J|A|C|S
ARTICLES
Published on Web 07/29/2006

^{19}F NMR Studies of the Native and Denatured States of Green Fluorescent Protein

Farid Khan,[†,‡] Ilya Kuprov,[§] Timothy D. Craggs,[†] P. J. Hore,[§,*] and Sophie E. Jackson[†,*]

Contribution from the University of Cambridge, Chemistry Department, Lensfield Road, Cambridge CB2 1EW, United Kingdom, and University of Oxford, Department of Chemistry, Physical and Theoretical Chemical Laboratory, South Parks Road, Oxford, OX1 3QZ, United Kingdom

Received January 26, 2006; E-mail: sej13@cam.ac.uk; peter.hore@chem.ox.ac.uk

Abstract: Biosynthetic preparation and ^{19}F NMR experiments on uniformly 3-fluorotyrosine-labeled green fluorescent protein (GFP) are described. The ^{19}F NMR signals of all 10 fluorotyrosines are resolved in the protein spectrum with signals spread over 10 ppm. Each tyrosine in GFP was mutated in turn to phenylalanine. The spectra of the Tyr → Phe mutants, in conjunction with relaxation data and results from ^{19}F photo-CIDNP (chemically induced dynamic nuclear polarization) experiments, yielded a full ^{19}F NMR assignment. Two ^{19}F-Tyr residues (Y92 and Y143) were found to yield pairs of signals originating from ring-flip conformers; these two residues must therefore be immobilized in the native structure and have ^{19}F nuclei in two magnetically distinct positions depending on the orientation of the aromatic ring. Photo-CIDNP experiments were undertaken to probe further the structure of the native and denatured states. The observed NMR signal enhancements were found to be consistent with calculations of the HOMO (highest occupied molecular orbital) accessibilities of the tyrosine residues. The photo-CIDNP spectrum of native GFP shows four peaks corresponding to the four tyrosine residues that have solvent-exposed HOMOs. In contrast, the photo-CIDNP spectra of various denatured states of GFP show only two peaks corresponding to the ^{19}F-labeled tyrosine side chains and the ^{19}F-labeled Y66 of the chromophore. These data suggest that the pH-denatured and GdnDCl-denatured states are similar in terms of the chemical environments of the tyrosine residues. Further analysis of the sign and amplitude of the photo-CIDNP effect, however, provided strong evidence that the denatured state at pH 2.9 has significantly different properties and appears to be heterogeneous, containing subensembles with significantly different rotational correlation times.

Introduction

Green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* is widely used as a fluorescent marker in many areas of biological research, owing to its unique photophysical properties.[1] The 238-residue protein undergoes an autocatalytic post-translational cyclization and oxidation of the polypeptide chain around residues Ser65, Tyr66, and Gly67, forming an extended and rigidly immobilized conjugated π-system, the chromophore, which emits green fluorescence.[2] No cofactors are necessary for either the formation or the function of the chromophore,[3] which is embedded in the interior of the protein surrounded by an 11-stranded β-barrel, Figure 1.[4,5] GFP has a high (~70%) fluorescence quantum yield due to the lack of vibrational relaxation, because in the native state, the chromophore is rigidly held and shielded from bulk solvent.[4,5] On protein denaturation, the chromophore remains chemically intact but fluorescence is lost. The green fluorescence is, therefore, a sensitive probe of the state of the protein.

GFP has been extensively engineered to modify and improve its spectroscopic and physical properties to facilitate its use as a marker of gene expression and protein localization, an indicator of protein-protein interactions, and its use as a biosensor.[1] In all cases, GFP needs to fold efficiently to function in these different biological assays. In addition to its use as a reporter protein, GFP is also a model system for photophysical studies[2] and for folding studies on large single domain proteins.[8−10] Despite the widespread use of GFP and level of interest in this protein, relatively little is known about its folding either in vitro or in vivo. A recent study by Kuwajima and coworkers has provided the most detailed information to date.[8]

[†] University of Cambridge.
[‡] Current address: Protein Technologies Laboratory, Babraham Institute, Babraham, Cambridge CB2 4AT, UK.
[§] University of Oxford.

(1) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509−544.
(2) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759−781.
(3) Reid, B. G.; Flynn, G. C. *Biochemistry* **1997**, *36*, 6786−6791.
(4) Ormo, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392−1395.
(5) Yang, F.; Moss, L.; Phillips, G. *Nat. Biotech.* **1996**, *14*, 1246−1251.
(6) Battistutta, R.; Negro, A.; Zanotti, G. *Proteins: Struct. Func., Gen.* **2000**, *41*, 429−437.
(7) DeLano, W. L. *World Wide Web* **2002**, http://www.pymol.org.
(8) Enoki, S.; Saeki, K.; Maki, K.; Kuwajima, K. *Biochemistry* **2004**, *43*, 14238−48.
(9) Fukuda, H.; Arai, M.; Kuwajima, K. *Biochemistry* **2000**, *39*, 12025−12032.
(10) Stepanenko, O. V.; Verkhusha, V. V.; Kazakov, V. I.; Shavlovsky, M. M.; Kuznetsova, I. M.; Uversky, V. N.; Turoverov, K. K. *Biochemistry* **2004**, *43*, 14913−14923.

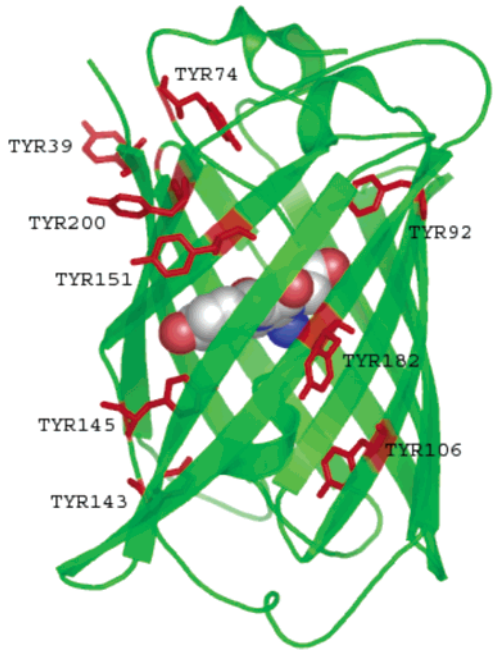

Figure 1. Ribbon diagram of the backbone of GFP showing the chromophore (space filled) and tyrosine residues (red). The structure is that of the GFPuv used in this study.[6] The figure was created using PyMol.[7]

In their study, fluorescence and far-UV circular dichroism were used to probe folding and unfolding and a model proposed in which GFP folds through several intermediate states. Further studies using complementary techniques and probes are clearly necessary to obtain a complete understanding of the folding pathway of this large, complex, and important protein.

NMR spectroscopy has proven to be a powerful method of studying structured, partially structured, and denatured states of proteins, including molten globules and kinetic intermediates.[11] In the cases where NMR can be used, it provides valuable residue-specific information on structure and dynamics and therefore complements optical techniques such as fluorescence and circular dichroism that only report on global properties. ^{1}H, ^{13}C, and ^{15}N are the most commonly used nuclei. Quite recently, however, much progress has been made on the study of proteins using ^{19}F NMR.[12−17]

A distinctive feature of ^{19}F NMR spectroscopy is the strong chemical shielding anisotropy (CSA) of the fluorine nucleus. The nuclear relaxation rate due to CSA is quadratic in magnetic induction, and for a protein the size of GFP at 10−20 T, the transverse ^{19}F relaxation is dominated by the CSA term and is very fast. This leads to broad spectral lines (50−400 Hz) and reduced detection sensitivity. Fast transverse relaxation of ^{19}F in proteins also makes magnetization transfer steps impossible in all but a few two-dimensional (2D) NMR experiments.[18]

Despite these seemingly formidable difficulties, ^{19}F NMR of fluorine-labeled biological macromolecules is gaining popularity, the reason being that the ^{19}F nucleus has a large chemical shift dispersion and its chemical shielding is very sensitive to changes in molecular environment.[19] In a protein containing 10 fluorinated tyrosines, one may generally expect that all 10 will be resolved in the native-state spectrum and that any structural transformation will be accompanied by significant changes in chemical shifts. Analysis of these changes can give direct information on the kinetics of the structural transformation in question.[12] In addition to the information obtained from chemical shifts, structural information can also be gathered from relaxation data. ^{19}F relaxation behavior is well understood and, in principle, allows the extraction of local and global motion correlation times, order parameters and rotational diffusion tensors,[18] which are useful as indicators of the local mobility and stability of the macromolecule.

Another unusual property of the fluorine nucleus, namely the large ^{19}F hyperfine coupling constants of fluorinated aromatic radicals, makes it particularly useful in NMR experiments with photochemical pumping of magnetization through the photo-CIDNP (chemically induced dynamic nuclear polarization) effect.[20−21] CIDNP is the non-Boltzmann population of nuclear spin states produced in chemical reactions that proceed through radical pair intermediates.[22] It is manifested by strongly enhanced NMR absorption or emission and has been used to detect solvent-accessible tyrosine, tryptophan, and histidine residues in proteins.[23−25] In high magnetic fields, the amplitude of the photo-CIDNP effect is directly proportional to the hyperfine coupling constant in the intermediate radical. Fluorine-containing molecules are quite unique in this respect because the ^{19}F hyperfine constants in aromatic radicals are large, owing to both the high magnetogyric ratio of the ^{19}F nucleus and the strong electronegativity of the fluorine atom.[26]

In this paper, we present NMR studies on the native and denatured states of GFP, which provide important information on the initial and final states of the protein on the folding pathway. We describe the strategies employed to produce 3-fluorotyrosine-labeled GFP. Protein engineering is used systematically to replace each of the tyrosines in turn to facilitate the assignment of the ^{19}F NMR spectrum. ^{19}F photo-CIDNP techniques are then used to characterize the native and denatured states of GFP. In particular, the solvent accessibility of the 10 tyrosines in GFP is assessed in both the native state and the acid- and guanidinium chloride-induced denatured states, providing important information on the starting point of kinetic refolding studies.

Materials and Methods

Construction of the trGFPuv Expression Vector. The gene encoding GFPuv (Clontech) was cloned into a modified pRSET vector (Invitrogen). GFPuv has three mutations (F99S, M153T, V163A) that enhance solubility and fluorescence.[27] The GFPuv gene was subsequently truncated to its minimal domain for green fluorescence (Met1-

(11) Redfield, C. In *Methods in Molecular Biology*; Protein NMR Techniques, Vol. 278; Humana Press: Totowa, NJ, 2004; pp 233−254.
(12) Bann, J. G.; Frieden, C. *Biochemistry* **2004**, *43*, 13775−13786.
(13) Hull, W. E.; Sykes, B. D. *Biochemistry* **1974**, *13*, 3431−3437.
(14) Li, H.; Frieden, C. *Biochemistry* **2005**, *44*, 2369−2377.
(15) Shu, Q.; Frieden, C. *J. Mol. Biol.* **2004**, *345*, 599−610.
(16) Sykes, B. D.; Hull, W. E. *Methods Enzymol.* **1978**, *49*, 270−295.
(17) Wang, X.; Mercier, P.; Letourneau, P.-J.; Sykes, B. D. *Prot. Sci.* **2005**, *14*, 2447−2460.
(18) Peng, J. W. *J. Magn. Res.* **2001**, *153*, 32−47.
(19) Gakh, Y. G.; Gakh, A. A.; Gronenborn, A. M. *Magn. Res. Chem.* **2000**, *38*, 551−558.
(20) Goez, M. *Concepts Magn. Reson.* **1995**, *7*, 69−86.
(21) Stob, S.; Kaptein, R. *Photochem. Photobiol.* **1989**, *49*, 565−577.
(22) Goez, M. *Adv. Photochem.* **1997**, *23*, 63−163.
(23) Hore, P. J.; Broadhurst, R. W. *Prog. Nucl. Magn. Reson. Spectrosc.* **1993**, *25*, 345−402.
(24) Kaptein, R. *Bio. Magn. Reson.* **1982**, *4*, 145−91.
(25) Kaptein, R.; Dijkstra, K.; Nicolay, K. *Nature* **1978**, *274*, 293−4.
(26) Brinkman, M. R.; Bethell, D.; Hayes, J. *J. Chem. Phys.* **1973**, *59*, 3431−4.
(27) Crameri, A.; Whitehorn, E. A.; Tate, E.; Stemmer, W. P. C. *Nat. Biotech.* **1996**, *14*, 315−319.

Ile229)[28] by the introduction of a stop codon at position 230 using PCR (polymerase chain reaction) techniques. The resulting plasmid (ptrGFPuv) was fully sequenced. The wild-type GFP referred to throughout this paper is a pseudo wild-type GFP corresponding to the trGFPuv construct described above.

Site-Directed Mutagenesis of Tyrosine Residues. All 10 tyrosine residues were systematically mutated to phenylalanine in ptrGFPuv using the QuikChange kit from Strategene. All mutations were confirmed by DNA sequencing of the truncated GFP gene.

Expression and Purification of Fluorotyrosine-Labeled GFP Mutants. The incorporation of 3-fluorotyrosine was successfully achieved using a modification of the method of Kim and co-workers.[29] In this approach, the herbicide glyphosate was used to inhibit the shikimate pathway in cultures of *Escherichia coli*. Briefly, single colonies of transformed *E. coli* cells (BL21 DE3) harboring ptrGFPuv were picked from TYE ampicillin plates and were used to inoculate 5 mL of 2 × TY media (containing 0.1 mg mL^{-1} ampicillin) and were grown overnight in a shaker at 37 °C. This overnight culture was used to inoculate 400 mL of 2 × TY media containing of 0.1 mg mL^{-1} ampicillin. This preculture was left to grow to a cell density with an A_{600} of between 0.8 and 1.0 in a shaker at 37 °C. The 2 × TY preculture was spun down (50 mL) at 25 °C for 10 min at 4000 rpm (SLC4000 rotor, Sorvall). Meanwhile, two 50-mL aliquots of minimal media were taken from a flask containing 500 mL of M9 minimal media. One 50-mL aliquot was used to resuspend the preculture cell pellet and was added to the remaining 400 mL minimal media. This main culture was allowed to grow to a cell density with an A_{600} of 0.6−0.7 in a shaker at 25 °C before GFP expression was induced by the addition of IPTG to a final concentration of 0.5 mM (for the Y66F and Y74F, the temperature was dropped to 18 °C). For selective fluorotyrosine labeling, 35 mg of 3-fluoro-D,L-tyrosine (96%, Lancaster), 30 mg of L-phenylalanine (Sigma), and 30 mg of L-tryptophan (Sigma) with 0.5 g of glyphosate (99.5% N-(phosphonomethyl)glycine, Sigma) were dissolved in the remaining 50 mL of minimal medium. This labeling mixture was added 30 min before induction to the cell culture, which was then left to grow overnight. The cultures were harvested by centrifugation (Sorvall, SLC4000 rotor, 15 min at 8000 rpm at 4 °C) and resuspended into 50 mL of buffer A (50 mM Tris, pH 8.0). The cells were lysed either by pulsed sonication on ice for 4 min or in a cell homogenizer (Glen Creston, UK) by pressure disruption at 4 °C and then sonicated for 1 min to make the lysate less viscous. The lysate was centrifuged at 18 000 rpm (SS34, Sorvall) at 4 °C for 45 min. The supernatant was then pooled and loaded onto a Q-Sepharose (Pharmacia) column at 2 mL min^{-1} until the green protein was visibly bound to the top of the column. After washing with 6 column volumes of buffer A (50 mM Tris, pH 8.0), trGFPuv was eluted with a 0−60 min gradient using buffer B (50 mM Tris, pH 8.0, 0.5 M NaCl). The green fractions were pooled and concentrated (Vivaspins 20, Vivascience) and washed through with PBS. The concentrated protein (∼10 mL) was then loaded onto a HiLoad 26/60 Superdex G75 column (Amersham Biosciences), pre-equilibrated in buffer A, and run at 2 mL min^{-1}. The visibly green fractions that eluted were pooled. Proteins were pure as assessed by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) and LC-MS. The labeling efficiency was also confirmed by LC-MS (for reference, the fully labeled, pseudo-wild-type protein showed a single HPLC peak of 25 806 Da corresponding to the incorporation of ten 3-fluorotyrosines). Pooled fractions were concentrated to 1 mM and either stored at −80 °C or used immediately.

Mass Spectrometry. Electrospray ionization mass spectrometry was performed using a Micromass/Waters Quattro LC instrument connected to a Prodigy C8 (250 mm × 10 mm) reverse phase (RP) column for protein analysis. Papain digested peptide mass analysis was done according to the method of Cody and co-workers[30] on a Jupiter RP−C18 (250 mm × 10 mm) column using a gradient of 0−100% acetonitrile (ACN)/0.1% trifluoroacetic acid (TFA) buffer over 30 min.

^{19}F NMR and CIDNP. Conventional ^{19}F NMR spectra were recorded with 5K − 10K scans, with a 12 kHz sweep width and free induction decays of 4K points. The pulse width was 60°, and the pulse spacing was 1.5 s (the ^{19}F T_1 times in GFP range from 0.5 to 2 s; T_2 times are 10−100 ms). The relatively short relaxation delay, which was essential for sensitivity reasons, has the consequence that the relative intensities of the ^{19}F signals are not proportional to numbers of spins as they would be in a fully relaxed spectrum. However, the distortion is not severe and at no point are the signal intensities interpreted quantitatively.

377 MHz ^{19}F NMR measurements were performed on 1 mM protein samples in 10% D_2O with PBS (phosphate-buffered saline) buffer at 300 K using a Bruker DRX400 spectrometer equipped with a 5-mm QNP probe. ^{19}F chemical shifts were referenced to external trifluoroacetic acid.

564 MHz ^{19}F NMR and photo-CIDNP spectra were recorded at 300 K on a Varian Inova 600 (14.1 T) NMR spectrometer equipped with a 5 mm ^{19}F{^{1}H} *z*-gradient probe. The light source was a Spectra Physics BeamLok 2080 argon ion laser, operating in single-line mode at 5 W output power at a wavelength of 514 nm. A mechanical shutter (NM Laser Products LS200) controlled by the spectrometer was used to produce 100 ms light pulses. The light was focused into a 6-m length of optical fiber (Newport F-MBE) using a Newport M-5X objective lens. The other end of the fiber was attached (via Newport SMA connectors) to a 2-m section of the same fiber whose stepwise-tapered tip[31] was held inside a 5-mm NMR tube by a truncated Wilmad WGS 5BL coaxial insert.

Before each photo-CIDNP experiment, the sample was repeatedly dialyzed against 50 mM phosphate buffer in D_2O at pH 7.2 (uncorrected for deuterium isotope effect) for several days to complete the proton exhange of both the solvent and the protein. Deuterated guanidinium chloride (GdnDCl), NaOD, and DCl were used as necessary to reach the prescribed conditions. During each scan of the photo-CIDNP experiment, a 0.2 mM fluorotyrosinated GFP sample containing 1 mM flavin mononucleotide as photosensitizer was irradiated for 100 ms and subjected to a 90° radio frequency pulse on ^{19}F followed by immediate acquisition of the free induction decay. After application of a shifted Gaussian window function, zero-filling, and Fourier transformation, the spectra were analyzed using mixed Lorentzian−Gaussian line-fitting. Photo-CIDNP spectra are shown as light minus dark difference spectra recorded with 16 scans, with a 60 s delay between scans, 12 kHz sweep width, and 4K points in the free induction decays. The large signal enhancement afforded by the CIDNP effect means that high-quality spectra can be obtained using relatively few scans together with a relaxation delay that is much longer than the ^{19}F T_1s. The photo-CIDNP spectra are thus "fully relaxed". In general, the CIDNP intensities in spectra recorded with a continuous wave laser reflect the spin−lattice relaxation rates of the individual nuclei in the protein; however, as the laser irradiation time (100 ms) is much shorter than the shortest T_1 (∼500 ms), the observed intensities are not distorted by relaxation in this case.[32] There are well-established differences between time-resolved ^{1}H CIDNP spectra, obtained with a nanosecond laser and microsecond time delays, and spectra recorded with a continuous wave laser and much lower time-resolution.[33] However, these effects are much smaller for ^{19}F than for ^{1}H CIDNP because of the very fast ^{19}F spin−lattice

(28) Li, X.; Zhang, G.; Ngo, N.; Zhao, X.; Kain, S. R.; Huang, C. C., *J. Biol. Chem.* **1997**, *272*, 28545−28549.
(29) Kim, H. W.; Perez, J. A.; Ferguson, S. J.; Campbell, I. D. *FEBS Lett.* **1990**, *272*, 34−36.
(30) Cody, C. W.; Prasher, D. C.; Westler, W. M.; Prendergast, F. G.; Ward, W. W. *Biochemistry* **1993**, *32*, 1212−1218.
(31) Kuprov, I.; Hore, P. J. *J. Magn. Res.* **2004**, *171*, 171−175.
(32) Hore, P. J.; Egmond, M. R.; Edzes, H. T.; Kaptein, R. *J. Magn. Res.* **1982**, *49*, 122−150.
(33) Morozova, O. B.; Yurkovskaya, A. V.; Sagdeev, R. Z.; Mok, K. H.; Hore, P. J. *J. Phys. Chem. B* **2004**, *108*, 15355−15363.

relaxation in the fluoro-tyrosine radical resulting from the highly anisotropic ^{19}F hyperfine interaction;[34] we believe them to be negligible here.

Relaxation Analysis. The relaxation of ^{19}F nuclei in GFP-sized diamagnetic molecules in solution occurs primarily as a result of rotational modulation of two anisotropic interactions: the dipole-dipole interaction with nearby protons (DD mechanism) and the Zeeman interaction with the applied magnetic field (chemical shift anisotropy, or CSA mechanism). A Bloch-Redfield-Wangsness relaxation theory treatment[35] results in the following expressions for the transverse relaxation rate of the fluorine nucleus:

$$\left(\frac{1}{T_2}\right)_{DD} = \frac{\delta_{DD}^2 + 3\eta_{DD}^2}{72}[4J(0) + 6J(\omega_H) + 3J(\omega_F) + 6J(\omega_H + \omega_F) + J(\omega_H - \omega_F)]$$

$$\left(\frac{1}{T_2}\right)_{CSA} = \frac{\delta_{CSA}^2 + 3\eta_{CSA}^2}{72}[4J(0) + 3J(\omega_F)] \quad (1)$$

where δ_A and η_A denote the axiality and rhombicity of the point dipolar interaction (A = DD) and the Zeeman interaction (A = CSA):

$$\delta_A = 2A_{zz} - (A_{yy} + A_{xx}) \quad \eta_A = A_{yy} - A_{xx}$$

$$\delta_{DD} = 3\frac{\gamma_H\gamma_F\hbar}{r_{HF}^3}\frac{\mu_0}{4\pi} \quad \eta_{DD} = 0$$

$$\delta_{CSA} = [2\sigma_{zz} - (\sigma_{yy} + \sigma_{xx})]B_0\gamma_F \quad \eta_{CSA} = (\sigma_{yy} - \sigma_{xx})B_0\gamma_F \quad (2)$$

r_{HF} is the proton-fluorine distance, ω_H and ω_F are the ^{1}H and ^{19}F Larmor frequencies, σ_{ii} are the eigenvalues of the shielding tensor, γ_H and γ_F are magnetogyric ratios, and B_0 is the magnetic field strength. We use the Lipari-Szabo restricted motion model for the spectral density $J(\omega)$[36,37]

$$J(\omega) = \frac{2}{5}\left(\frac{S^2\tau_c}{1 + \omega^2\tau_c^2} + \frac{(1 - S^2)\tau_e}{1 + \omega^2\tau_e^2}\right) \quad \tau_e = \left(\frac{1}{\tau_c} + \frac{1}{\tau_i}\right)^{-1} \quad (3)$$

in which S^2 is an order parameter indicating the extent to which motion is restricted, τ_c is the global molecular rotational correlation time (assumed to be isotropic), and τ_i is the internal motional correlation time of the amino acid residue in question. Note that τ_i does not include contributions from fluorotyrosine ring flips (see below), which are too slow to make a significant contribution to the relaxation.

At magnetic field strengths of 10−15 T, the typical Zeeman interaction anisotropy of an aromatic ^{19}F nucleus is $\delta_{CSA} = (5-8) \times 10^5$ rad s^{-1},[38] compared to $\delta_{DD} = (1-2) \times 10^5$ rad s^{-1}, meaning that $\delta_{DD}^2 + 3\eta_{DD}^2 \ll \delta_{CSA}^2 + 3\eta_{CSA}^2$. The DD contribution to the relaxation rate may, therefore, be neglected.

It has relatively recently become possible to obtain the chemical shielding anisotropy parameters from *ab initio* calculations.[38] Values of the anisotropy of the ^{19}F chemical shielding tensor were obtained using the Gaussian03 program[39] at three different levels of theory (Table 1). The CSGT DFT B3LYP/cc-pVDZ method is recommended by Tormena et al.[40] for accurate calculation of ^{13}C chemical shifts. The intermediate GIAO DFT B3LYP/6-311++G(2d,2p) method is a general-purpose technique for chemical shift calculation that has been shown to provide reasonably accurate results for most Period II elements. Although, in general, one would expect a better estimate from a higher level method, ^{19}F chemical shift calculations are known to disobey this rule, and the least computationally expensive in vacuo GIAO HF/6-311++G(2d,2p) method yields values of the ^{19}F chemical shifts that are superior to those obtained with DFT and MP2.[41] We therefore chose to use the GIAO HF value for the relaxation analysis. DFT methods are known to overestimate ^{19}F shielding in fluorinated aromatics,[38] as is also seen here.

Table 1. Computed Parameters of the ^{19}F Chemical Shift Tensor in 3-Fluorotyrosine

method (geometry/NMR)	$\lvert 2\sigma_{zz} - (\sigma_{xx} + \sigma_{yy})\rvert$ /ppm	$\lvert \sigma_{xx} - \sigma_{yy}\rvert$ /ppm
DFT B3LYP 6-311++G(2d,2p)/ CSGT DFT B3LYP cc-pVDZ	238	44
DFT B3LYP 6-311++G(2d,2p)/ GIAO DFT B3LYP 6-311++G(2d,2p)	271	54
DFT B3LYP 6-311++G(2d,2p)/ GIAO HF 6-311++G(2d,2p)	210	44

Once the chemical shift tensor is known, measuring relaxation times at two different magnetic fields allows extraction of the order parameters and the correlation times. It is particularly convenient to use T_2 values for fluorinated proteins, both because of signal-to-noise and degassing difficulties associated with T_1 measurements and because, for signals broader than approximately 50 Hz, accurate estimates of T_2s may be obtained simply from line widths (the broadening arising from field inhomogeneity being negligible).

Results

Mass Spectrometric Analysis of Labeled and Nonlabeled GFPs. The incorporation of the fluorinated tyrosine analogue into wild-type and mutant GFP was analyzed using liquid chromatography electro-spray ionization mass spectrometry (LC-ESI-MS). In all cases, the major mass product obtained corresponded to the protein with a fully cyclized chromophore that had a loss of 20 Da consistent with dehydration and aerial oxidation. The high-resolution HPLC step allowed the separation of the small quantities of unlabeled protein from fluorinated protein. In all cases, using peak-area integration, the expected fluorinated protein was the major product (> 95% labeled) (data not shown). In addition, the labeled and nonlabeled proteins were enzymatically digested using papain and separated, and the resulting peptides were identified using LC-ESI-MS and, in each case, compared. No unlabeled peak was detected in the fluorinated protein digest HPLC chromatogram, suggesting that the ^{19}F-Tyr labeling efficiency was very high (> 95% fluorinated chromo-peptide).

^{19}F NMR of Tyr → Phe Mutants. The ^{19}F assignments were obtained using a combination of site-directed mutagenesis, calculated solvent accessibilities and photo-CIDNP spectroscopy, corroborated by NMR line-fitting analysis and line width measurements. Wild-type GFP has 10 tyrosine residues distributed throughout the structure of the protein (Figure 1). The ^{19}F NMR spectrum of wild-type GFP shows a distinctive spread of ^{19}F chemical shifts from −50 to −60 ppm (Figure 2), as expected for a folded protein. To assign the ^{19}F NMR spectrum, each of the 10 tyrosine residues was mutated, one at a time, to phenylalanine. If the mutation does not significantly perturb the structure, the spectrum of each of the tyrosine mutants should lack one or two signals compared to the spectrum of fully ^{19}F-Tyr labeled wild-type protein. As will be seen, the number of missing signals is determined by the existence of ring-flip conformers.

(34) Kuprov, I.; Goez, M.; Abbott, P. A.; Hore, P. J. *Rev. Sci. Instrum.* **2005**, *76*, 84103.
(35) Goldman, M. *J. Magn. Res.* **2001**, *149*, 160−87.
(36) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.
(37) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.
(38) de Dios, A. C.; Oldfield, E. *J. Am. Chem. Soc.* **1994**, *116*, 7453−4.
(39) Frisch, M. J., et al. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(40) Tormena, C. F.; da Silva, G. V. J. *Chem. Phys. Lett.* **2004**, *398*, 466−470.
(41) Sanders, L. K.; Oldfield, E. *J. Phys. Chem. A* **2001**, *105*, 8098−8104.

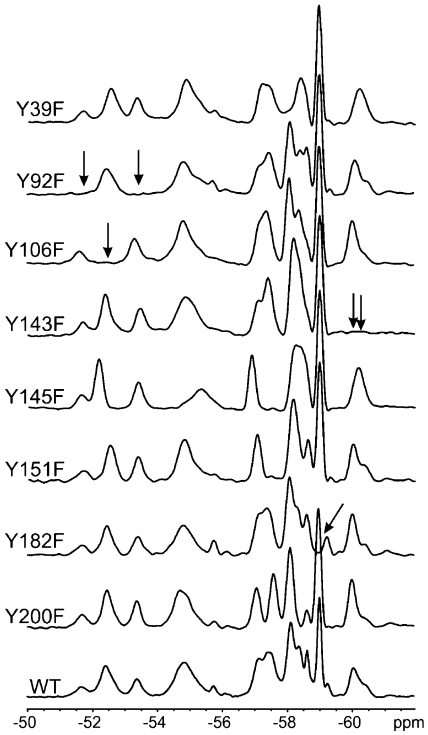

Figure 2. ^{19}F spectra (564 MHz) of wild-type and Tyr → Phe mutant GFPs used for signal assignment. For Y92F, Y106F, Y143F, and Y182F, the arrows indicate the loss of a signal corresponding to that residue. For Y92F and Y143F, two signals are lost as these residues populate two distinct rotameric states (see section on ring-flip conformers).

For Y92F, Y106F, Y143F, and Y182F, the spectra clearly lack one or two signals allowing these fluorotyrosines to be assigned immediately, Figure 2. In the other cases, however, the mutation resulted in noticeable chemical shift changes (e.g., the Y145F mutant, Figure 2), most likely a result of small structural perturbations caused by loss of the polar hydroxyl group when the 3-fluorotyrsine residue is mutated to phenylalanine. Note that because of the short relaxation delay between signal acquisitions, the NMR intensities in Figure 2 are weighted by the ^{19}F relaxation rates and, therefore, vary from residue to residue.

For Y66, which forms part of the chromophore, the mutant protein (Y66F) could not be prepared in sufficient yield to acquire an NMR spectrum, most likely because the Tyr → Phe mutation in the chromophore disrupts the extensive hydrogen bond network and destabilizes the protein. However, Y66 may be assigned on the basis of its relaxation behavior. After the post-translational cyclization, Y66 is incorporated into the chromophore, which is known to be a very rigidly immobilized structure.[4,5] It should therefore be expected that Y66 would have the slowest motion of all 10 tyrosines and, therefore, the broadest NMR line. By far, the broadest signal in the spectrum is located around −55 ppm (Figure 2); we assign this peak to Y66. However, it is clear from the majority of the spectra in Figure 2 that this resonance is asymmetric, suggesting a contribution from another tyrosine residue. As this signal partially vanishes in the spectrum of the Y74F mutant (Figure 3), we attribute it to a superposition of peaks from Y66 and Y74.

Table 2 gives the approximate chemical shifts of the 11 resolvable ^{19}F NMR lines in the wild-type spectrum (two lines for each of Y92 and Y143, one line for Y66/Y74, and one line for each of the other six tyrosine residues).

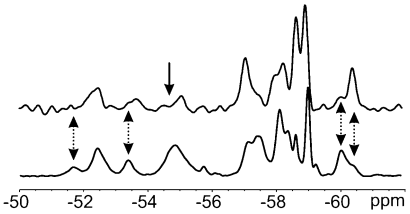

Figure 3. ^{19}F NMR spectra of the Y74F mutant (above) and the wild-type protein (below). The Y74F spectrum was recorded at a different temperature (5 °C) and, consequently, shows a different ring-flip conformer distribution for Y92 and Y143 compared to that in Figure 2 (double-headed dashed arrows). Note the attenuated signal at around −55 ppm in the Y74F spectrum (single-headed solid arrow).

The as yet unassigned NMR signals, which fall in the range −57.1 to −58.4 ppm, are labeled A−D (Table 2); the unassigned residues at this point are Y39, Y145, Y151, and Y200.

Solvent Accessibility Data. The generation of CIDNP in tyrosine residues in proteins involves an electron transfer from an amino acid residue to the excited photosensitizer,[42] a process that only occurs with the necessary kinetics if the residue is exposed to the bulk solvent.[43,44] Solvent accessibility data were therefore calculated using the web-based program GetArea[45] and the crystal structure of GFP[6] (PDB code 1b9c), using either 0.14 or 0.30 nm as the radius of the solvent probe (to model accessibility by water and the flavin, respectively). Both whole-residue accessibilities and atomic accessibilities were calculated. Although there have been attempts in the past to correlate CIDNP intensities with whole-residue accessibilities,[46,23] it is more likely to be the accessibility of the orbital from which the electron is removed that determines the magnitude of the nuclear polarization. Accordingly, the solvent accessibilities of the highest occupied molecular orbital (HOMO) of the 10 tyrosine residues were computed as the weighted sum of atomic accessibilities, the weights corresponding to the HOMO density on each atom.

Only four of the 10 tyrosines (Y39, Y151, Y182, Y200) have significant HOMO accessibilities (21.7, 12.8, 11.8, and 9.9%, respectively); the others are essentially buried (accessibility ≤ 1%). This result suggests that only Y39, Y151, Y182, and Y200 should show significant polarization in a photo-CIDNP experiment.

Photo-CIDNP of the Native State. ^{19}F photo-CIDNP spectra were recorded for wild-type and mutant GFP proteins. GFP absorbs strongly at the usual photo-CIDNP excitation wavelength of 488 nm, so the illumination system was modified to use fiber optics specifically designed to deal with optically dense samples.[31] It was found that the UV lines of the argon ion laser degrade GFP, so the experiment was further adjusted to operate in a single-line, rather than multi-line, laser mode at a wavelength of 514 nm.

Most of the mutants exhibit four emissively polarized overlapping peaks in the ^{19}F photo-CIDNP spectrum (Figure 4). In addition to signals attributable, by their chemical shifts and line widths, to the native states of the proteins, there appears to be a small admixture of unfolded, misfolded, or degraded

(42) Tsentalovich, Y. P.; Lopez, J. J.; Hore, P. J.; Sagdeev, R. Z. *Spectrochim. Acta, Part A* **2002**, *58*, 2043−2050.
(43) Berliner, L. J.; Kaptein, R. *Biochemistry* **1981**, *20*, 799−807.
(44) Kaptein, R. *J. Chem. Soc. D: Chem. Commun.* **1971**, *14*, 732−733.
(45) Fraczkiewicz, R.; Braun, W *J. Comput. Chem.* **1998**, *19*, 319−333.
(46) Feeney, J.; Roberts, G. C. K.; Kaptein, R.; Birdsall, B.; Gronenborn, A. M.; Burgen, A. S. V. *Biochemistry* **1980**, *19*, 2466−2472.

Table 2. ^{19}F Chemical Shifts and Partial Assignments for the 10 ^{19}F-Tyr Residues in Wild-Type Fluorotyrosinated GFP

chemical shift/ppm	−51.7	−52.4	−53.4	−54.9	−57.1	−57.4	−58.1	−58.4	−59.0	−60.1	−60.4
assignment	Y92(I)	Y106	Y92(II)	Y66/74	A	B	C	D	Y182	Y143(I)	Y143(II)

Table 3. Assignment of the ^{19}F Chemical Shifts, the Values of the Generalized Order Parameters, the Residue Accessibilities, and HOMO Accessibility Data for the 10 ^{19}F-Tyr Residues in Wild-Type Fluorotyrosinated GFP[a]

residue	chemical shift[b] ppm	S^2	SASA[c] %	SASA[d] %	HOMO accessibility[e]
Tyr92(I)	−51.65 ± 0.04	0.58 ± 0.03	0.6	0.0	0.0
Tyr106	−52.40 ± 0.08	0.61 ± 0.04	0.0	0.0	0.0
Tyr92(II)	−53.36 ± 0.04	0.51 ± 0.03	0.6	0.0	0.0
Tyr66/74	−54.86 ± 0.10	1.00[f]	0.5/1.9	0.0	0.0
Tyr145	−57.06 ± 0.04	0.49 ± 0.04	6.4	2.6	0.0
Tyr151	**−57.41 ± 0.06**	**0.40 ± 0.03**	**35.3**	**21.8**	**12.8**
Tyr39	**−58.11 ± 0.04**	**0.38 ± 0.03**	**45.0**	**35.1**	**21.7**
Tyr200	**−58.36 ± 0.02**	**0.41 ± 0.06**	**21.2**	**12.8**	**9.9**
Tyr182	**−58.98 ± 0.02**	**0.21 ± 0.02**	**31.4**	**17.3**	**11.8**
Tyr143(I)	−60.08 ± 0.06	0.44 ± 0.04	22.8	10.6	1.0
Tyr143(II)	−60.37 ± 0.06	0.44 ± 0.04	22.8	10.6	1.0

[a] The residues whose signals appear in the photo-CIDNP spectra of the native protein are shown in bold. [b] The signal of residual free 3-fluorotyrosine was used as a chemical shift reference, at −58.60 ppm. [c] Solvent accessible surface area (SASA) calculated using a 0.14 nm probe. [d] SASA calculated using a 0.30 nm probe. [e] Atomic accessibility weighted by the Mulliken atomic populations for the HOMO (see Supporting Information), the latter obtained from a DFT B3LYP 6-311G(d,p) *ab initio* calculation using the GAMESS program.[47] The assignments of Y39 and Y200 are tentative, as described in the text. [f] The Y66 residue is assumed to be rigidly immobilized in the protein core.

GFP (the sharp signal marked by a cross in the spectra of Y200F and Y151F in Figure 4). The exact nature of the admixture is unknown; it could not be separated by HPLC or dialysis, appears to have higher side chain mobility than GFP (hence the sharper NMR signal) and to have much greater solvent accessibility of its fluorotyrosine(s) (accounting for the strong signal in the CIDNP spectra).

The polarized signal at −58.98 ppm (wild-type chemical shift) was previously assigned to Y182 (Figure 2) and is one of the four residues with significant HOMO accessibility (11.8%). Comparison of the various CIDNP spectra in Figure 4 indicates that B−D are polarizable but that A is not. Peak A can therefore be assigned to Y145, the only unassigned residue with negligible HOMO accessibility.

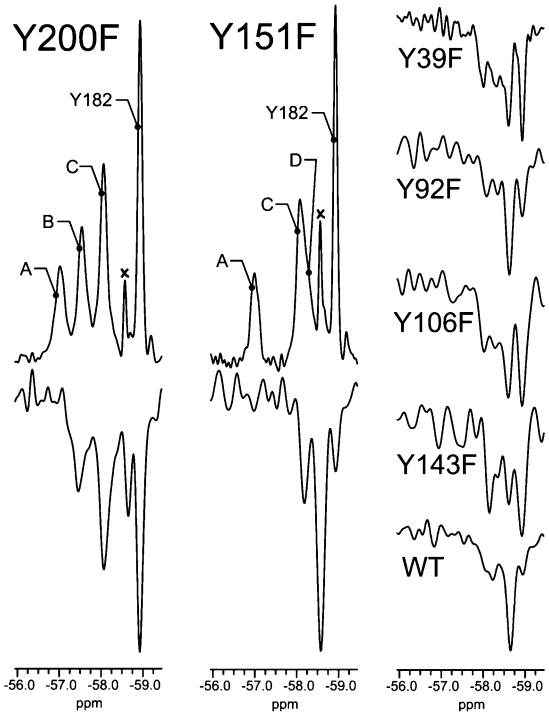

Figure 4. ^{19}F photo-CIDNP spectra of GFP mutants and the wild-type protein. In all cases, the ^{19}F polarizations are emissive. The NMR spectra of Y200F and Y151F are shown above the corresponding photo-CIDNP spectra (on the left of the Figure).

Peak B is absent from the NMR and CIDNP spectra of the Y151F mutant and can thus be assigned to Y151. The two remaining peaks, C and D, cannot unambiguously be assigned to the remaining two residues (Y39 and Y200). Judging by the line shapes of the NMR peaks close to −58.2 ppm in the Y200F and Y151F spectra, it appears that D is absent from the former, suggesting that D should be assigned to Y200 and, by elimination, C to Y39. However, C and D are so strongly overlapping in the Y151F spectrum that this conclusion cannot be definitive.

Table 3 gives the final assignments, the chemical shifts obtained from line-fitting, side chain accessibilities (which will be discussed further below), and the values of the generalized order parameters obtained from a relaxation analysis (see below). The standard deviations reported in Table 3 reflect the range of chemical shifts observed for each residue in the different mutants. Figure 5 shows the results of the spectral line-fitting procedure, together with the assignments. Note that because of the short relaxation delay between signal acquisitions, the NMR intensities in this spectrum, as in Figures 2−4, are weighted by the ^{19}F relaxation rates and, therefore, vary from residue to residue.

^{19}F Relaxation Data. The chromophore of GFP is a rigidly immobilized structure[4,5] which must therefore tumble with the overall molecular correlation time of the protein. It would therefore be reasonable to assume that it has a unit order parameter, $S^2 = 1$. Using this assumption and the *ab initio* ^{19}F shielding tensor parameters given in Table 1, it is possible to extract the overall molecular rotation correlation time τ_c and the order parameters for each residue from the experimental T_2 data at 377 and 564 MHz. The resulting value of τ_c is 14.8 ns, in excellent agreement with the value of 13.5 ns obtained from a HYDRONMR[48] calculation using the X-ray diffraction

(47) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(48) García de la Torre, J.; Huertas, M. L.; Carrasco, B. *J. Magn. Reson.* **2000**, *147*, 138−146.

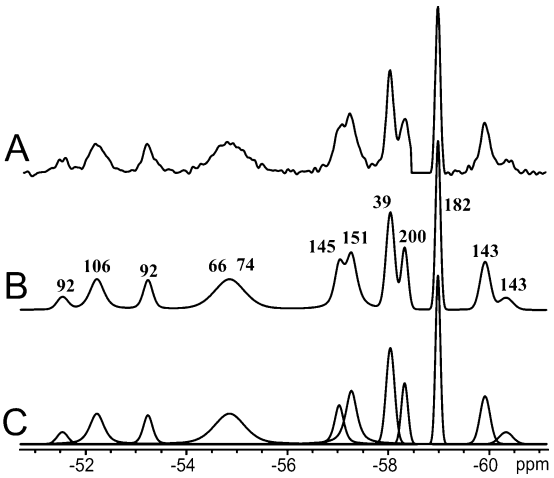

Figure 5. (A) ^{19}F spectrum (564 MHz) of GFP with Lorentzian-to-Gaussian resolution enhancement. (B) Line-fitting. (C) Individual lines in the fitting. The narrow resonance of free 3-fluorotyrosine at −58.7 ppm has been zeroed to facilitate least squares line-fitting. The assignments of Y39 and Y200 are tentative, as described in the text.

structure (PDB code 1B9C). Resulting values of S^2 for the other nine fluorotyrosine residues of GFP are listed in Table 3. The four residues that appear in the photo-CIDNP spectra, namely Y39, Y151, Y182, and Y200, have the narrowest NMR signals and have order parameters in the range 0.21−0.41; the other six residues, which have near-zero HOMO accessibilities, all have larger S^2 ($\geq$0.44). Moreover, the same four fluorotyrosines correspond to the ^{19}F resonances with chemical shifts closest to that of 3-fluorotyrosine (−58.7 ppm). Taken together, these results corroborate the assignments presented above. Residues with high HOMO accessibilities are close to the surface of the GFP molecule, are likely to be less constrained by neighboring side chains, and should therefore exhibit greater internal mobility.

Ring-Flip Conformers. An interesting feature of 3-fluorotyrosinated proteins is the existence of fluorotyrosine ring-flip conformers. Introduction of fluorine into the phenyl moiety of tyrosine breaks its symmetry, and therefore, two ^{19}F signals per fluorotyrosine residue can be expected, because the chemical environment of the two possible fluorine positions will, in general, be different.

Two such signal pairs (Y92 and Y143) are observed in the wild-type GFP spectrum (Figure 2); the rightmost pair, belonging to Y143, displays characteristic behavior: the relative intensity of the two signals varies from mutant to mutant, the signals sometimes coalesce into one apparent signal, and both signals vanish simultaneously when the residue is mutated (as in the Y143F spectrum, Figure 2). The same behavior is also demonstrated by Y92; in this case, however, the signals of the two conformers are substantially further apart. All spectra also contain from two to five low-intensity peaks of variable intensity, which are likely to be weakly populated rotamers of the other fluorotyrosine residues. On the basis that the ring-flipping must be slow compared to the difference in ^{19}F NMR frequencies to observe separate lines from the two conformers, we can put upper limits on the flipping rate of $\sim$160 s^{-1} for Y143 and $\sim$960 s^{-1} for Y92. In the relaxation analysis described above, we have assumed that the exchange is slow enough that its effect on the line widths can be neglected. Any significant contribution from chemical exchange would have made the lines asymmetric and non-Lorentzian.

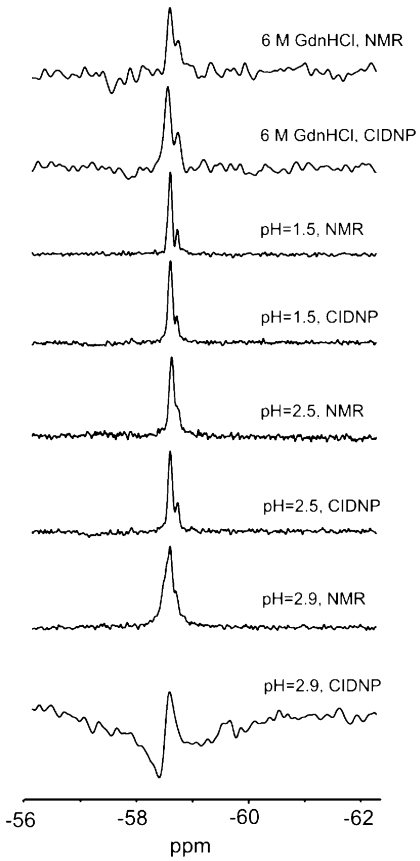

Figure 6. ^{19}F NMR and photo-CIDNP spectra of denatured states of GFP. The NMR spectra were recorded with 256 scans, the CIDNP spectra with 16 scans.

Photo-CIDNP of Denatured GFP. ^{19}F photo-CIDNP spectra were recorded for denatured samples of GFP, in acid solution (pH 1.5, 2.5, and 2.9) and in 6 M GdnDCl at pH 7.2. These conditions were chosen to replicate the initial conditions in ongoing fluorescence kinetic studies.

The ^{19}F NMR spectra of GdnDCl and acid-denatured states are shown in Figure 6 along with the corresponding photo-CIDNP spectra. In all cases, two peaks are clearly resolved; the major peak corresponds to the nine fluorotyrosines that are evidently in chemically equivalent average environments in the denatured state. The minor peak results from the single fluorotyrosine (Y66) that forms part of the chromophore and has a different chemical environment due to the chemical modification of the side chain during chromophore formation. These results suggest that the denatured states (at least in terms of the chemical environments of the tyrosines) are similar in the four different denaturing conditions.

However, the photo-CIDNP spectra (Figure 6) provide more information on the nature of these denatured states and show that the pH 2.9 denatured state is qualitatively different from the other three denatured states.

Extensive experiments on 3-fluorotyrosine (not shown) reveal that for small motional correlation times (less than about 1 ns), the sign of the ^{19}F CIDNP enhancement conforms to Kaptein's rules[44] and the ^{19}F is absorptively polarized. For correlation times greater than about 1 ns at 14.1 T, the polarization changes sign to emission. Briefly, the origin of this effect has been traced

to a cross-correlated relaxation pathway involving the ^{19}F hyperfine tensor and the *g*-tensor of the intermediate fluorotyrosyl radical, which becomes fast enough when $\tau_c > 1$ ns to interfere with the production of CIDNP in the fluorotyrosylflavin radical pair. The sign of the ^{19}F CIDNP polarization therefore provides a probe of the local motional correlation time. Full details of this phase change will be published separately in due course.

This relaxation effect explains the phases of the ^{19}F enhancements shown in Figures 4 and 6. The surface tyrosine residues of the folded GFP molecule have negative CIDNP enhancements, indicating that the effective rotational correlation time is larger than approximately 1 ns. In the unfolded states, however, the enhancement is positive, reflecting much less constrained motion of the polypeptide chain. The partially folded pH 2.9 state features a sharp positive signal (an unfolded, or less restricted ensemble of structures) on top of a broad negative signal (a folded, or more restricted ensemble).

Discussion

^{19}F-fluorotyrosine Labeling Strategy. Many strategies now exist for the incorporation of nonnatural amino acids into proteins.[49−51] Here, we have modified one of these methods[29] to incorporate 3-fluorotyrosine efficiently into GFP for ^{19}F NMR studies. The method uses glyphosate to inhibit the biosynthesis of aromatic amino acids in *E. coli*, the culture media being supplemented with fluorinated tyrosine and nonfluorinated phenylalanine and tryptophan. The same method can be used to introduce fluorinated phenylalanine and tryptophan (F.K. and S.E.J., unpublished results). The efficiency of labeling and the extent of incorporation of labeled tyrosine into GFP was assessed using LC-MS and shown to be greater than 95%. Interestingly, Budisa's group also successfully replaced the tyrosines in the enhanced variant of GFP (eGFP) with 2- and 3-fluorotyrosine derivatives using a Tyr-auxotrophic strain of *E. coli*, work which was published recently.[52] We also tried this method but were less successful. In their study, Budisa and co-workers crystallized the 3-fluorotyrosine-labeled eGFP and found the structure to be indistinguishable from the non-labeled GFP.[52]

Characterization of the Native State of GFP by ^{19}F NMR. The ^{19}F NMR spectrum of ^{19}F-Tyr-labeled wild-type GFP shows multiple peaks corresponding to the 10 tyrosine residues of GFP, and the spectrum shows a distinctive spread of chemical shifts as expected for a natively folded protein. A full assignment was made using a combination of Tyr → Phe mutants, photo-CIDNP spectroscopy, and accessibility calculations, corroborated by relaxation data. Two of the tyrosine residues (Y92 and Y143) were each found to give rise to two ^{19}F peaks corresponding to the two ring-flip conformers, suggesting that the side chains of these residues are in a rigid environment. This is not unexpected, as these side chains are buried in the native structure. Surprisingly, the signals from Y106 and Y145, which are also buried (Table 3), do not appear to be split to such a degree (there is, however, evidence of minor populations of other ring-flip conformers). This may be because the two conformers have fortuitously identical chemical shifts (although this seems unlikely), or because one conformer is highly favored over the other due to specific packing interactions in the native state. Alternatively, although the crystal structure indicates that these side chains are buried and rigidly held, local breathing motions might conceivably allow the conformers to interconvert rapidly. Nevertheless, the splitting of the ^{19}F signal for Y92 and Y143 gives a sensitive probe of the local conformational flexibility of the protein, which can be utilized in future unfolding/folding experiments.

Further experiments were undertaken to characterize the native and denatured states of the protein under a variety of experimental conditions. Four fluorotyrosine residues were found to exhibit nuclear polarization and were assigned using the calculated solvent accessibilities of the molecular orbitals responsible for donating electrons to the photoexcited triplet flavin. Only four of the tyrosine residues have significant HOMO accessibilities (≥ 10%), the others being ≤ 1%. Previous attempts to correlate whole-residue solvent accessibilities (SASA) with observed photo-CIDNP effects have not always been totally successful.[23] In the case of GFP, we also find that the SASAs, whether calculated with a 0.14 nm or a 0.30 nm probe, are qualitatively in disagreement with the observed CIDNP effects. The SASA data in Table 3 suggest that five, rather than four, fluorotyrosine residues should be polarizable in GFP, i.e., Y39, Y151, Y182, Y200, *and* Y143. Although Y143 has very similar overall solvent accessibility to Y200, the HOMO accessibilities differ by an order of magnitude. Inspection of the crystal structure[6] shows that the part of Y143 which protrudes into the solvent is an unreactive $C_\beta H_2-C_\alpha H-NH$ fragment that bears little HOMO electron density, whereas it is the reactive aromatic side chain that is exposed for Y200.

Characterization of the Denatured States of GFP. In general, it is considerably more difficult to obtain detailed structural information on the denatured states of proteins than on native states. Here, we have used ^{19}F photo-CIDNP to probe the environment of the tyrosine side chains of GFP under unfolding conditions (low pH and high concentrations of chemical denaturant). The photo-CIDNP spectra of denatured GFP at pH 1.5, 2.5 and 6 M GdnDCl are all very similar (Figure 6) and exhibit a positive enhancement. In contrast, the pH 2.9 denatured state exhibits both a positive and a negative polarization. As the CIDNP amplitude is related to the correlation time, as summarized above, it would appear that there is some heterogeneity in the size/mobility of the molecule at this pH. This could be due to two structures in equilibrium, one more compact than the other. This is consistent with far-UV circular dichroism studies of the pH 2.9 denatured state which has considerably more secondary structure than the denatured states at pH 1.5 or in 6 M GdnDCl, see Supporting Information. Residual structure, such as this, may be very important in the folding mechanism of GFP, effectively restricting conformational space at a very early stage during the folding process. It may also explain the results of folding measurements on GFP that suggest that folding is considerably faster from the pH 2.9 denatured state than from the pH 1.5 denatured state (data not shown).

In conclusion, we have fully assigned the ^{19}F NMR spectrum of GFP including the identification of two fluorotyrosine ring-

(49) Bayley, H.; Jayasinghe, L. *Mol. Membr. Biol.* **2004**, *21*, 209−220.
(50) Hahn, M. E.; Muir, T. W. *Trends Biochem. Sci.* **2005**, *30*, 26−34.
(51) Petersson, E. J.; Brandt, G. S.; Zacharias, N. M.; Dougherty, D. A.; Lester, H. A. *Methods Enzymol.* **2003**, *360*, 258−273.
(52) Pal, P. P.; Bae, J. H.; Azim, M. K.; Hess, P.; Friedrich, R.; Huber, R.; Moroder, L.; Budisa, N. *Biochemistry* **2005**, *44*, 3663−3672.

flip conformers which provide important probes of the dynamics of the protein. Further, we have shown that photo-CIDNP techniques can be successfully applied to ^{19}F-labeled proteins and provide important structural information on denatured states. A strong correlation has been established between the observed CIDNP effect and the solvent accessibility of the HOMO of the tyrosines. The sign and amplitude of the CIDNP effects have proved particularly useful in identifying differences in the four denatured states of GFP, providing important structural information complementary to other spectroscopic methods.

Acknowledgment. This work was funded in part by the Welton Foundation. F.K. and T.D.C. were funded by MRC and BBSRC Ph.D. studentships, respectively. F.K., T.D.C., and S.E.J. thank Prof. Sir Alan Fersht and Prof. Chris Dobson for access to biophysical equipment and helpful discussions. I.K. thanks the Scatcherd European Foundation and Hill Foundation for a Ph.D. studentship. I.K. and P.J.H. thank the Oxford Supercomputing Centre for a generous allocation of CPU time. P.J.H. acknowledges the financial support of INTAS (Project No. 02-2126), the Royal Society (International Joint Project Grant Program), and the BBSRC.

Supporting Information Available: Expression and purification of fluorotyrosine labeled GFP mutants; far-UV circular dichroism spectra; complete list of authors for ref 39; atomic accessibilities and Mulliken atomic populations for the HOMO of tyrosine. This material is available free of charge via the Internet at http://pubs.acs.org.

JA060618U