

HHS Public Access

Phys Chem Chem Phys. Author manuscript; available in PMC 2022 March 21.

Published in final edited form as: Phys Chem Chem Phys. 2021 March 21; 23(11): 6433–6437. doi:10.1039/d0cp06710e.

Tuning the Electronic Transition Energy of Indole via Substitution: Application to Identify Tryptophan-Based Chromophores That Absorb and Emit Visible Light

Robert J. Micikas#1, **Ismail A. Ahmed**#2, **Arusha Acharyya**1, **Amos B. Smith III**1,* , **Feng Gai**1,*

¹Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United **States**

²Department of Biochemistry and Biophysics Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

These authors contributed equally to this work.

Author manuscript

Abstract

Fluorescent amino acids (FAAs) offer significant advantages over fluorescent proteins in applications where the fluorophore size needs to be limited or minimized. A long-sought goal in biological spectroscopy/microcopy is to develop visible FAAs by modifying the indole ring of tryptophan. Herein, we examine the absorption spectra of a library of 4-substituted indoles and find that the frequency of the absorption maximum correlates linearly with the global electrophilicity index of the substituent. This finding permits us to identify two promising candidates, 4-formyltryptophan (4CHO-Trp) and 4-nitrotryptophan (4 $NO₂-Trp$), both of which can be excited by visible light. Further fluorescence measurements indicate that while 4CHO-indole (and 4CHO-Trp) emits cyan fluorescence with a reasonably large quantum yield (ca. 0.22 in ethanol), $4NO₂$ -indole is essentially non-fluorescent, suggesting that $4CHO-Trp$ $(4NO₂-Trp)$ could be useful as a fluorescence reporter (quencher). In addition, we present a simple method for synthesizing 4CHO-Trp.

Introduction

Because of the high sensitivity, low cost and ease of use, fluorescence spectroscopy/ microscopy is one of the most widely used techniques in biochemistry, biophysics, and biology. Applications range from tracking biomolecules in the cell¹ to probing protein folding and conformational dynamics.^{2,3} While exceedingly useful, this technique requires the presence of at least one fluorophore in the biological system of interest, and in many cases, the chosen fluorophore dictates the type of information extractable from the corresponding fluorescence measurements. Due to this requirement, as well as the fact that

ESI contains the synthesis details of 4X-indoles as well as additional results.

Conflicts of interest

^{*}To whom correspondence should be addressed, smithab@sas.upenn.edu and gai@sas.upenn.edu, Feng Gai, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, PA 19104, USA.

Electronic supplementary information

There are no conflicts to declare.

the majority of biological molecules are intrinsically nonfluorescent or lack a useful fluorescent reporter, most fluorescence studies necessitate the introduction of one or multiple external fluorophores. Therefore, one of the major efforts in chemical biology and biophysics is to develop a diverse set of fluorophores that are useful under different requirements and conditions.

Currently, the available biological fluorophores can be broadly characterized as fluorescent proteins (FPs), organic fluorescent dyes, fluorescent nano-sensors, quantum dots, fluorescent transition metal complexes, and/or fluorescent amino acids (FAAs). Among them, FAAs are particularly suitable for applications where the fluorophore size needs to be minimized, or the chemical structure needs to be similar to the native amino acid it replaces, due to concerns of structural or physicochemical perturbations. While several recent studies $4-9$ have reinforced this notion, the choice of minimally perturbative FAAs is currently limited, thus requiring further development.

Among the naturally occurring FAAs, tryptophan (Trp) is the most fluorescent. However, the utility of tryptophan in biological milieu is limited by several undesirable photophysical properties (e.g., low fluorescence quantum yield (QY), low photostability, and excitable only with UV photons).^{10,11} Importantly, the source of tryptophan's fluorescence, the indole ring, provides a template for engineering FAAs with improved photophysical properties. Therefore, many past studies have focused on developing Trp-based FAAs that can offer new and/or improved spectroscopic utility.^{12–19} Despite the past efforts only recently did Hilaire et al.²⁰ discover that 4-cyanotryptophan (4CN-Trp) exhibits significantly different photophysical properties than Trp, making it a promising FAA for biological spectroscopic and imaging applications.21 However, while 4CN-Trp emits in the blue region of the visible spectrum with a large \overline{OY} (>0.8), its one-photon excitation wavelength is still in the UV region (up to ca. 360 nm). Therefore, in this study we seek to identify indole-based chromophores that can absorb and emit visible light.

A previous study has shown that amongst all CN-substituted indoles, 4CN-indole produces the most red-shifted absorption and emission spectra.²² A similar trend is also observed for azaindoles²³ and formyl indoles.²⁴ These findings suggest the 4-position of the indole ring is special in that its local electronic density may have a greater effect on the electronic transition dipole moment(s) of the indole chromophore. While validation of this notion needs further quantum mechanical calculations, we hypothesize that the absorption and emission wavelengths of 4-substituted indoles could be tuned by changing the electronic properties of the substituent. If confirmed, such tunability would help discover Trp derivatives that afford novel spectroscopic properties and utilities. In an effort to test our hypothesis, herein we first determine the absorption and emission spectra of a diverse library of 4X-indoles and then evaluate how their electronic transition energy (i.e., that corresponding to the maximum absorption wavelength) correlates with various scales characterizing the electronic effect of X. Based on the results obtained with 4X-indoles, we further identify two Trp-based unnatural amino acids, 4-formyltryptophan (4CHO-Trp) and 4-nitrotryptophan ($4NO₂-Trp$), that can be excited by visible light and hence could have useful biospectroscopic applications. In addition, we describe a facile synthetic route to construct 4CHO-Trp.

Experimental details

Materials

4-Bromoindole, indole-4-carboxaldehyde, indole-4-carboxylic acid, 4-cyanoindole, 4 fluoroindole, 4-methylindole, 4-nitroindole, and L-4-bromo-tryptophan were purchased from Chem-Impex Int'l Inc. 4-Aminoindole was purchased from Ark Pharm, Inc. All commercial reagents were used as received. All other $4X$ -indoles as well as $L\text{-}N^a$ -acetyl-4formyltryptophan were synthesized using the protocols described in detail in the ESI.

Absorption measurements

UV-vis absorption spectra were collected on a Jasco V-650 UV-vis spectrophotometer using a quartz cuvette of 1.0 cm path length at room temperature. The concentrations of all samples were kept at ca. 25 μM.

Fluorescence measurements and quantum yield (QY) determination

Steady state fluorescence measurements were carried out on a Jobin Yvon Horiba Fluorolog 3.10 spectrofluorometer at 20 °C using a quartz cuvette of 1.0 cm path length. All the measurements were done using 1 nm excitation/emission slits, an integration time of 1.0 nm/s, and an excitation wavelength of either 270 nm (for 4X-indoles whose absorption spectra are similar to that of indole) or 350 nm (for 4X-indoles whose absorption spectra are similar to, or red-shifted from, that of 4CN-indole). The sample concentrations were ca. 5 μM.

For each sample, QY was determined at 20 °C using either quinine sulfate in 0.5 M sulfuric acid (when λ_{ex} = 350 nm was used) or 4CN-indole in water (when λ_{ex} = 270 nm was used) as the reference and the following relationship:

$$
QY_S = QY_R \frac{I_S}{I_R} \frac{A_R}{A_S} \frac{n_S^2}{n_R^2}
$$

where I is the integrated fluorescence intensity, A is the optical density of the fluorophore at $\lambda_{\rm ex}$, *n* is the refractive index of the solvent, and the subscripts *S* and *R* represent the sample and reference, respectively. In addition, QY_R of 0.543 (0.85) was used for quinine sulfate (4CN-indole).

Results and discussion

The 4X-indole library was assembled to include both electron-donating and -withdrawing substituents varying in their degree and manner (e.g., induction and π conjugation) of electronic perturbation. Specifically, the X substituents include -F, -Cl, -Br, -I, -OH, -CN (cyano), -NC, -N₃, -NO₂ (nitro), -NH₂, -NHCH₃ (-NHMe), -N(CH₃)₂ (-NMe₂), -N(CH₃)₃⁺ $(-NMe₃⁺)$, -NHC(O)Me $(-N-Ac)$, -NHC(O)CF₃ $(-N-TFA)$, -CH₃ $(-Me)$, -CHO (formyl), - $C(O)OH (-CO₂H)$, $-C(O)OCH₃ (-CO₂Me)$, $-OC(O)Me (-O₄Ca)$, and $-OC(O)CF₃ (-O₄Te)$. Derivatives not commercially available were obtained through chemical synthesis (Scheme S1 in the ESI). Many of the 4X-indoles examined have diminishingly low solubility in water

thus necessitating the use of ethanol as the solvent. As expected, the absorption spectra of 4X-indoles depend on the nature of the substituent (Fig. S1, ESI). As indicated (Fig. 1 and Table 1), a general trend that emerged from this data is that an electron-withdrawing group (EWG) shifts the wavelength of maximum absorbance (λ_{ab}^{max}) to the red relative to that of indole, whereas an electron-donating group (EDG) does the opposite. In addition, a strong EWG (e.g., $NO₂$) tends to increase the overall width of the absorption spectrum. To provide a more quantitative assessment of the effect of the substituent, X, on the absorption spectra of 4X-indoles, we sought to identify possible correlations between $\lambda_{ab}^{\text{max}}$ and the electronattracting propensity of X. A widely used empirical for quantifying the electronwithdrawing or donating nature of a substituent is the Hammett parameter (δ).This parameter was originally formulated to quantify the effect of a substituent (relative to a hydrogen atom) on various physicochemical properties (e.g., ionization constant and reaction rate) of an aromatic molecule in the ground state, $2⁵$ but has since found similar utility in assessing the excited-state processes and photochemical behaviors of substituted aromatic molecules.^{26,27} As shown (Fig. S3, ESI), the frequencies of maximum absorbance (v_{ab}^{max}) of 4X-indoles exhibits only a weak correlation with either the δ_{meta} ($R^2 = 0.20$) or δ_{para} (R^2 = 0.30) values²⁸ of X, indicating that δ is not a good predictor of the effect of a substituent on v_{ab}^{max} or λ_{ab}^{max} . However, the observed trend is consistent with our initial expectation that a strong EWG (e.g., NO₂ and CHO) does lead to an increase in λ_{ab}^{max} .

Since δ contains contributions from a field (or induction) component (through parameter F) and a resonance component (through parameter R), $28,29$ we also analyzed the dependence of v_{ab}^{max} on F and R. As indicated (Fig. S4 in the ESI), neither F ($R^2 = 0.10$) nor R ($R^2 = 0.34$) significantly improves the strength of the correlation.

Another more recently proposed scale for measuring the propensity of a molecule (or group) to acquire additional electronic density from the environment is the global electrophilicity index (ω),^{30–32} which is defined as $\mu^2/2\eta$, where μ and η are the chemical potential and hardness of the molecule, respectively. Chemical hardness (η) is defined as the change in chemical potential with the change in the number of electrons, hence describing a molecule's resistance towards changes in electron density and distribution. Considering that indole's electronic transition(s) to the lowest excited state(s) involves a partial electron transfer from the pyrrole to the benzyl portion of the ring^{33,34} and that ω measures the energy of stabilization of a species due to acquirement of additional electronic charge, we hypothesized that the stability of the excited state(s) of 4X-indole would correlate well with the ω value of X. As shown (Fig. 2), $v_{\rm sh}^{\rm max}$ indeed exhibits a strong positive correlation with ω, which was obtained from the computational study of Domingo *et al.* using monosubstituted ethylene as a reduced model for more complex aromatic systems, 35 as judged by the R^2 value (0.84).

The observed correlation suggests that the ability of a substituent, X, to stabilize a 4Xindole's excited state is a function of both X's affinity for the indole's π electron density and the polarizability of X's own electron density as dictated by η . Moreover, this result

validates the ability of ω to predict the electronic transition energy of 4X-indoles. However, further studies are necessary to determine whether ω can also be used in a predictive fashion for other aromatic fluorophores.

As observed for the absorption spectra, the fluorescence spectra of these 4X-indoles in ethanol are all broad and featureless (Fig. S2, ESI). In addition, most display a large (>60 nm) Stokes shift (Fig. 1 and Table 1). Furthermore, like that observed for v_{ab}^{max} , the frequency of maximum emission intensity (v_{em}^{max}) only correlates well with ω (Fig. 2 and Fig. S3 and S4, ESI).

For any molecule to be useful as a fluorescence reporter, it must have a reasonably large QY. Therefore, we also characterized the QYs of these 4X-indoles. As indicated (Table 1), most of the 4X-indoles examined have a low QY $(0.1), making them less useful as stand-alone$ fluorophores in biological studies. However, $4CO₂H$ -indole, $4CO₂Me$ -indole, and $4CHO$ indole all have a relatively high QY (>0.2).

The experimental results presented above provide us with the needed information to identify new 4X-indole-based FAAs, especially those whose fluorescence can be excited with visible light. In this regard, 4-formyltryptophan (4CHO-Trp) is the most attractive fluorophore as 4 formylindole has a fluorescence QY of 0.22 in ethanol and an absorption spectrum extending beyond 390 nm. A recent study on formylindoes by Zhang and coworkers²⁴ reached a similar conclusion. While 4-nitroindole (4NO₂-indole) has the most red-shifted absorption spectrum and is a true visible chromophore, the very low fluorescence QY (ca. 0.001) makes 4-nitrotryptophan (4NO₂-Trp) unsuitable for fluorescence measurements. However, it could be used as a fluorescence resonance energy transfer (FRET) quencher of biological fluorophores whose absorption spectra are in the wavelength range of 380–450 nm, such as the nuclear stain DAPI and enhanced green fluorescent protein (EGFP).

The fluorescence QY of Trp is significantly smaller than that of indole. Therefore, to confirm that 4CHO-Trp possesses similar fluorescent properties as 4CHO-indole, we devised a short and efficient route (Scheme 1) to synthesize N-acetyl-4-formyl-tryptophan-OMe (**1**) and also determined its QY in ethanol. Briefly, commercially available 4-bromo-Ltryptophan (**2**) was first esterified and acetylated to deliver N-acetyl-4-bromo-tryptophan-OMe (**3**) (58% yield). The aldehyde functionality was next installed via a convenient two step protocol wherein Suzuki cross-coupling of **3** with vinyl trifluoroborate (**5**) ³⁶ afforded Nacetyl-4-vinyl-tryptophan-OMe (**4**) which underwent Lemieux-Johnson oxidation to yield the desired 4CHO-Trp (**1**) (48% over 2 steps).

The results of Hilaire *et al.* indicated that the absorption and emission spectra of 4CN-Trp are further red-shifted from those of 4CN-indole.20 As shown (Fig. 3), 4CHO-Trp (**1**) follows the same trend as its λ_{ab}^{max} and λ_{em}^{max} values are increased by 13 and 23 nm, respectively, from those of 4CHO-indole in ethanol. As a result, the absorption spectrum of 4CHO-Trp (**1**) in ethanol extends beyond 410 nm, making the fluorescence excitable by blue light sources. In addition, the fluorescence spectrum of 4CHO-Trp (**1**), which peaks at ca. 500 nm in ethanol, makes this Trp derivative a cyan fluorophore (Fig. 3). Finally, it is worth

noting that Ladner *et al.*³⁷ have employed various photoreactions of the indole ring of Trp38,39 to generate new Trp-based fluorophores and, interestingly, they found that one of the photoproducts, 4-imido-tryptophan, which was formed by the initially produced photoproduct 4-formyltryptophan via reaction between its aldehyde and amine groups, exhibited yellow color and had a $\lambda_{\text{em}}^{\text{max}}$ of 535 nm.

Conclusions

In summary, we surveyed the absorption and emission properties of a large array of 4 substituted indoles (4X-indoles), aiming to identify new biologically useful fluorescent amino acids. Our results clearly demonstrate that the excitation energy (i.e., the frequency of the maximum of the absorption spectrum) of a 4X-indole depends on the global electrophilicity index (ω) of X. This correlation provides a convenient means to identify Trp derivatives whose absorption spectra are red-shifted from that of Trp and, especially, whose fluorescence can be excited by visible light. Moreover, we believe that our experimental data also provide important benchmark information for studies aiming to understand how a substitution on the indole ring affects its ground and excited state electronic structures via both theoretical and experimental approaches. $40-42$ Among the 4X-indoles studied, we find that 4CHO-indole and $4NO₂$ -indole meet this requirement. However, $4NO₂$ -indole is essentially non-fluorescent, making it only useful in FRET-based applications (i.e., serving as a FRET quencher).43 On the other hand, 4CHO-indole and its amino acid form, 4CHO-Trp, have a reasonably large fluorescence QY in ethanol (ca. 0.22), suggesting that this chromophore could be useful as a biological fluorescence reporter. However, to validate this point, further studies are needed to fully characterize the chemical and photophysical stabilities, the fluorescence decay kinetics, and the factors affecting the fluorescent properties of this FAA in a peptide or protein environment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge the Research Foundation of the University of Pennsylvania for support. R.J.M. is supported by NIAID (8PO1-A1-150471) and I.A.A. is supported by a NINDS D-SPAN Fellowship (F99 NS108544-01).

REFERENCES

- 1. Chudakov DM, Matz MV, Lukyanov S and Lukyanov KA, Physiol. Rev, 2010, 90, 1103–1163. [PubMed: 20664080]
- 2. Michalet X, Weiss S and Jäger M, Chem. Rev, 2006, 106, 1785–1813. [PubMed: 16683755]
- 3. Serrano AL, Waegele MM and Gai F, Protein Sci, 2012, 21, 157–170. [PubMed: 22109973]
- 4. Rubini M, Lepthien S, Golbik R and Budisa N, Biochimica et Biophysica Acta 2006, 1764, 1147– 1158. [PubMed: 16782415]
- 5. Hammond JW, Blasius TL, Soppina V, Cai D and Verhey KJ, J. Cell Biol 2010, 189, 1013–1025. [PubMed: 20530208]
- 6. Stadler C, Rexhepaj E, Singan VR, Murphy RF, Pepperkok R, Uhlen M, Simpson JC and Lundberg E, Nat. Methods, 2013, 10, 315–323. [PubMed: 23435261]

- 7. Luo J, Uprety R, Naro Y, Chou C, Nguyen DP, Chin J and Dieters A, J. Am. Chem. Soc, 2014, 136, 15551–15558. [PubMed: 25341086]
- 8. Stout RF Jr., Snapp EL and Spray DC, J. Biol. Chem, 2015, 290, 23497–23514. [PubMed: 26265468]
- 9. Hostetler ZM, Ferrie JJ, Bornstein MR, Sungwienwong I, Petersson EJ and Kohli RM, ACS Chem. Biol, 2018, 13, 2855–2861. [PubMed: 30216041]
- 10. Chen RF, Anal. Lett, 1967, 1, 35–42.
- 11. Roy R, Hohng S and Ha T, Nat. Methods, 2008, 5, 507–516. [PubMed: 18511918]
- 12. Negrerie M, Bellefeuille SM, Whitham S, Petrich JW and Thornburg RW, J. Am. Chem. Soc, 1990, 112, 7419–7421.
- 13. Ross JBA, Szabo AG and Hogue CWV, Methods Enzymol, 1997, 278, 151–190. [PubMed: 9170313]
- 14. Budisa N and Pal PP, Biol. Chem, 2004, 385, 893–904. [PubMed: 15551863]
- 15. Talukder P, Chen S, Arce PM and Hecht SM, Org. Lett, 2014, 16, 556–559. [PubMed: 24392870]
- 16. Talukder P, Chen S, Roy B, Yakovchuk P, Spiering MM, Alam MP, Madathil MM, Bhattacharya C, Benkovic SJ and Hecht SM, Biochemistry, 2015, 54, 7457–7469. [PubMed: 26618501]
- 17. Shen JY, Chao WC, Liu C, Pan HA, Yang HC, Chen CL, Lan YK, Lin LJ, Wang JS, Lu JF, Chou SCW, Tang KC and Chou PT, Nat. Commun, 2013, 4, 2611. [PubMed: 24177573]
- 18. Chalyavi F, Gilmartin PH, Schmitz AJ, Fennie MW and Tucker MJ, Angew. Chem. Int. Ed. Engl 2018, 57, 7528–7532. [PubMed: 29710418]
- 19. Zhang K, Ahmed IA, Kratochvil HT, DeGrado WF, Gai F and Jo H, Chem. Commun, 2019, 55, 5095–5098.
- 20. Hilaire MR, Ahmed IA, Lin C, Jo H. l., DeGrado WF and Gai F, Proc. Natl. Acad. Sci, 2017, 114, 6005–6009. [PubMed: 28533371]
- 21. Acharyya A, Ahmed IA and Gai F, Methods Enzymol, 2020, 639, 191–215. [PubMed: 32475401]
- 22. Hilaire MR, Mukherjee D, Troxler T and Gai F, Chem. Phys. Lett, 2017, 685, 133–138. [PubMed: 29225366]
- 23. Lepthien S, Hoesl MG and Merkel L, Proc. Natl. Acad. Sci, 2008, 105, 16095–16100. [PubMed: 18854410]
- 24. You M, Fan H, Wang Y and Zhang W, Chem. Phys, 2019, 526, 110438.
- 25. Hammett LP, J. Am. Chem. Soc, 1937, 59, 96–103.
- 26. Papper V and Likhtenshtein GI, J. Photochem. Photobiol. A, 2001, 140, 39–52.
- 27. Sadlej-Sosnowska N and Kijak M, Struct. Chem, 2012, 23, 359–365.
- 28. Hansch C, Leo A and Taft RW, Chem. Rev, 1991, 91, 165–195.
- 29. Hansen LD and Hepler LG, Can. J. Chem, 1972, 50, 1030–1035.
- 30. Maynard AT, Huang M, Rice WG and Covell DG, Proc. Natl. Acad. Sci, 1998, 95, 11578–11583. [PubMed: 9751708]
- 31. Parr RG, Szentpaly L. v. and Liu S, J. Am. Chem. Soc, 1999, 121, 1922–1924.
- 32. Chattaraj PK, Sarkar U and Roy DR, Chem. Rev, 2006, 106, 2065–2091. [PubMed: 16771443]
- 33. Andrews LJ and Forster LS, Biochemistry, 1972, 11, 1875–1879. [PubMed: 5025629]
- 34. Levy RM, Westbrook JD, Kitchen DB and Krogh-Jespersen K, J. Phys. Chem, 1991, 95, 6756– 6758.
- 35. Domingo LR, Perez P and Contreras R, J. Org. Chem, 2003, 68, 6060–6062. [PubMed: 12868951]
- 36. Molander GA and Rivero MR, Org. Lett, 2002, 4, 107–109. [PubMed: 11772102]
- 37. Ladner CL, Tran K, Le M, Turner RJ and Edwards RA, Photochem. Photobiol, 2014, 90, 1027– 1033. [PubMed: 24738707]
- 38. Voropei AV, Chernitskii YA, Konev SV, Krivitskii AP, Pinchuk SV and Shchukanova NA, Biophysics, 1992, 37, 743–745.
- 39. Edwards RA, Jickling G and Turner RJ, Photochem. Photobiol, 2002, 75, 362–368. [PubMed: 12003125]

- 41. Abou-Hatab S and Matsika S, J. Phys. Chem. B, 2019, 123, 7424–7435. [PubMed: 31373821]
- 42. Henrichs C, Reineke M, M-L. Hebestreit and M. Schmitt, J. Mol. Struct, 2021, 1223, 129241.
- 43. Ahmed IA, Rodgers JM, Eng C, Troxler T and Gai F, Phys. Chem. Chem. Phys, 2019, 21, 12843– 12849. [PubMed: 31179453]

Dependence of the frequency of the maximum absorbance/emission of 4X-indoles on the global electrophilicity index (ω) .

Micikas et al. Page 11

Fig. 3.

Normalized absorption (blue) and fluorescence (red) spectra of 4CHO-Trp (**1**) in ethanol. The excitation wavelength for the fluorescence measurement was 350 nm. As shown in the inset, 4CHO-Trp (**1**) is a cyan FAA.

Synthesis of 4CHO-Trp (**1**).

Table 1.

Spectroscopic properties and physical constants of the 4X-indoles.

