. Biol.* **1999**, *6*, 557 – 560.
[13] G. Striker, V. Subramaniam, C. A. M. Seidel, A. Volkmer, *J. Phys. Chem. B* **1999**, *103*, 8612 – 8617.
[14] M. Cotlet, J. Hofkens, M. Maus, T. Gensch, M. van der Auwerear, J. Michiels, G. Dirix, M. van Guyse, J. Vanderleyden, A. J. W. G. Visser, F. de Schryver, *J. Phys. Chem. B* **2001**, *105*, 4999 – 5006.
[15] D. A. De Angelis, G. Miesenböck, B. V. Zemelmann, J. E. Rothman, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12 312 – 12 316.
[16] W. W. Ward, H. Prentice, A. F. Roth, C. W. Cody, S. C. Reeves, *Photochem. Photobiol.* **1982**, *31*, 803 – 808.
[17] M. A. Lill, V. Helms, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2778 – 2781.
[18] C. R. Cantor, P. R. Schimmel, *Biophysical Chemistry*, W. H. Freeman and Co., San Francisco, **1980**, Chapter 2.
[19] E. L. Wehry, L. B. Rogers, *J. Am. Chem. Soc.* **1965**, *87*, 4234 – 4238.
[20] A. Kummer, J. Wiehler, H. Rehaber, C. Kompa, B. Steipe, M. Michel-Beyerle, *J. Phys. Chem. B* **2000**,*104*, 4791 – 4798.
[21] G. H. Patterson, S. M. Knobel, W. D. Sharif, S. R. Kain, D. W. Piston, *Biophys. J.* **1997**, *273*, 2782 – 2790.
[22] S. C. Gill, P. H. von Hippel *Anal. Biochem.* **1989**, *182*, 319 – 326.
[23] M. W. Cutler, W. W. Ward, *Photochem. Photobiol.* **1993**, *57*, 63S (Abstract).
[24] G. N. Phillips, Jr., Cu*rr. Opin. Struct. Biol.* **1997**, *7*, 821 – 827.
[25] C. E. Mills, **1999—2002**, http://faculty.washington.edu/cemills/Aequorea.html
[26] U. C. Storz, R. J. Paul, *J. Comp. Physiol. A* **1998**, *183*,709 – 717.
[27] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, D. C. Prasher, *Science* **1994**, *263*, 802 – 804.
[28] G. Jung, J. Wiehler, W. Göhde, J. Tittel, T. Basché, B. Steipe, C. Bräuchle *Bioimaging* **1998**, *6*, 54 – 61.
[29] T. A. Kunkel, Proc. *Natl. Acad. Sci. USA* **1985**, *82*, 488 – 492.

Received: March 10, 2003
Revised: July 16, 2003 [F595]

 © 2003 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim