J|A|C|S
ARTICLES
Published on Web 06/13/2002

Slow Exchange in the Chromophore of a Green Fluorescent Protein Variant

Markus H. J. Seifert, Dorota Ksiazek, M. Kamran Azim, Pawel Smialowski, Nediljko Budisa, and Tad A. Holak*

Contribution from the Max-Planck-Institute for Biochemistry, 82152 Martinsried, Munich, Germany

Received January 31, 2002

Abstract: Green fluorescent protein and its mutants have become valuable tools in molecular biology. They also provide systems rich in photophysical and photochemical phenomena of which an understanding is important for the development of new and optimized variants of GFP. Surprisingly, not a single NMR study has been reported on GFPs until now, possibly because of their high tendency to aggregate. Here, we report the ^{19}F nuclear magnetic resonance (NMR) studies on mutants of the green fluorescent protein (GFP) and cyan fluorescent protein (CFP) labeled with fluorinated tryptophans that enabled the detection of slow molecular motions in these proteins. The concerted use of dynamic NMR and ^{19}F relaxation measurements, supported by temperature, concentration- and folding-dependent experiments provides direct evidence for the existence of a slow exchange process between two different conformational states of CFP. ^{19}F NMR relaxation and line shape analysis indicate that the time scale of exchange between these states is in the range of 1.2−1.4 ms. Thermodynamic analysis revealed a difference in enthalpy $\Delta H_0 = (18.2 \pm 3.8)$ kJ/mol and entropy $T\Delta S_0 = (19.6 \pm 1.2)$ kJ/mol at $T = 303$ K for the two states involved in the exchange process, indicating an entropy−enthalpy compensation. The free energy of activation was estimated to be approximately 60 kJ/mol. Exchange between two conformations, either of the chromophore itself or more likely of the closely related histidine 148, is suggested to be the structural process underlying the conformational mobility of GFPs. The possibility to generate a series of single-atom exchanges ("atomic mutations") like H → F in this study offers a useful approach for characterizing and quantifying dynamic processes in proteins by NMR.

Introduction

Pioneered by Chalfie et al.[1] green fluorescent proteins (GFPs) are now ubiquitously used in molecular biology for monitoring gene expression, protein movement, and protein interaction.[2] The optical absorption, fluorescence, and ultrafast dynamic fluorescence properties of GFP are commonly explained by a three-state model[3−6] that assumes that the chromophore exists either in a neutral (395 nm) or an anionic (475 nm) state; the latter state exists in a thermodynamically unstable intermediate form (493 nm) and a stable low-energy form.[7] The model of an intramolecular Förster-cycle[5,8] rationalizes the fluorescence dynamics. GFPs exhibit a variety of photochemical processes including photoconversion[3,9,10] and complex on−off blinking and switching behavior in single-molecule measurements,[11−13] which cannot be easily explained on the basis of the canonical three-state model.[14]

Until now, 18 high-resolution crystal structures of GFPs have been available (http://www.rcbs.org). The overall structure of GFP consists of an 11-stranded β-barrel with a central helix that carries the chromophore.[15,16] The X-ray diffraction studies and a variety of physicochemical methods highlight an apparent exceptional stability of the GFP fold in which the chromophore lies rigidly inside the conformationally inflexible GFP molecule.[17] The stability of GFP against denaturants and proteases is very high,[2] the melting temperatures for *Aequorea*, *Renilla*,

* Corresponding author. E-mail: holak@biochem.mpg.de.

(1) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science* **1994**, *263*, 802−805.
(2) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509−544.
(3) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362−8367.
(4) Lossau, H.; Kummer, A.; Heinnecke, R.; Pöllinger-Dammer, F.; Kompa, C.; Bieser, G.; Jonsson, T.; Silvia, C.; Yang, M.; Youvan, D.; Michel-Beyerle, M. E. *Chem. Phys.* **1996**, *213*, 1−16.
(5) Prendergast, F. *Methods Cell Biol.* **1999**, *58*, 1−18.
(6) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormo, M.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2306−2311.
(7) Creemers, T. M. H.; Lock, A. J.; Subramaniam, V.; Jovin, T. M.; Völker, S. *Nat. Struct. Biol.* **1999**, *6*, 557−560.
(8) Förster, T. *Z. Elektrochemie* **1950**, *54*, 42−46.
(9) van Thor, J. J.; Pierik, A. J.; Nugteren-Roodzant, I.; Xie, A.; Hellingwerf, K. J. *Biochemistry* **1998**, *37*, 16915−16921.
(10) van Thor, J. J.; Gensch, T.; Hellingwerf, K. J.; Johnson, L. N. *Nat. Struct. Biol.* **2002**, *9*, 37−41.
(11) Dickson, R. M.; Cubitt, A. B.; Tsien, R. Y.; Moerner, W. E. *Nature* **1997**, *388*, 355−358.
(12) Creemers, T.; Lock, A.; Subramaniam, V.; Jovin, T.; Völker S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2974−2978.
(13) Garcia-Parajo, M.; Segers-Nolten, G.; Veerman, J.; Greve, J.; von Hulst, N. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7237−7242.
(14) Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6177−6182.
(15) Ormoe, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392−1395.
(16) Yang, F.; Moss, L. G.; Phillips Jr.; G. N. *Nat. Biotechnol.* **1996**, *14*, 1246−1251.
(17) Striker, G.; Subramaniam, V.; Seidel, C. A. M.; Volkmer, A. *J. Phys. Chem. B* **1999**, *103*, 8612−8617.

10.1021/ja0257725 CCC: $22.00

and *Phialidium* GFP are 78, 70, and 69 °C, respectively.[18] However, even the earliest absorption spectroscopy studies by Ward et al. indicated some conformational flexibility in the chromophore binding site.[19] More recently, several fluorescence, X-ray crystallographic, and computational studies tried to address the contributions of various fast dynamic processes to the fluorescence spectra of GFPs. At room temperature, denatured GFP or model chromophores in solution show no significant fluorescence. However, when frozen as ethanol glass at 77 K, such compounds become highly fluorescent.[20,21] Inside the protein environment the chromophore is very sensitive to mutations of surrounding residues.[22,23] Competition of radiationless photoisomerization and fluorescence emission has been suggested to provide an explanation for the differences in quantum yield of fluorescence in various GFP mutants.[22,24] It was proposed that any increase in the degrees of motional freedom has to be avoided when designing new GFP variants.[24] For photoisomerization pathways the rotations around either of the two bridging bonds between the two rings of the chromophore and a so-called hula-twist have been suggested, but it was not possible to determine the time scales of conformational changes quantitatively by molecular dynamics simulations.[14,25]

In contrast to many spectroscopic techniques, nuclear magnetic resonance (NMR) spectroscopy provides a large frequency range for studying dynamical processes from picosecond-to-second time scales and even longer at atomic resolutions. For example, motional and thermodynamical information for backbone amides in proteins can be obtained from measurements of ^{15}N relaxation rates in ^{15}N-labeled proteins.[26,27] Dynamic NMR spectroscopy[28] is able to reveal populations and rate constants of exchange processes in proteins. Further information, for example, on the energy landscape of proteins, may be revealed by ^{17}O and ^{2}H NMR relaxation measurements.[29] Side-chain dynamics can be investigated by measurement of, for example, ^{13}C relaxation of methyl groups[26,30,31] or ^{19}F relaxation[32−34] in fluorine-labeled aromatic rings. ^{19}F NMR spectroscopy offers the advantage over ^{1}H, ^{15}N, and ^{13}C NMR that only specifically labeled sites give a rise to an NMR signal which circumvents in many cases the assignment of highly crowded ^{1}H, ^{15}N, and ^{13}C NMR spectra. In addition many proteins of known structure, which are too large (M > 30 kDa) for conventional multidimensional NMR studies, fall within the range of molecular weights accessible to ^{19}F NMR (M < 100 kDa).[35] In addition it is usually assumed that fluorine substitution has negligible steric effects on the protein structure,[35] although indications exist that it may influence catalytic behavior[36] and cause subtle differences in side-chain and backbone conformation.[37] Fluorine NMR has been used for ligand binding[38] and folding studies[39,40] and for probing conformational[35,41,42] and light-induced structural changes.[43]

We present here the first quantitative ^{19}F NMR characterization of the conformational dynamics in the cyan fluorescent protein (CFP), a GFP mutant in which tyrosine 66 is replaced with a tryptophan. The incorporation of different fluorotryptophans (i.e., 4-, 5-, and 6-fluoro-Trp) results in direct post-translational integration of the indole moiety and concomitantly ^{19}F nucleus into the CFP chromophore providing an almost ideal reporter of its dynamics. Using this approach we were able to demonstrate that either CFP chromophore itself or residues in its vicinity interconvert between multiple conformations on the time scale of milliseconds. These slow motions should therefore be considered in the interpretation of structural and spectroscopic properties of the ground states of GFPs.

Methods and Experimental Section

Sample Preparation. The incorporation of fluorinated analogues 4-fluoro-Trp (4FW), 5-fluoro-Trp (5FW), and 6-fluoro-Trp (6FW) into Clontech GFPuv (F99S, M153T, V163A), Clontech EGFP (F64L, S65T), and Clontech ECFP (F64L, S65T, Y66W, N146I, M153T, V163A) was achieved in *Escherichia coli* Trp-auxotrophic strain ATCC49980[44] using the selective pressure incorporation (SPI) method developed in our lab.[45] The protein expression host *E. coli* ATCC49980 was routinely co-transformed with two plasmids: ampicillin-resistant pQE−30-PP4 harboring the GFP and CFP gene sequence under the control of the T5 promoter, and kanamycin-resistant pREP4 containing a repressor gene *lacI*^{q}. The native and fluorinated proteins included a polyhistidine tag and were purified using Ni-chelate columns. The purity of the recombinant proteins was checked by SDS-PAGE (Commassie staining). The quantitative replacement of the native Trp-residues by its noncanonical analogues was routinely confirmed by electrospray mass spectrometric analyses (ESI-MS, API III, Sciex Perkin-Elmer) as described earlier.[44] In addition UV-absorption spectra were routinely recorded in PBS buffer solution, pH 7.3, as described later. Before recording NMR spectra the proteins were dialyzed against bi-distilled water and the samples were concentrated to 5−10 mg/mL. D_2O (10%)

(18) Ward, W. W. In *Bioluminescence and Chemiluminescence: Basic Chemistry and Analytical Applications*; DeLuca, M. A., McElroy, W. D., Eds.; Academic Press: New York, 1981; pp 235−242.
(19) Ward, W. W.; Prentice, H. J.; Roth, A. F.; Cody, C. W.; Reeves, S. C. *Photochem. Photobiol.* **1982**, *35*, 803−808.
(20) Niwa, H.; Inouye, S.; Hirano, T.; Matsuno, T.; Kojima, S.; Kubota, M.; Ohashi, M.; Tsuji, F. I. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13617−13622.
(21) Bell, A. F.; He, X.; Wachter, R. M.; Tonge, P. J. *Biochemistry* **2000**, *39*, 4423−4431.
(22) Kummer, A. D.; Kompa, C.; Lossau, H.; Pöllinger-Dammer, F.; Michel-Beyerle, M. E.; Silva, C.; Bylina, E.; Coleman, W.; Yang, M.; Youvan D. *Chem. Phys.* **1998**, *237*, 183−193.
(23) Kummer, A. D.; Wiehler, J.; Rehaber, H.; Kompa, C.; Steipe, B.; Michel-Beyerle M. E. *J. Phys. Chem. B.* **2000**, *104*, 4791−4798.
(24) Voityuk, A. A.; Michel-Bayerle, M. E.; Rösch, N. *Chem. Phys. Lett.* **1998**, *296*, 269−276.
(25) Chen, M. C.; Lambert, C. R.; Urgitis, J. D.; Zimmer, M. *Chem. Phys.* **2001**, *270*, 157−164.
(26) Spyracopoulos, L.; Sykes, B. D. *Curr. Opin. Struct. Biol.* **2001**, *11*, 555−559.
(27) Kay, L. E. *Nat. Struct. Biol.* **1998**, NMR Supplement, 513−517.
(28) Sandström, J. *Dynamic NMR Spectroscopy;* Academic Press: London, 1982.
(29) Denisov, V. P.; Peters, J.; Hörlein, H. D.; Halle, B. *Nat. Struct. Biol.* **1996**, *3*, 505−509.
(30) Palmer, A. G., III; Wright, P. E.; Rance, M. *Chem. Phys. Lett.* **1991**, *185*, 41−46.
(31) Mulder, F. A. A.; Mittermaier, A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *Nat. Struct. Biol.* **2001**, *8*, 932−935.
(32) Hull, W. E.; Sykes, B. D. *Biochemistry* **1974**, *13*, 3431−3437.
(33) Hull, W. E.; Sykes, B. D. *J. Chem. Phys.* **1975**, *63*, 867−880.
(34) Hull, W. E.; Sykes, B. D. *J. Mol. Biol.* **1975**, *98*, 121−153.
(35) Danielson, M. A.; Falke, J. J. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 163−195.
(36) Dominguez, M. A.; Thornton, K. C.; Melendez, M. G.; Dupureur, C. M. *Proteins:. Struct., Funct., Genet.* **2001**, *45*, 55−61.
(37) Xiao, G.; Parsons, J. F.; Tesh, K.; Armstrong, R. N.; Gilliland, G. L. *J. Mol. Biol.* **1998**, *281*, 323−339.
(38) Rong, D.; Lin, C.-L. S.; d'Avignon D. A.; Lovey, A. J.; Rosenberger, M.; Li, E. *FEBS Lett.* **1997**, *402*, 116−120.
(39) Bann, J. G.; Pinkner, J.; Hultgren, S. J.; Frieden, C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 709−714.
(40) Ropson, I. J.; Frieden, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7222−7226.
(41) Cho, B. P., Zhou, L. *Biochemistry* **1999**, *38*, 7572−7583.
(42) Kwon, K.; Jiang, Y. L.; Song, F.; Stivers, J. T. *J. Biol. Chem.* **2002**, *277*, 353−358.
(43) Klein-Seetharaman, J.; Getmanova, E. V.; Loewen, M. C.; Reeves, P. J.; Khorana, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13744−13749.
(44) Minks, C.; Huber, R.; Moroder, L.; Budisa, N. *Biochemistry* **1999**, *38*, 10649−10659.
(45) Minks, C.; Alefelder, S.; Moroder, L.; Huber, R.; Budisa, N. *Tetrahedron* **2000**, *56*, 9431−9442.

was added to allow the field locking of the NMR spectrometer. All chemicals were purchased from Sigma or Aldrich unless stated otherwise.

Optical Spectroscopy. UV-absorption spectra of the proteins in the phosphate buffered saline (PBS: 115 mM NaCl, 8 mM KH_2PO_4, 16 mM Na_2HPO_4, pH 7.3) were recorded on a Perkin-Elmer Lambda 17 UV/vis spectrophotometer. Molar extinction coefficients (ϵ_M) for native and fluorinated proteins were determined at room temperature (20 °C) according to the procedure described by Mach et al.[46] Native and 6FW-EGFP have essentially the same absorbance ($\lambda_{max} = 277$ nm: $\epsilon_M = 21000\ M^{-1}\ cm^{-1}$; $\lambda_{max} = 488$ nm: $\epsilon_M = 38000\ M^{-1}\ cm^{-1}$). Molar extinction coefficients (ϵ_M) for native and substituted ECFP variants were different: native ECFP: $\lambda_{max} = 278$ nm: $\epsilon_M = 25000\ M^{-1}\ cm^{-1}$; $\lambda_{max} = 434$ nm: $\epsilon_M = 25500\ M^{-1}\ cm^{-1}$; 6FW-ECFP: $\lambda_{max} = 280$ nm: $\epsilon_M = 23300\ M^{-1}\ cm^{-1}$; $\lambda_{max} = 430$ nm: $\epsilon_M = 22900\ M^{-1}\ cm^{-1}$. Detailed analysis of these and other fluorinated variants of ECFP and EGFP will be reported in a related paper. Fluorescence spectra were recorded on a Perkin-Elmer spectrometer (LS50B) equipped with digital software. Protein samples (0.25 μM; slit 2.5 nm) were excited at 488 nm (EGFP) or at 450 nm (ECFP), and the emission spectra were recorded in the 490−540 nm (EGFP) or 460−540 nm (ECFP) range. Emission and excitation spectra of 6FW ECFP were measured at temperatures of 288, 293, 298, 303, 308, 313, 318, and 323 K.

NMR Spectroscopy. ^{19}F NMR measurements were carried out on a Bruker DRX500 spectrometer equipped with a dual $^1H-^{19}F$ probehead. The fluorine Larmor frequency was 470 MHz. All measurements were performed at a temperature of 303 K with exception of six spectra taken at temperatures between 293 and 318 K. Due to the negative NOE enhancement factor of fluorine nucleii in proteins, no proton decoupling was used. All ^{19}F NMR spectra were referenced to an external reference consisting of 10 mM TFA in 90% H_2O, 10% D_2O at a temperature of 303 K. For a typical one-dimensional ^{19}F NMR spectrum of 14 kHz sweep-width up to 16k scans with an interscan delay of 1 s were recorded. Baseline distortions caused by a relatively long prescan delay were corrected by manual baseline fitting.

The $^1H-^{19}F$ *J*-coupling constants in 4-, 5-, and 6-fluorotryptophan were measured in ^{19}F NMR spectra of pure tryptophan samples. To evaluate the effect of *J*-coupling on the results of line shape fitting experimentally, spectra of 6FW CFP with and without proton decoupling were recorded. Despite of a strong reduction of signal intensity upon decoupling no difference in line shape of the W66 resonances was visible (data not shown). Therefore, it was concluded that the calculated values for the rate constants are not affected by *J*-coupling.

pH titrations of the 4FW- and 6FW-labeled CFP sample were carried out by dissolving the protein in 50 mM borate buffer at pH 9.1 with 100 mM NaCl and then lowering the pH by adding the appropriate amount of 1% H_3PO_4 for each step. ^{19}F spectra at pH 9.1, 8.1 and 7.1 were recorded. To record a ^{19}F NMR spectrum of unfolded 6FW CFP, urea was added, and heating to 95 °C for 10 min was applied. To analyze the influence of protein concentration on ^{19}F NMR spectra a sample of 6FW ECFP was diluted in four steps to 1/16th of the original concentration. The ^{19}F NMR spectra of the reference and of the diluted samples were recorded with 8k scans (reference, 8 mg/mL, 0.3 mM monomeric), 16k scans (4 mg/mL, 0.16 mM), 32k scans (2 mg/mL, 0.08 mM), 64k scans (1 mg/mL, 0.04 mM), and 128k scans (0.5 mg/mL, 0.02 mM). The effect of UV irradiation on 6FW ECFP was checked by recording ^{19}F NMR spectra after irradiating the NMR sample with a 450 W Xe-lamp (~5 mW light power at 400 nm) for 4 h.

NMR Relaxation Measurements. The longitudinal relaxation time T_1 of ^{19}F nuclei was measured by standard inversion recovery methods with inversion times of 0.2 10, and 100 ms and 0.5, 1, and 2 s. The interscan delay was 3 s. The T_1 data was fitted by a single-exponential function.[33] The heteronuclear $\{^1H\}-^{19}F$ Overhauser effect was measured by comparing one-dimensional ^{19}F spectra with and without presaturation for 1 s on water (NOE = I_{sat}/I_{ref}).[33] Presaturation was used for saturating protons due to a reduced heating effect during the application of the 1 s pulse. Usually broadband decoupling is required for measuring the heteronuclear Overhauser effect due to the chemical shift dispersion of protons. Broadband decoupling was evaluated and showed no significant differences in the resulting heteronuclear Overhauser effect values. The peak integrals A and the line widths $\Delta\nu_{exp}$ of the peaks in all spectra were extracted by fitting the spectra with the appropriate number of Lorentzians.

To determine the overall tumbling time of GFP, T_1, T_2 and $\{^1H\}-^{15}N$ heteronuclear Overhauser effect measurements were performed on a ^{15}N-labeled sample of GFPuv.[47] The overall tumbling time τ_C was calculated using $\tau_C = (2\omega_N)^{-1}\sqrt{(6T_1/T_2-7)}$.[48] Only amide groups in structured parts of the protein were taken into account for this calculation.

Simulation of ^{19}F Relaxation. For the interpretation of ^{19}F relaxation data no standard method of analysis, such as that of reduced spectral density mapping for ^{15}N relaxation, is available. Therefore, a simulation of the ^{19}F relaxation in 5FW due to dipolar interactions with protons and fluorine chemical shift anisotropy was used to enable the data interpretation in terms of motional frequencies. The simulation was based on the following equations (see ref 49):

$$\frac{1}{T_1} = R_1 = \frac{1}{20}K^2[J(\omega_H - \omega_F) + 3J(\omega_F) + 6J(\omega_H + \omega_F)] + \frac{1}{15}\omega_F^2\Delta\sigma^2\left(1+\frac{\eta^2}{3}\right)J(\omega_F)$$

$$\frac{1}{T_2} = R_2 = \frac{1}{40}K^2[4J(0) + J(\omega_H - \omega_F) + 3J(\omega_F) + 6J(\omega_H) + 6J(\omega_H + \omega_F)] + \frac{1}{15}\omega_F^2\Delta\sigma^2\left(1+\frac{\eta^2}{3}\right)[3J(\omega_F) + 4J(0)]$$

$$f = \frac{I_{sat}}{I_{ref}} - 1 = \frac{1}{20}\frac{\gamma_H}{\gamma_F}K^2[6J(\omega_H + \omega_F) - J(\omega_H - \omega_F)]/R_1$$

with $K = (\mu_0/4\pi)\hbar\gamma_H\gamma_F r^{-3}$. $r = 1/N\sum_i r_i$ is defined as an averaged fluorine−proton distance, with *N* indicating the number of contributing protons. This distance *r* is set to 2 Å in our calculations. This is valid since the dominating contribution to the observed relaxation arises from interactions with protons on other residues[50] The spectral density is defined by $J(\omega) = (2S^2\tau_C)/(1 + \omega^2\tau_C^2) + [2(1 - S^2)\tau_e]/[1 + \omega^{2r}\tau_e^2]$ with the effective correlation time $\tau_e^{-1} = \tau_C^{-1} + \tau_i^{-1}$, where τ_c is the overall tumbling time correlation time and τ_i the correlation time of internal motion. Using this model-free approach the spectral density is characterized by two correlation times τ_i and τ_c and the order parameter S^2.[51] An overall tumbling time of $\tau_c = 22$ ns was assumed. The chemical shift anisotropy was set to $\Delta\sigma\sqrt{1+\eta^2/3} = -93.5$ ppm.[50] The simulation was done with Mathematica (Wolfram Research).

Analysis of Thermodynamic Parameters. To evaluate the populations and rate constants, the resonances of W66 in the 6FW ECFP fluorine spectra were fitted with Lorentzians functions, as well as with the McConnell equations[52] according to the complete band shape (CBS) method.[28] This approach is also called dynamic NMR.

(46) Mach, H.; Middaugh, C. R.; Lewis, R. V. *Anal. Biochem.* **1992**, *200*, 74−80.
(47) Skelton, N. J.; Palmer, A. G., III; Akke, M.; Kördel, J.; Rance, M.; Chazin, W. J. *J. Magn. Reson. B* **1993**, *102*, 253−264.
(48) Fushman, D.; Weisemann, R.; Thüring, H.; Rüterjans, H. *J. Biomol. NMR* **1994**, *4*, 61−78.
(49) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, 2nd ed.; Wiley−VCH: New York, 2000.
(50) Luck, L. A.; Vance, J. E.; O'Connell, T. M.; London, R. E. *J. Biol. NMR* **1996**, *7*, 261−272.
(51) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4570.
(52) McConnell, H. M. *J. Chem. Phys.* **1958**, *28*, 430.

In the slow exchange limit the CBS method allows the determination of the populations and rate constants as well as the T_2 without exchange broadening. Therefore this method is not affected by different magnetic field inhomogeneties in different samples. The CBS method comprises a fit of the experimental data to the McConnell equations for two-site exchange spectra with neglect of *J*-coupling.[52] This approach is valid in this case since the ^{1}H−^{19}F *J*-coupling is much smaller than the observed line widths. The fitting was done using the Levenberg−Marquardt algorithm. It was assumed that T_2 is the same for both states. It was found that different settings of T_2 in the McConnell equations do not affect the fitting result in our case unless T_2 gets unrealistically short. Comparison of fits where T_2 was neglected to the fits with variable T_2 showed no difference.

To further confirm the results, the data from the Lorentzian fit was used to calculate the equilibrium constant $\bar{K} = k_B/k_A = p_A/p_B$ of this exchange process in an independent way. The free energy ΔG_0, enthalpy ΔH_0 and entropy ΔS_0 were obtained from nonlinear fit of $\bar{K}$ according to the Arrhenius-type equation $\bar{K} = \exp[-(\Delta H_0 - T\Delta S_0)/RT]$ and $\Delta G_0 = \Delta H_0 - T\Delta S_0$.[28] The activation energy of the exchange process was estimated according to the formula $\Delta G^{act}_{A,B} = aT\,[10.319 + \log(T/k_{A,B})]$ with the constant $a = 19.14$ J K^{-1} mol^{-1}, which is derived from the well-known Eyring equation.

Modeling of the CFP Structure. To evaluate the possible structural origins of exchange the structure of CFP was modeled based on the crystal structure coordinates of 4-amino-tryptophan CFP (Bae, personal communication, see also ref 53). In this CFP variant the two tryptophans (W57 and W66) are replaced by the noncanonical amino acid 4-amino-tryptophan. To construct the CFP model the 4-amino group of the tryptophan (W66) involved in chromophore formation was removed and a short run of energy minimization was performed using the modeling software package SYBYL (Tripos, Inc., St. Louis, MO). No fluorine atom was modeled in the chromophore since the replacement of a hydrogen atom with a fluorine atom is supposed to have negligible steric effect.[35] To reveal gap regions around the chromophore the program SURFNET[54] was used. SURFNET generates gaps by fitting spheres between all pairs of atoms and computing three-dimensional density maps.

Results

UV and Fluorescence Spectroscopy. As expected, the UV absorption profile as well as fluorescence spectra of EGFP (F64L, S65T) were not affected by fluorination of W57 which is placed about 15 Å apart from the chromophore. In fact, single absorption peak at 488 nm of both EGFP variants as well as emission maximum at 510 nm clearly indicate anionic state of the chromophore (Figure 1A). Conversely, native ECFP (F64L, S65T, Y66W, N146I, M153T, V163A) exhibits two chromophore absorption bands (434 and 452 nm) and two distinct fluorescence emission maxima (476 and 500 nm). Fluorination of W57 and chromophore indole induced 4 nm blue shift in the absorption maxima of 6FW-ECFP (430, 448 nm) as well as decrease in the absorbance intensity by about 10 %. These effects are even more pronounced in fluorescence properties of 6FW-ECFP: while emission maxima are negligibly blue shifted by about 2 nm, the fluorescence intensity is lowered by about 50% compared to the native ECFP (Figure 1B). It should be however noted that basic features of spectral behavior of ECFP (i.e. the presence of two spectral bands in UV and fluorescence) are not changed upon fluorination. A detailed study on the effects of fluorination on the optical properties of GFP variants will be published in a related paper.[53]

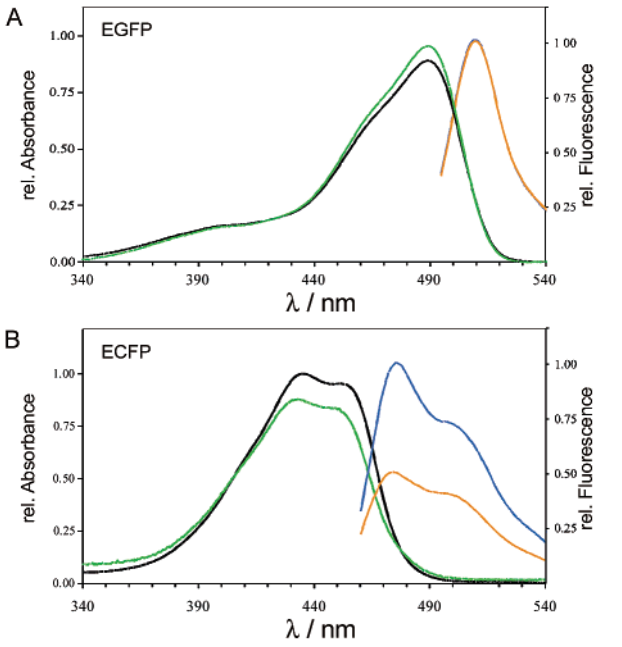

Figure 1. UV absorption and fluorescence emission spectra of EGFP (A, black and blue), 6FW EGFP (A, green and orange), ECFP (B, black and blue) and 6FW ECFP (B, green and orange) were recorded at neutral pH. Note that absorption band around 395 nm characteristic for wild-type GFP is suppressed in EGFP and its fluorinated counterpart. Only one absorption band at 488 nm and emission maxima at 510 nm are characteristic for the anionic state of chromophore in EGFP. On the other hand the UV-absorbance and fluorescence spectra for both native and fluorinated ECFP are characterized with two spectroscopically distinguishable states (for more details see Results section).

NMR Assignment. The ^{19}F NMR resonance of free 4-fluorotryptophan is split into a quartet ($^3J_{FH} = 12$ Hz, $^4J_{FH} = 4.6$ Hz), whereas 5- and 6-fluorotryptophan show a sextet structure ($^3J_{FH} = 11$ Hz, $^4J_{FH} = 5.5$ Hz) (data not shown). These splittings are obscured in the broad protein signals when these tryptophans are incorporated into EGFP and ECFP proteins.

The ^{19}F NMR spectra of ECFP and EGFP labeled with 4-, 5-, or 6-fluorotryptophan exhibited four and two resonances, respectively (Figure 2). CFP contains only two tryptophan residues, W57 and W66, whereas GFP has only one tryptophan, W57. This clearly indicates that each of the two tryptophans in CFP exists in two states. The more and less populated states of W57 and W66 are marked with subscript A and B, respectively (Figure 2). Since one of the CFP tryptophans, W66, is incorporated into the CFP chromophore during protein folding by cyclization of the backbone, two states exist within the chromophore. The resonances of 4FW and 6FW ECFP can be unambiguously assigned by comparison with the EGFP spectra. The assignment of the resonances of 5FW ECFP is complicated by the fact that a second peak is not visible in the 5FW EGFP spectrum. To assign the 5FW resonances, the peak integral values given in Table 1 were used. Since W66 and W57 are incorporated in the protein in a 1:1 ratio, the sum of integrals for both states of each residue must give the same value. Peaks assigned with this method are marked with *f* in Table 1. The ratio of the integrals of states A and B for W57 and W66 are consistent within the 4FW and 5FW CFP samples ($W66_A/W66_B = 0.3$ and $W57_A/W57_B = 0.2$). 6FW CFP shows integral ratios of $W66_A/W66_B = 0.6$ and $W57_A/W57_B = 0.6$. The difference in chemical shift for the two states of W66 is 2.6 ppm in 4FW and 6FW ECFP. The ^{19}F NMR spectra of 4FW

(53) Budisa, N.; Bae, J.; Rubini, M.; Weyher, E.; Wenger, W.; Azim, M. K.; Moroder, L.; Huber, R. Manuscript in preparation, 2002.
(54) Laskowski, R. A. *J. Mol. Graphics* **1995**, *13*, 323−330.

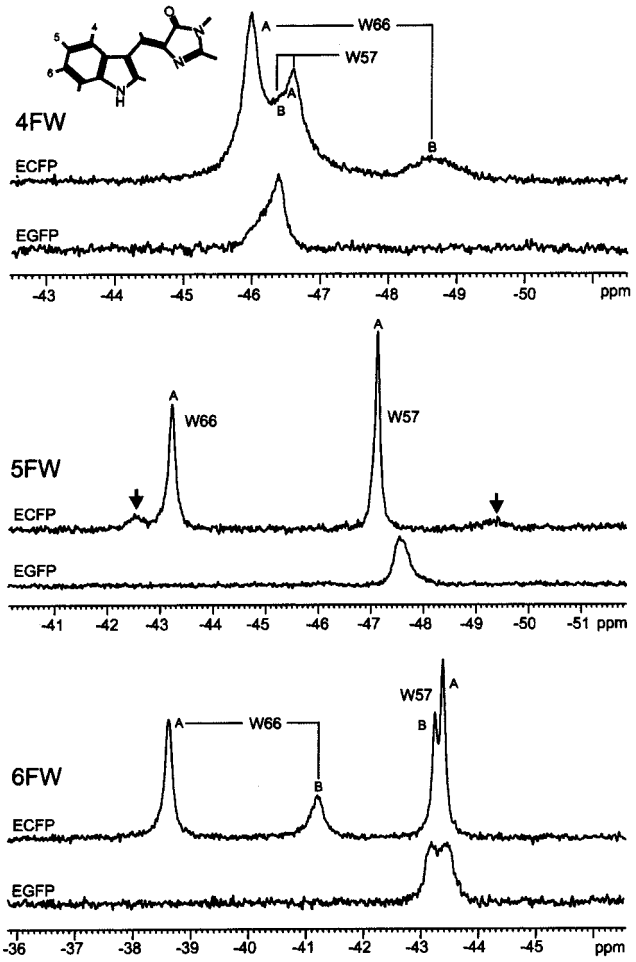

Figure 2. ^{19}F NMR spectra of ECFP and EGFP labeled with 4FW, 5FW, and 6FW: The chromophore of CFP and the fluorinated sites are shown in the upper left corner. The spectra were recorded at a temperature of 303 K without proton decoupling and referenced relative to TFA. An exponential line-broadening of 3 Hz was used to increase signal-to-noise ratio. The arrows in the spectrum of 5FW ECFP indicate two resonances which were assigned by comparison of peak integrals (see text).

Table 1. ^{19}F NMR Relaxation of 4-, 5-, and 6-FW CFP

peak	δ/ppm [a]	A_{rel} [b]	T_1/s [c]	$\Delta\nu_{exp}$/Hz [d]	het NOE [e]
4FW $W66_A$	−46.0	0.87 ± 0.03	0.85 ± 0.02	123 ± 3	0.19
4FW $W57_A$	−46.4	1.00 ± 0.06	−	320 ± 18	−
4FW $W57_B$	−46.8	0.25 ± 0.07	0.87 ± 0.04	104 ± 15	0.19
4FW $W66_B$	−48.6	0.30 ± 0.01	1.00 ± 0.28	290 ± 19	0.30
5FW $W57_B$ [f]	−42.5	0.18 ± 0.03	−	168 ± 33	<0.1
5FW $W66_A$	−43.2	0.84 ± 0.02	0.64 ± 0.07	72 ± 2	0.13
5FW $W57_A$	−47.1	1.00 ± 0.01	0.55 ± 0.04	55 ± 1	0.21
5FW $W66_B$ [f]	−49.3	0.26 ± 0.03	−	361 ± 67	−
6FW $W66_A$	−38.6	1.00 ± 0.02	0.57 ± 0.01	72 ± 1	0.14
6FW $W66_B$	−41.2	0.56 ± 0.02	0.42 ± 0.03	118 ± 5	0.13
6FW $W57_B$	−43.2	0.56 ± 0.02	0.58 ± 0.01	47 ± 2	0.16
6FW $W57_A$	−43.4	0.96 ± 0.02	0.61 ± 0.01	51 ± 1	0.20

[a] ^{19}F chemical shift. [b] Peak integrals relative to the integral of the most intense peak of the corresponding sample. [c] Spin−lattice relaxation time. [d] Measured line widths including an artificial line broadening of 3 Hz. [e] Heteronuclear {^{1}H}−^{19}F Overhauser effect (I_{sat}/I_{ref}). [f] Resonances assigned by comparison of peak integrals.

and 6FW GFPuv did not show double peaks (data not shown) indicating a different time scale of exchange.

The appearance of two peaks for each tryptophan in the spectra of ECFP and EGFP and the different broadening of those peaks provided a first indication for the existence of slow exchange processes between different states of the tryptophan side chains. This is especially important since it gives a strong hint that a slow exchange process exists within the chromophore of CFP.

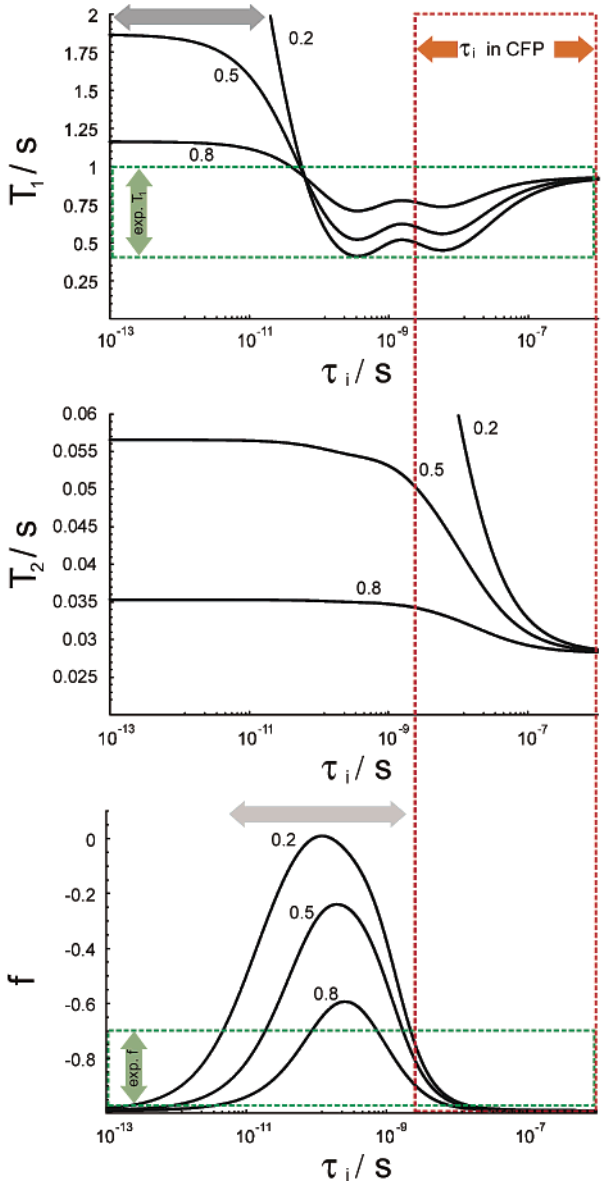

Figure 3. ^{19}F relaxation simulation: the relaxation times T_1 and T_2 as well as the heteronuclear {^{1}H}−^{19}F NOE factor f were calculated for a protein labeled with 5FW. An overall correlation time of τ_c = 22 ns was used as determined by ^{15}N relaxation measurements. The results are plotted for different values of the order parameter S^2. A smaller order parameter indicates larger amplitudes of motion. Green vertical arrows indicate the range of the experimental values. Therefore significant motions on the 10^{-12} to 10^{-9} s time scale are excluded by the experimental values of f, as indicated by a light gray bar. Strong motions on time scales shorter than 10^{-12} s are not present due to low experimental T_1 values, as shown by a dark gray bar. A red, horizontal arrow indicates the range of time scales for indole flexibility that is in agreement with the experimental data. Experimental T_2 values are shorter than calculated due to exchange broadening.

^{19}F NMR Relaxation. The results of the ^{19}F NMR relaxation measurements are summarized in Table 1. The spin−lattice relaxation time T_1 is in the range of 0.42 s to 1.0 s. The line widths $\Delta\nu_{exp}$ are in the range of 47−360 Hz. Therefore, the values of the effective spin−spin relaxation time $T_2^* = (\pi\Delta\nu_{exp})^{-1}$ of 0.9−7 ms are very short compared to the values expected from theoretical calculations even when the artificial line broadening of 3 Hz and the obscured J-splitting is taken

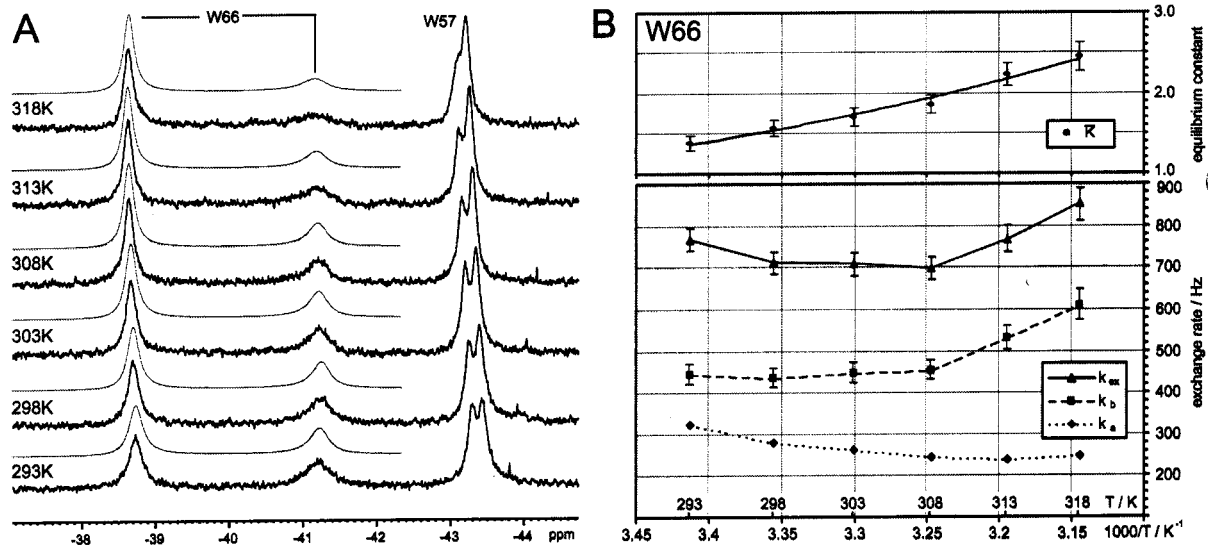

Figure 4. Temperature dependence of the ^{19}F NMR spectra of 6FW ECFP: (A) Spectra were recorded at six different temperatures from 293 K to 318 K. State A of W66 gets more populated, whereas resonance of state B is reduced and broadened. Peak shapes obtained by fitting the McConnell equations to the experimental spectra are shown as thin lines on top of the spectra. The numerical fitting results for W66 are listed in Table 2. (B) The exchanges rates k_A, k_B and $k_{ex} = k_A + k_B$ and the equilibrium constant $\bar{K} = k_B/k_A$ are plotted against temperature. The total exchange rate k_{ex} shows a minimum around 303 K and increases then with temperature. The equilibrium constant $\bar{K}$ increases with temperature.

into account (Figure 3). The hetero-NOE factors are mostly in the range of f = NOE − 1 = −0.8 to −0.9, which is in-line with nearly complete cancellation of the signal when proton decoupling is used.

For defining the valid time scales of motions in CFP the experimental data are compared with the theoretical values calculated by ^{19}F relaxation simulation. Several time scales can be excluded because the theoretical values and experimental data do not coincide, including 10^{-12} -10^{-9} s (hetNOE data) and 10^{-12} s and shorter (T_1 data) time scales (Figure 3). Therefore, significant motions are only present on time scales longer than nanoseconds.

Thermodynamic Analysis. The temperature dependence of ^{19}F NMR spectra of 6FW ECFP provides additional evidence for the existence of slow exchange processes (Figure 4A). These dynamic NMR measurements show that the exchange rate and state populations change with temperature. The results for the rate constants and populations are summarized in Table 2 and Figure 4B. The mean lifetime of the W66 states was found to be $\tau = 1/k_{ex} = 1.2$ ms to 1.4 ms for temperatures in the range of 293−318 K. Remarkably, the rate of interconversion from $W66_A$ to $W66_B$ (k_A) decreases with temperature, whereas the rate of interconversion from $W66_B$ to $W66_A$ (k_B) increases exponentially with temperature. This leads to a minimum of the total exchange rate $k_{ex} = k_A + k_B$ at approximately 303 K. The resonances of W57 were not fitted because the condition of well-separated peaks was not fulfilled, which is important for a reliable McConnell fitting. Two methods were used to determine the equilibrium constant $\bar{K}$ (see Figure 5 and Table 2): Lorentz and McConnell fitting. Despite a small offset in the equilibrium constant resulting from both methods, the calculated enthalpy and entropy values are in agreement within the experimental error. The data point for 318 K of the Lorentzian fit data has been neglected since its large deviation from the expected behavior can be explained by the proximity to the coalescence point. Therefore, the McConnell equations provide the more appropriate model at higher temperatures. The

Table 2. Temperature Dependence of W66 State Populations and Exchange Rates in 6FW CFP

Lorentzian fit

T/K[a]	$\Delta\nu_A$/Hz[b]	p_A[c]	$\Delta\nu_B$/Hz[d]	p_B[e]	$\bar{K}_{Lorentzian} = p_A/p_B$[f]
293	95 ± 2	0.55 ± 0.01	168 ± 5	0.45 ± 0.01	1.21 ± 0.02
298	82 ± 1	0.58 ± 0.01	154 ± 4	0.42 ± 0.01	1.39 ± 0.03
303	79 ± 1	0.62 ± 0.01	145 ± 4	0.38 ± 0.01	1.65 ± 0.03
308	73 ± 1	0.64 ± 0.01	167 ± 4	0.36 ± 0.01	1.77 ± 0.04
313	71 ± 1	0.66 ± 0.01	201 ± 7	0.34 ± 0.01	1.96 ± 0.04
318	71 ± 1	0.64 ± 0.01	330 ± 17	0.36 ± 0.01	1.75 ± 0.04

McConnell fit

T/K[a]	k_A/Hz[g]	p_A[h]	k_B/Hz[i]	p_B[j]	$\bar{K}_{McConnell} = k_B/k_A$[k]
293	322 ± 4	0.58 ± 0.01	445 ± 24	0.42 ± 0.01	1.38 ± 0.09
298	278 ± 3	0.61 ± 0.01	435 ± 23	0.39 ± 0.01	1.56 ± 0.10
303	263 ± 3	0.63 ± 0.01	448 ± 24	0.37 ± 0.01	1.70 ± 0.11
308	244 ± 2	0.65 ± 0.01	453 ± 24	0.35 ± 0.01	1.86 ± 0.11
313	238 ± 2	0.69 ± 0.01	530 ± 29	0.31 ± 0.01	2.23 ± 0.14
318	248 ± 3	0.71 ± 0.01	607 ± 37	0.29 ± 0.01	2.45 ± 0.18

[a] Temperature. [b] Line width of state A (corrected for the artificial line broadening of 3 Hz). [c] Population of state A. [d] Line width of state B (corrected for the artificial line broadening of 3 Hz). [e] Population of state B. [f] Equilibrium constant calculated from state populations. [g] Rate of conversion from A to B. [h] Population of state A. [i] Rate of conversion from B to A calculated from $k_B = k_A\, p_A/p_B$. [j] Population of state B calculated from $p_B = 1 - p_A$. [k] Equilibrium constant calculated from exchange rates.

McConnell fit is based on the assumption of T_2 being the same for both states. This is valid since the rigid and planar chromophore is located in the core of the protein[5] and therefore is likely to show the same overall rotational diffusion as the protein itself.[17] Despite of the deviation of the individual reaction rates from simple Arrhenius behavior, the relationship of $\ln(\bar{K})$ to $1/T$ is linear, according to the regression coefficient R (Figure 5B). In most cases it is assumed that deviations from Arrhenius behavior do not occur for the range of temperatures used in this study.[55] In our case this applies at least for the equilibrium constant. Therefore it is valid to fit the equilibrium constant with an exponential function.

(55) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598−1603.

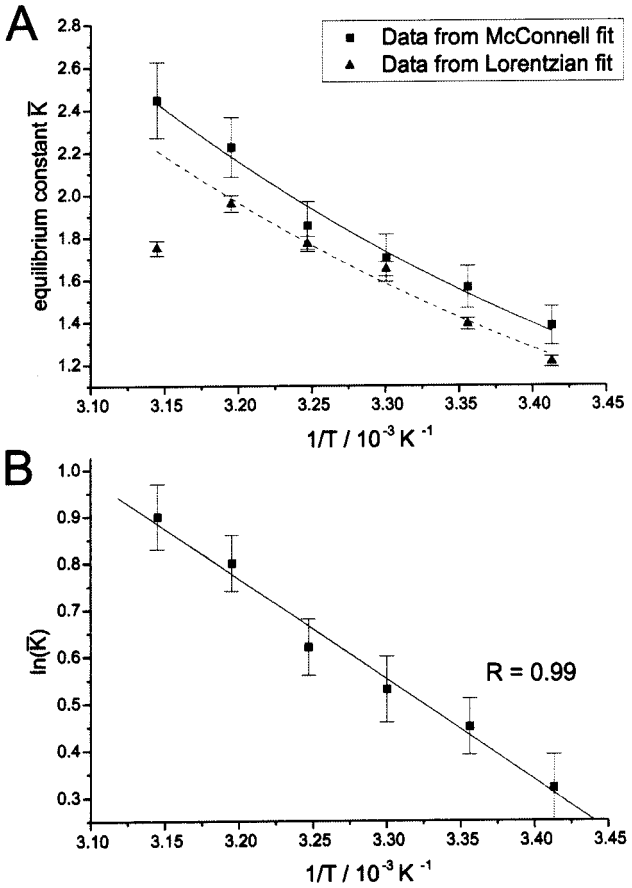

Figure 5. (A) Comparison of data obtained by fitting the McConnell equations and Lorentzian peaks to the experimental data. Despite a small offset in the curves of the equilibrium constant, the resulting enthalpy and entropy values are identical within the experimental error. (B) Graph of the natural logarithm of the equilibrium constant $\bar{K}$ against the inverse temperature. The relationship of $\ln(\bar{K})$ to $1/T$ is linear with a very high probability (regression coefficient $R = 0.99$). Therefore, the equilibrium constant can be fitted with an exponential function.

This results in a molar enthalpy of $\Delta H_0 = (18.2 \pm 3.8)$ kJ/mol and an entropy of $\Delta S_0 = (64.6 \pm 3.8)$ J K^{-1} mol^{-1}. At temperatures around $T = 303$ K the difference in entropy $T\Delta S_0 = (19.6 \pm 1.2)$ kJ/mol compensates the difference in enthalpy which is in-line with the small value of $\Delta G_0 = -RT \ln(\bar{K}) = -(1.3 \pm 0.1)$ kJ/mol. The negative sign of ΔG_0 is a consequence of the definition of the equilibrium constant being $\bar{K} = p_A/p_B$; therefore, it does not indicate an exothermic reaction. It is known that for an endothermic reaction higher temperatures favor the products. Therefore, state $W66_A$ may be attributed to a state of the chromophore higher in the energy than state $W66_B$. The difference in free energy of both states is of the same magnitude as the thermal energy $RT_{303K} = 2.5$ kJ/mol. The free energy of activation for the exchange process was estimated to be in the range of 57−63 kJ/mol.

Influence of Denaturation, pH, Protein Concentration, and Irradiation with UV Light. To clarify the nature of the two states of ECFP found by NMR, the effect of denaturation was examined (Figure 6). The fluorine resonances experience a pronounced upfield shift of 3−5 ppm upon protein unfolding. This can be attributed to the change of environment from the hydrophobic core of GFP to the hydrophilic solution. The assignment of the resonances of the unfolded protein was based on the chemical difference of W57 and W66. One resonance

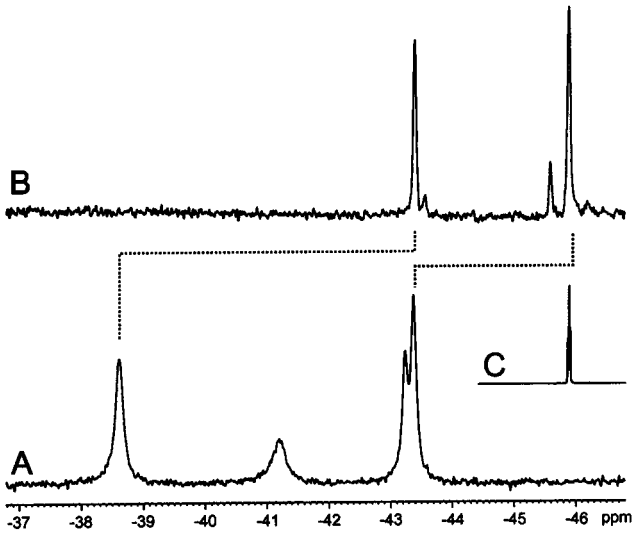

Figure 6. ^{19}F NMR spectra of folded 6FW ECFP (A), unfolded 6FW ECFP (B) and 50 mM 6FW (C): Even unfolded CFP clearly shows double peaks for the W57 resonance. For the W66 only a very small second peak is observed.

has exactly the chemical shift of pure 6-fluorotryptophan in solution. This resonance is attributed to W57. The indole of W66 is bound to an imidazolinone ring and should therefore exhibit a different chemical shift for the fluorine atom at the 6 position. Therefore, the resonance at −43.4 ppm is assigned to W66. Despite unfolding the W57 shows two distinct resonances. This is attributed to proline *cis* and *trans* conformations[56] in the P56−W57−P58 sequence. In the vicinity of the W66 resonance a small peak is detectable whose origin, however, could not be established without being too speculative. The resonances of unfolded CFP are significantly narrower compared to those of folded CFP, indicating higher flexibility and the absence of exchange broadening. The "ground state" $W66_B$ in folded CFP is always shifted upfield compared to the "excited state" $W66_A$. But due to several significant contributions to ^{19}F chemical shift[57] this observation cannot be attributed to a specific origin at the moment.

pH titration with 6FW (see Supporting Information Figure S1) and 4FW ECFP (data not shown) showed no significant changes for W66 resonances, but a line narrowing can be observed for the W57 resonances. Protein concentration has no effect on the two states observed by ^{19}F NMR in the 6FW ECFP sample (see Figure 7). The dilution spectra show that the two states of W66 and W57 cannot be attributed to monomer−dimer states of ECFP. No significant changes in ^{19}F NMR spectra of 6FW ECFP after irradiation with UV light for 4 h (see Supporting Information Figure S2) could be detected.

Discussion

Fluorination Influence on Absorbance and Fluorescence. In the last three decades ^{19}F NMR studies have taken advantage of high sensitivity of ^{19}F chemical shifts to the changes in local environments including van der Waals packing interactions and local electrostatic fields with no background signals from solvent. In most cases, crystallographic studies confirmed that structures of fluorinated proteins are indistinguishable from those

(56) Wüthrich, K. *NMR of Proteins and Nucleic Acids;* John Wiley & Sons: New York, 1994.
(57) Lau, E. Y.; Gerig, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 4408−4417.

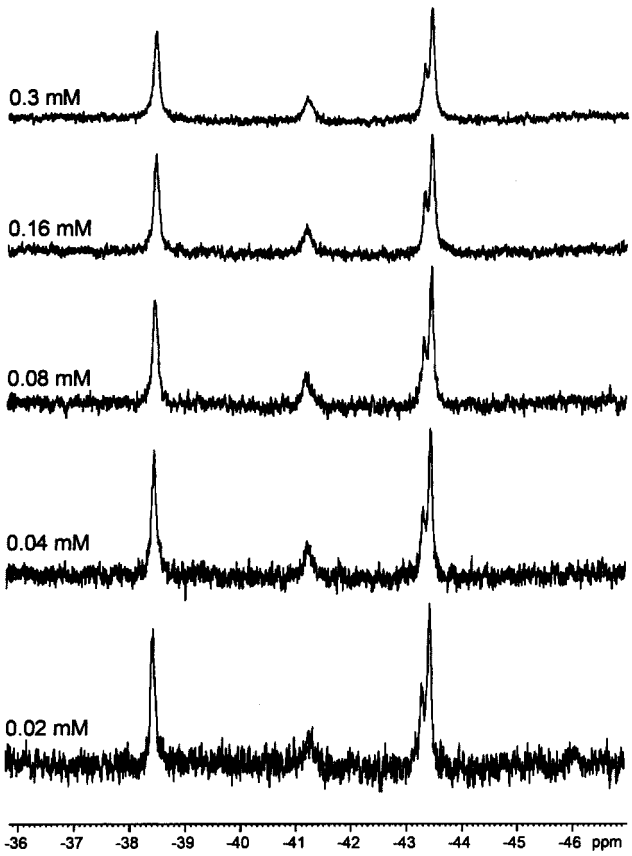

Figure 7. Concentration dependence in ^{19}F NMR spectra of 6FW ECFP: protein concentration does not affect the exchanging states in ECFP significantly. This excludes dimerization as source of the observed states. The concentration is given as the total concentration of monomeric protein (27 kDa). The decrease in signal-to-noise ratio results from the linear increase of signal intensity with the number of scans per spectrum but a quadratic increase of noise level.

of wild-type protein forms.[44] This makes ^{19}F NMR a powerful technique as it allows to probe subtle conformational changes and to reveal intrinsic dynamic features of the proteins. Aromatic residues are especially suitable targets for fluorination because they are normally less likely to be involved in internal motions than smaller amino acids, thus facilitating chemical shift calculations. In this context, the residue specific labeling of EGFP and ECFP with fluorine containing tryptophan analogues offers a good model for NMR studies as well as the opportunity to study possible effects of fluorination on the dynamic properties of protein chromophore.

Fluorination and its effects on the indole moiety of the tryptophan-residues have been well studied in the literature. For example, the UV transition and fluorescence emission of indole is composed of the two overlapping transitions ^1L_a and ^1L_b that are nearly orthogonal in polarization. The relative energies and oscillator strengths of these two electronic transitions depend on the solvent as well as on nature and position of the substituent in the indole ring. Thus the almost isosteric H $\rightarrow$ F (e.g. C−F bond is 0.3 Å longer than the aromatic C−H bond) single atom exchange ("atomic mutations") in indole moiety will result in inverted polarities and, thus, an increased charge separation in the indole ring. Since fluorine is the most electronegative element its electron-withdrawing properties should affect dipole moments (e.g. 6-F-indole has a dipole moment of $\mu = 3.05$ D while indole only $\mu = 2.10$ D[58]) and consequently spectroscopic properties of the substituted tryptophans.

Our spectroscopic investigations of substituted proteins at neutral pH clearly indicate that incorporation of electron-withdrawing fluorine in the environment (i.e. W57) of EGFP chromophore does not affect absorbance and steady-state fluorescence properties of this molecule (Figure 1). On the other hand, additional fluorination of W66 of ECFP provides direct integration of the fluorine nucleus into the chromophore. Therefore, the limited blue shift in both fluorescence and absorbance as well as the dramatic decrease in fluorescence intensity upon fluorination can be attributed to this H $\rightarrow$ F substitution. Detailed spectroscopic and structural characterization of these and other fluorinated variants will be reported in a related paper. However, it should be kept in mind that the general behavior of ECFP is not substantially changed by fluorination since two states of chromophore seen as absorption bands at around 430 and 450 nm as well as in fluorescence are preserved (Figure 1). Thus these spectral properties represent relatively static properties of the GFP chromophore whose dynamic behavior can certainly be better dissected using ^{19}F NMR.

Presence of Slow Exchange. Since light absorption takes place on a femtosecond time scale − much faster than any structural process in proteins − the information on structural dynamics of proteins obtained by UV absorption spectroscopy is limited to an instantaneous average of the immediate environment of the absorbing moiety on this time scale.[59] Compared to the UV spectroscopy NMR is able to characterize motions on a much broader range of time scales. Each of the ECFP residues W57 and W66 exhibits two ^{19}F NMR resonances, which is a clear sign for slow exchange processes.

For both tryptophans, W66 and W57, the very low $f =$ NOE − 1 values exclude the existence of pronounced motions on the picosecond to nanosecond time scale. Excessive motions on time scales shorter than picoseconds are also excluded by the small values of T_1. ^{19}F line widths indicate the existence of motions on a time scale of nanoseconds and longer. The temperature dependence and denaturation behavior observed in the ^{19}F NMR spectra of 6FW ECFP provides direct evidence for the existence of slow exchange processes. The mean lifetime of the W66 states was found to be 1.2−1.4 ms for temperatures in the range of 293−318 K. Temperature-dependent measurements revealed that at temperatures around 303 K the difference in entropy $T\Delta S_0 = (19.6 \pm 1.2)$ kJ/mol compensates the difference in enthalpy $\Delta H_0 = (18.2 \pm 3.8)$ kJ/mol. This is in-line with a small value of ΔG_0. This entropy−enthalpy compensation is very prevalent in biological systems[60] and is also characteristic for an order−disorder transition.[29,31] But an order−disorder transition can be excluded since the NOE values are similar and the T_1 values of the lower-energy state $W66_B$ are shorter than those of $W66_A$ indicating even a reduced flexibility of the higher-energy state $W66_A$ on the nanosecond-to-microsecond time scale. The enthalpy and entropy values are comparable to those found by Mulder et al. for the conformational exchange process in a cavity mutant of T4 lysozyme.[31]

(58) Cotton, M.; Tian, C.; Busath, D. D.; Shirts, R. B.; Cross, T. A. *Biochemistry* **1999**, *38*, 9185−9197.
(59) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 2nd ed.; Kluwer Academic: New York, 1999.
(60) Dunitz, J. D. *Chem. Biol.* **1995**, *2*, 709−712.

Slow conformational exchange between two ground states of the chromophore is also in-line with the existence of two UV absorption and fluorescence peaks for CFP and their temperature dependence (see Supporting Information Figure S3). The pH titration suggests that the two exchanging states of the chromophore in ECFP cannot be attributed to different protonation states of the indole nitrogen. This leads to the conclusion that the two states of the CFP chromophore are formed in a different way compared to GFP where the protonation state of the phenolic oxygen determines the state of the chromophore.

Effect of Fluorine Location on Exchange. As stated before, in 6FW CFP the energetically lower state ($W66_B$) is more populated and therefore more stabilized compared to that in 4FW CFP and 5FW CFP. This difference may be attributed to a specific interaction of 6-fluorotryptophan with surrounding residues. The difference of $\Delta G_0 = -RT \ln(p_B/p_A)$ between 6FW CFP ($p_B/p_A = 0.6$, $\Delta G_0 = 1$ kJ/mol) and 4FW/5FW CFP ($p_B/p_A = 0.3$, $\Delta G_0 = 3$ kJ/mol) is only approximately $\Delta\Delta G_0 = 2$ kJ/mol. Therefore, weak dipole-induced dipole interactions or London dispersion forces can account for the observed $\Delta\Delta G_0$. The different interaction energies can be explained by the different polarity of the CF bond compared to a CH bond. Depending on the position of the fluorine atom the electrostatic interactions of the chromophore with surrounding residues are modulated. This would affect the interaction energy, and therefore flip rates and state populations of the exchanging moiety. The experimental data show the existence of slow exchange in all three samples indicating that the influence of different fluorination patterns ("atomic mutations") on chromophore dynamics does not alter the regime of exchange in the case of ECFP. But the observed effect on dynamics suggests that in other cases NMR studies using only a single fluorinated analogue may not be sufficient for correct interpretation and quantification of the results.

Structural Basis of Exchange. On the basis of the ^{19}F NMR data discussed above the time scale and thermodynamic properties of the exchange process between two different conformational states of ECFP were fully characterized, but the molecular structures of the two states are not determined yet.

Both tryptophans in ECFP show double peaks. Therefore, it could be thought first that this double peak behavior may be general for tryptophans, but the ^{19}F NMR spectrum of GFPuv showed only a single peak for W57. The more probable explanation of double signals for both tryptophans in the NMR spectrum of ECFP is that the signals of W57 originate from different processes than those of W66, as the resonances of W57 can be easily explained by the exchange between *cis* and *trans* proline conformation[56] in the P56−W57−P58 sequence. Whereas an exchange process of different origin causes the resonances of the chromophore tryptophan W66.

To determine the exchange process the steric interactions in the chromophore cavity of CFP have to be examined for which a crystal structure would be desirable. Since attempts to crystallize fluorinated CFP were not yet successful the structure CFP was modeled on the basis of X-ray data from the closely related 4-amino-tryptophan CFP.[53] It is well-known that even minor substitutions in the core of proteins can cause substantial structural rearrangements therefore the model may give only a first impression on the steric interactions going on in CFP; nevertheless, this provides substantial aid for interpreting the NMR data.

In principle, three possibilities for the exchange process within CFP exist that may explain our experimental data: (i) the two states correspond to the monomer−dimer states of CFP, (ii) they are related to the exchanging ground states of the chromophore itself, or (iii) side chains in the vicinity of the chromophore exchanges between two conformations.

Dimerization. The explanation that the two states are simply the monomer and dimer of the protein is the least plausible interpretation. Although GFPs are known to dimerize, the dissociation constant of wild-type GFP is approximately $K_D = 0.1$ mM as measured by analytical ultracentrifugation.[61−63] The K_D indicates weak binding for the dimer formation; in addition the symmetric dimers seen in the X-ray structures are rather loose without any interleaved monomer units[16] so that even if GFP formed a tight complex the immediate atomic environment, which to the first approximation is responsible for the difference in the chemical shift, would be almost identical for a given nucleus in the monomer and dimer unit. This is in strong contrast to the large difference in ^{19}F chemical shift ($\Delta\delta = 2.6$ ppm) between the two states, $W66_A$ and $W66_B$. An example of NMR characterization of a tight nM range dimer, with interleaved monomers, is provided by Rous sarcoma virus protease.[64] Even in this case, a single set of NMR peaks was observed in the ^1H-^{15}N correlation spectrum, indicating the monomer−dimer equilibrium to be in fast exchange with respect to the NMR time scale.

The observed NMR resonances in our case are also a population-weighted monomer−dimer ensemble average. To definitely exclude that the $W66_A$ and $W66_B$ states represent the monomer−dimer equilibrium, we have recorded the ^{19}F NMR spectra at different concentrations of ECFP. Since the total exchange k_{ex} rate is in the order of 700 s^{-1}, any concentration dependent population change should be detectable within the long measurement time. Assuming that the two states are caused by the exchange between monomers and dimers, the higher ($W66_A$) and lower ($W66_B$) energy states had to be attributed to the monomer and dimer states, respectively. At a total protein concentration of 0.3 mM (Figure 7, top spectrum) the populations of the monomer $W66_A$ is 63%, which would lead to a dissociation constant of $K_D = [W66_A]^2/[W66_B] = 0.4$ mM. At the lowest total protein concentration of 0.02 mM (Figure 7, bottom spectrum) a K_D of 0.4 mM would result in a monomer fraction of 95%.[65] As can be seen from Figure 7 variation of the protein concentration does not affect the relative proportions of the $W66_A$ and $W66_B$ states.

Chromophore Ring Rotation. Aromatic ring flipping is a well-known phenomenon found in proteins (e.g., ref 66). To evaluate the possibility of the CFP chromophore being directly

(61) Phillips, G. N. In *Green Fluorescent Protein*; Chalfie, M., Kain, S., Eds.; Wiley-Liss: New York, 1998.
(62) Phillips, G. N. *Curr. Opin. Struct. Biol.* **1997**, *7*, 821−827.
(63) Ward, W. W. In *Green Fluorescent Protein*; Chalfie, M., Kain, S., Eds.; Wiley-Liss: New York, 1998.
(64) Schatz, G. W.; Reinking, J.; Zippin, J.; Nicholson, L. K.; Vogt, V. M. *J. Virol.* **2001**, *75*, 4761−4770.
(65) Korchuganov, D. S.; Nolde, S. B.; Reibarkh, M. Y.; Orekhov, V. Y.; Schulga, A. A.; Ermolyuk, Y. S.; Kirpichnikov, M. P.; Arseniev, A. S. *J. Am. Chem. Soc.* **2001**, *123*, 2068−2069.
(66) Skalicky, J. L.; Mills, J. L.; Sharma, S.; Szyperski, T. *J. Am. Chem. Soc.* **2001**, *123*, 388−397.

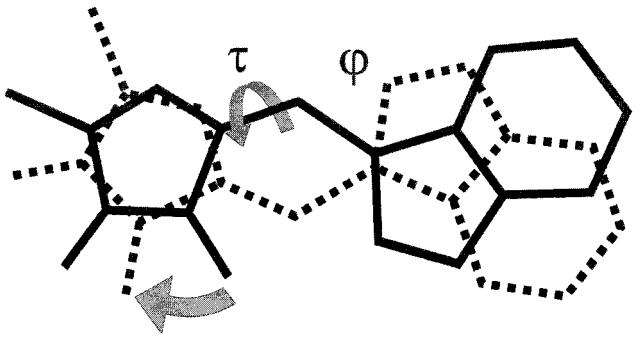

Figure 8. Possible rotamer conformations of the CFP chromophore are illustrated. A rotation around τ is accompanied by an in-plane rotation of the imidazolinone ring.

involved in such an exchange process, an examination of the known GFP structures is useful. In GFP, the chromophore is stabilized by a hydrogen bond network, which mainly involves its phenolic hydroxyl group and H148 and T203. Despite tight packing in the chromophore cavity, it was shown by molecular dynamic simulations that concerted rotations around the τ and φ angles, so-called hula-twists, are a possible pathway of photoisomerization of the GFP chromophore since electronic excitation reduces the energy barrier for these rotations.[14,25] Without electronic excitation the hula-twist will occur only with a reduced rate.

The model of CFP structure predicts an overall structure similar to the well-known GFP structures, as expected. In CFP, a rotation of the chromophore around the φ axis is very unlikely since this would cause a rotation of the long axis of the indole, which is hindered by surrounding residues. But a rotation around the τ axis in concert with an in-plane rotation of the imidazolinone ring appears to be possible (Figure 8), since both rotamers $\tau = 0°$ and $\tau = 180°$ fit into the same chromophore cavity.

Weber et al. showed that the *cis* and *trans* isomers of the same protonation state in GFP are similar in their heats of formation (<5 kJ/mol).[14] Applied to ECFP, this is in agreement with the small energy difference determined by NMR. Weber et al. also calculated the energy barriers for bond rotations in the ground state of the GFP chromophore.[14] Most of them are in the range of 100−200 kJ/mol, except for example the φ rotation in the neutral and anionic state or the τ rotation in a zwitterionic state, which have lower energy barriers. Since the free energy of activation for the exchange process was estimated to be in the range of 57−63 kJ/mol, the rotation around one of the bonds connecting the two rings of the chromophore appears to be feasible for the indole−imidazolinone chromophore of CFP. One might argue that in CFP a hydrogen bond between the indole amide and T203 may exist. However, the breaking of a hydrogen bond introduces an energy barrier of typically 20 kJ/mol, which is only a fraction of the 60 kJ/mol estimated as the height of the overall energy barrier, therefore not prohibitive for the proposed exchange process.

Multiple forms of the chromophore should be detectable as a poorly defined or extra electron density of the chromophore in crystal structures. In the crystal structures of GFP the chromophore is usually well defined, although some structures with a poor electron density around the chromophore have been reported (e.g., ref 67). The model of the CFP structure indicates a well-defined chromophore. Therefore chromophore isomerization is relatively unlikely, but cannot be excluded by the present data.

Conformational Exchange of H148. Another possible reason for slow motion processes is the exchange between different conformations of neighboring side chains. In GFP, two residues, H148 and T203, are closest to the phenolic ring of the chromophore. In most of the crystal structures of GFPs, the β-barrel is somewhat perturbed around H148 with residues 144−150 not being hydrogen bonded to the adjacent backbone residues 165−170.[68] This suggests that H148 has an increased flexibility compared to other residues within the β-barrel. The comparison of the CFP model with the GFP structure suggests that it is sterically feasible for H148 in CFP to adopt two conformations: one with the H148 side chain oriented toward the chromophore like in GFP and another one with H148 pointing toward the solvent (see Figure 9).

The different conformations are expected to introduce distortions in the backbone conformation of the adjoining residues. The hydrogen bond between H148 and the chromophore is removed in CFP, at least in one conformation. If the histidine side chain flips from a chromophore-neighboring conformation to a solvent-exposed conformation, the chemical shift of the fluorine atoms at the 4-, 5-, or 6-position of the indole will be affected by the change in electronic environment. A flipped H148 side chain opens a channel to the solvent and additionally gives the chromophore a larger degree of motional freedom. But the NMR pH titration clearly shows that the chromophore is not significantly affected by pH changes. Either surrounding hydrophobic residues hinder the entry of water molecules or the imino nitrogen is involved in a hydrogen bond and therefore does not titrate in the measured range of pH. Assuming that the solvent exposed conformation of H148 is higher in energy, the increased disorder associated with H148 flipping from its position in the protein core to a more flexible, solvent-exposed position may account for the observed enthalpy−entropy compensation.

Taken together the exchange of H148 between two conformational states is the most probable explanation for our NMR observations, but more X-ray studies are necessary to provide direct evidence.

Effects of Exchange on Optical Properties. The quantum yield of CFP (QY = 0.4) is approximately one-half compared to GFP (QY = 0.8).[2] The differences in quantum yield are commonly explained by the competition of radiationless and emitting decay channels (e.g. ref 14). A large indole-bearing chromophore, such as the one of CFP, is expected to show a reduced rate of isomerization compared with phenyl or imidazole rings. This explains the observed differences in fluorescence intensities (F < H < W),[14] but the fluorescence of CFP is still substantially weaker than in GFP. Although the exchange process observed in CFP occurs on a much longer time scale (~1.3 ms) than that for the electronic excited-state lifetime (e.g., ~3.3 ns for GFP[3,4]), it still may influence fluorescence properties and therefore explain the differences in quantum yield of fluorescence for GFP and CFP. For example, the flipping of

(67) Battistutta, R.; Negro, A.; Zanotti, G. *Proteins: Struct., Funct., Genet.* **2000**, *41*, 429−437.
(68) Wachter, R. M.; Elsiger, M. A.; Kallio, K.; Hanson, G. T.; Remington, S. J. *Structure* **1998**, *6*, 1267−1277.
(69) Palm, G. J.; Zdanov, A.; Gaitanaris, G. A.; Stauber, R.; Pavlakis, G. N.; Wlodawer, A. *Nat. Struct. Biol.* **1997**, *4*, 361.
(70) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.

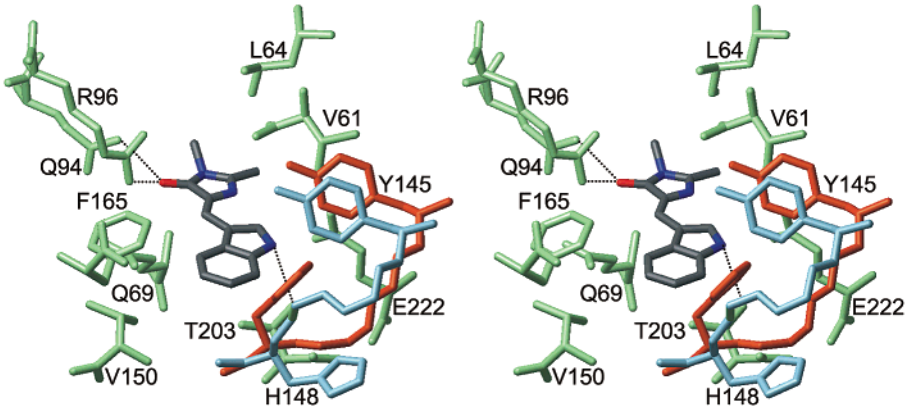

Figure 9. Stereoview of the chromophore environment in the CFP model structure: Residues surrounding the chromophore are shown in green. The chromophore itself is shown as a stick representation; the carbon, nitrogen, and oxygen atoms are depicted in black, blue, and red, respectively. The comparison of the conformation around residue H148 for GFP (in this case taken from PDB entry 1EMM[69]) and the CFP model is shown in orange and cyan, respectively. Close inspection of both structures suggests that the conformation present in GFP may also be adopted in CFP. Hydrogen bonds are depicted as dotted lines. The figure was prepared using the program MOLMOL.[70]

H148 opens a channel to solvent that may allow fluorescence quenchers such as oxygen to enter the chromophore cavity more easily. Additionally tryptophan is uniquely sensitive to collisional quenching, apparently due to a tendency of indole to donate electrons while in the excited state.[59]

For applications in molecular biology a higher quantum yield of CFP fluorescence is desirable. To achieve an enhancement of the fluorescence the influence of the slow exchange process has to be reduced. Therefore replacement of H148 by an amino acid that (i) is not able to get in contact with the chromophore and (ii) stabilizes the backbone conformation should enhance the fluorescence of CFP. If the chromophore itself exchanges between different conformations, it has to be stabilized. For example, the mutation of T203 in CFP to an aromatic amino acid may lead to the formation of a stacked π-electron system like in YFP, which is likely to reduce slow exchange motions within the chromophore and therefore may increase the quantum yield of fluorescence.

Conclusions

In summary, in our study the concerted use of ^{19}F NMR relaxation and dynamic NMR provides direct evidence for the existence of a slow exchange process between two states on the millisecond time scale (1.3 ms) including or near the chromophore of CFP. The differences in entropy and enthalpy between the two states were quantified and showed an enthalpy−entropy compensation character. The chromophore itself exhibits no pronounced flexibility on the picosecond to nanosecond time scale. The structural origin of the exchanging states has not yet been fully clarified, but it seems to be clear that either the chromophore itself or more likely the side chain of H148 exchanges between two conformations. The differences in thermodynamic properties found for the different fluorination sites show that the introduction of fluorine into the protein has a small but observable effect on dynamic parameters. Therefore, fluorine cannot be considered purely as a noninteracting "reporter atom" in ^{19}F NMR studies. However, since the introduction of fluorine does not alter the time scales of dynamic processes, useful information can be extracted by comparing proteins with different fluorine-labeled sites. This study not only proves the applicability of ^{19}F NMR to explore the chromophore dynamics of green fluorescent proteins but also demonstrates the use of an "atomic mutation" methodology to investigate proteins that contain aromatic amino acids. The knowledge of dynamics of ECFP opens several possibilities for manipulating the chromophore of CFP to improve the fluorescence quantum yield of CFPs. This may lead to the development of new, optimized variants of GFPs.

Abbreviations

GFP, green fluorescent protein; CFP, cyan fluorescent protein; YFP, yellow fluorescent protein; EGFP, enhanced GFP; ECFP, enhanced CFP; Trp, tryptophan; 4FW, 4-fluorotryptophan; 5FW, 5-fluorotryptophan; 6FW, 6-fluorotryptophan; TFA, trifluoroacetic acid; δ, chemical shift; $\Delta\nu$, NMR line width; {I}-S NOE, nuclear Overhauser effect on nucleus S by saturating nucleus I; f, NOE enhancement factor; T_1, longitudinal relaxation time; T_2, transversal relaxation time; J, spectral density of rotational diffusion; S^2, generalized order parameter; τ_C, overall rotational correlation time; τ_i, correlation time of internal motions; p_A, p_B, populations of states A and B; k_A, k_B, off-rates of states A and B; $\bar{K}$, equilibrium constant; R, universal gas constant; ΔH_0, reaction enthalpy; ΔS_0, reaction entropy; ΔG_0, reaction free energy; ΔG^{act}, activation energy; $\Delta\sigma$, anisotropy of chemical shift tensor; η, asymmetry of chemical shift tensor; $^3J_{FH}$, three-bond J-coupling constant; $^4J_{FH}$, four-bond J-coupling constant; ω, Larmor frequency.

Acknowledgment. This work was supported by a Grant of Sonderforschungsbereich 533. M.H.J.S. thanks Dr. Ingo Krossing and Dr. Ruth Pfänder for stimulating discussions. We acknowledge the work of Dr. Julia Georgescu on ^{15}N labeled GFPuv. M.K.A. is a receipient of a fellowship from the Alexander von Humbolt Foundation. We are indebted to Jae Hyun Bae for providing us with his unpublished experimental data.

Supporting Information Available: Figures illustrating the effect of pH and UV irradiation on ^{19}F NMR spectra of 6FW ECFP and temperature-dependent optical emission and excitation spectra (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0257725