Computational and Experimental Exploration of Stryl Indole Analogs
Kenneth O'Dell

I. Introduction

The usage of fluorescence spectroscopy, and luminescence in general, has extended to many biochemical applications in protein research, contributing to insight into protein structure, function, and interaction. The possible approaches that can be undertaken with fluorescent probes are open to flexibility, e.g., oxidation of the protein of interest, creating new functional groups for subsequent fluorescent labeling, or, for the investigation of peptide-binding proteins, the substrate of interest may also be subject to fluorescent tagging (Roeser, 2010; Kitamatsu, 2010).

Another fascinating method considers the interactions between an excited molecule and neighboring molecules, which can also be used to measure the nanometer scale distances moved during conformational changes. This phenomenon, known as quenching, is yet another helpful tool employed in protein investigations. For example, the shifting of a helix that conceals an active site can be elucidated with the use of energy-donors (excited species) and energy-acceptors (called 'quenchers') bound to the protein, which may consist of the natural, aromatic amino acids of the protein (such as tryptophan) and an unnatural-acceptor such as IEDANS, FITC, or Dansyl (Atkins, 849-853). However, there are a variety of acceptors and donors. The donor species may very well be a non-peptide species like naphthalene, whereas acceptors species may also be inorganic ions in solution, like iodide (Atkins, 851). Such a condition can allow for the distinction between surface residues and interior residues when surface residues are quenched by ions in solution. Thus, by measuring the relative differences between a species' fluorescence quantum yield in the absence and the presence of a quencher, we can determine the changes in distances that have occurred during a conformational change or the relative positions of certain residues.

The two common natural amino acids used for fluorescent studies are tyrosine and tryptophan. Synthesized using the Heck coupling reaction, 3,5-tyrosine analogs consisting of styryl functional groups displayed unique fluorescent properties, and these unnatural amino acids (UAAs) can be promising fluorescent labels for future proteomic applications (Cheruku, 2015).

Tryptophan derivatives synthesized by C-C coupling reactions (i.e. extending the conjugated system) have not been explored to the same extent as tyrosine derivatives vis-à-vis photochemical properties (Li,1995). Tryptophan itself is an interesting amino acid chosen for this investigation for several reasons. The chemistry of tryptophan stems from its indolic side-chain, which contributes to its unique biochemical roles, e.g., as a structural anchor for membrane proteins due to its large size and hydrophobicity (de Jesus, 2013). The amino acid also plays an important role as a precursor to essential biomolecules, such as serotonin, melatonin, NAD, and NADP (Sainio, Pulkki, and Young,

1996) Tryptophan, owing to its large conjugated system that allows for some fluorescence, has been employed in quenching studies, and the natural tryptophan residues may be used to measure close-range conformational changes, using a technique called tryptophan-induced quenching (TrIQ) (Mansoor, DeWitt, and Farrens, 2010).

Exploring the luminescent properties of a molecule invariably involves consideration of its excitation energies. Following a similar approach to Cheruku, unnatural tryptophan derivatives were originally considered for this study.

These unnatural tryptophan derivatives (UTD) may, for example, be synthesized from C-C coupling reactions, allowing for the extension of the π-conjugated system. These additions may consist of styrene and functional derivatives of styrene such as vinylbenzoic acid or vinylaniline. The addition of functional substituents may allow for microenvironmental conditions (such as solvent effects) to affect the molecule's fluorescence similarly to that observed for the tyrosine derivatives in Cheruku's study. However, certain limitations of economy (price of reagents) and software (atom-count limits) have altered the approach to this study.

The salient feature of tryptophan, discussed earlier, lies in its indolic side-chain. By making the approximation that the backbone structure contributes very little to the actual luminescent properties of the modified side-chain, the backbone structure can be omitted and we can focus on the functional fluorophore (that is, indole) and bypass the atom-count limitation (which is 30 atoms for *Spartan Student*). Thus, this study aims at exploring the potential of various derivatives of tryptophan through a comparative, qualitative analysis of the electronic structure of various analogs of its side-chain, indole, alongside stilbene analogs, which will represent the tyrosine derivatives in Cheruku's study. Methylation of what would have been the backbone sites (C3 for indole, C4 for phenol) was considered to allow for the inclusion of the effects an alkyl substituent may have on the fluorescence of the compounds; however, the additional four atoms contributed by the methyl group precluded several of the indole and stilbene analogs from *Spartan's* calculations.

Henceforth, styrl-*mono*-indole analogs with C-H at C3 and stilbene analogs with C-H at C4, both sites representing where the backbone would have been, will be the species of focus throughout this study.

The indole derivatives were chosen based off of the most common commercially available brominated derivatives, i.e. pre-functionalized, which are the 4-, 5-, 6-, and 7-substituted derivatives. Studies have focused on Palladium-catalyzed C-C coupling reactions by C-H functionalization of indole, a synthetic method that would prove useful should pre-functionalized indole not be used. Luckily, the availability of brominated indole simplifies this decision. Among other things to consider are reactions that allow for N-substituted (Heller, 2012), boronated (Bartolucci, 2011), mono- and dibrominated derivatives (Bittner, 2007), and phenyl and methoxy derivatives (Janczuk, 2002). As a sidenote, it is interesting to note that additional variations of the brominated derivatives

have not been found to exist freely in nature and instead result from post-translational modification of proteins in marine sponges, where they are involved in immune defense mechanism (Bittner, 2007; Mollica, 2012).

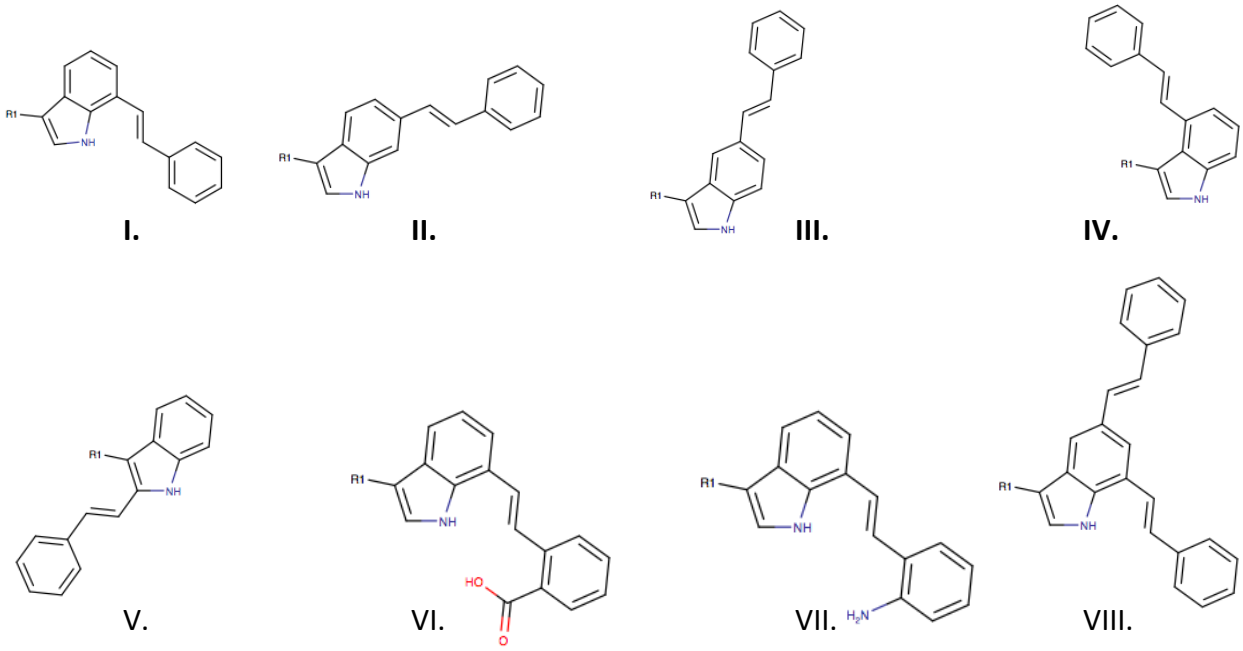


Fig. 1 - Examples of derivatives of tryptophan that may be possibly synthesized from pre-functionalized indoles. I. 7-styrylindole II. 6-styrylindole III. 5-styrylindole IV. 4-styrylindole V. 2-styrylindole VI. 7-(-2-vinylbenzoic acid)-indole VII. 7-(2-vinylaniline)-indole VIII. 5,7-distyrylindole. Not all of these derivatives are examined in this study. Drawn with ChemAxon's Marvin Sketch.

The actual indole analogs themselves may lend to applications in biochemical assays. A structurally similar stilbenyl boronic acid analog was used as a cofactor for an antibody that could then fluoresce an intense blue, which would subsequently be quenched by Hg^{2+} ions should they be present (Masayuki, 2005). In this case, a compound similar to those examined in this study 'chemically programmed' a protein so that it would fluoresce and specifically be quenched by mercury ions, leading to the creation of a novel biosensor for mercury ions. This, again, exemplifies the many possible applications for fluorescent compounds, whether they be unnatural amino acids, indole analogs, or stilbene analogs.

Quantum Mechanical (QM) calculations are an excellent way of studying molecules that may be difficult to synthesize, expensive, or simply theoretical. The downside of QM calculations lies in the savviness, per se, of the practitioner and the price and ac-

cess to the software that can run the desired model. QM calculations performed *ab initio,* despite advances in computing, are still approximations. Different models utilize different approximations and basis set functions to obtain a balance between the computational costs and the level of accuracy required for a given system, the molecule(s) of interest.

Various options exist at this point in proceeding with the study. For example, rather than an immediate experimental approach to synthesizing and analyzing the excited-state properties of theoretical indole analogs, computational analysis and theoretical absorption and emission spectra is an alternative method and has, in fact, been used for nucleotide derivatives (Gustavssonm, 2006).

The computational analysis of excited states would entail a time-dependent model, such as TD-DFT (Time-dependent density functional theory), which would allow for ground state and excited state calculations, enabling us to determine excitation energies and absorption spectra and make comparisons between empirical observations and computational models. One study did exactly this, using UV-Photoelectron Spectroscopy to measure HOMO/LUMO energies and correlate them with computational models (Chrostowska, et al). UV-Photoelectron spectroscopy, which utilizes photons of known energy, is an ideal technique to pair with quantum calculations since orbital energies can be empirically determined via the kinetic energies of electrons ejected from the molecule orbitals. This allows for an ideal model to be chosen for the given molecule(s) and for a confirmation as to a model's accuracy. Chrostowska et al also utilized a time-dependent model (CAM-B3LYP/6-311++G(d,p)) which is unnamable in Spartan Student.

TD-DFT calculations can be rather involved, requiring expensive software, and be difficult to use for the non-specialist (Jacquemin, Mennucci, and Adamo, 2011). Therefore, a second part of this study will involve synthesis of one indole analog (chosen after a computational analysis of a set of analogs) and an effort to characterize its fluorescent properties obtain a UV-Vis spectra. Thus, the computational part will help to steer the exploration and offer some structural details of the analogs of interest.

The program that will be used in this study, Spartan Student Edition, contains several available models (Hartree-Fock, B3LYP, EDF2, MP2) to use with default basis sets and parameters. The Hartree-Fock (HF) models stem from the Schrödinger equation, treating electrons as separate, individual particles, and molecular orbitals are approximated by summing the relevant atomic orbitals (*Spartan '10 Tutorial and User Guide*). Differences within the Hartree-Fock models depend on the basis sets, which are sets of basis functions that describe the atomic orbitals centered at each atom. Thus, by varying the number and type of atomic orbitals (for example, including d- and f-orbitals, though unnecessary for this study), we arrive at different models that may perform better for a given system than other models. For example, the basis sets available with Spartan Student are 6-31G*, 6-31G** and 3-21G. Between these three basis sets, the 6-31G* basis set offers sufficient flexibility that is suitable for the needs of this study (particularly,

qualitatively examining the stabilities and electronic structures of the chosen indole derivatives and comparison with phenolic derivatives). The 6-31G* basis set uses 6 functions to describe each inner-shell orbital, sets of 3 and 1 to describe valence-shell orbitals, and provides polarization functions for non-hydrogen atoms that aid in describing the distribution of electrons in-between atoms; the 6-31G** basis set places polarization functions on hydrogen atoms *(Spartan Student Edition v. 6 Topics: Theoretical Models)*. Larger basis sets and some hybrid models would obviously add to the computational cost, which is an important consideration for this study given the 30-atom limit.

Models besides the HF models account for electron-electron coupling of motions (electron correlations) and the energy differences between experimental energies and HF energies (*Spartan Student Edition v. 6 Topics: Theoretical Models).* These models include the Møller-Plesset model (MP), which adds an energy correction to the HF approximation, and the density functional model (EDF), which adds a correction term that is dependent on the electron density and the gradient of this density. It should be noted that additions of these corrections for electron correlations expand the functions, significantly increasing computational cost—especially for the Møller-Plesset model. Three models will be used (EDF, MP, and B3LYP) and the relative differences of energies across models will help determine the ideal model to use.

Following the quantum mechanical assessment of the analogs, an attempt of synthesizing the analog of choice (in this case, 5-*mono*-styryl-indole) will be conducted using the procedure outlined in a study by Cui et al, except with the replacement of phenyl bromide with the heteroaromatic compound, indole. One study had reported the successful synthesis of (E)-5-styryl-indole, using a Wittig reaction with benzyl bromide and indole-5-carbaldehyde; the product was then subsequently biologically evaluated and shown to serve as probes for Beta-amyloid plaques (Yang, Jia, and Liu 2012).

For the purposes of this study, the Heck reaction will be evaluated as a method to synthesize the indole analog and, consequently, tryptophan analogs.

A phosphine free catalyst (Pd(quinoline-8-carboxylate)$_2$) will be used for both its low-price and the promising high-yield outcomes. 5-bromoindole will be used as the aryl bromide instead of phenyl bromide. 5-*mono*-styrylindole was chosen as the target molecule for the synthesis due to the low-cost and availability of 5-bromoindole. Also, should this procedure be extrapolated to the synthesis of an UAA, there would be less steric hindrance should the styryl group be placed on the C-5 position.

In consideration of the sensitivity of the the reaction to air, precautions will be taken to perform the reaction under an inert atmosphere by setting up a Schlenk line using N_2 gas instead of Ar (again, for cost considerations), drying and storing the catalyst under a vacuum desiccator, storing the pre-measured reagent mix in a vacuum desiccator, and preparing the reaction flasks inside a homemade glove box of positive pressure with N_2. Two runs will be performed simultaneously. The reaction mixture will be diluted with ether, washed with water, and extracted, allowing the solvent to dry.

A TLC will be performed to check for the presence of the product. Should the synthesis be successful, a UV-vis absorption spectrum will be obtained (following purification, if necessary).

II. **Methods and Experimental Observations**

Quantum Calculations

Spartan Student, a quantum chemistry software package was used to compute the energies of indole and its stilbene analogs. Various *ab initio* methods were tested (EDFT, B3LYP, and MP2), and the most ideal was chosen. It is important to take note of the limitations of the student edition, as these limitations very much provided obstacles and steered the direction of this exploration.

The analogs were chosen based on the products of a simple Heck coupling reaction with an aryl bromide and styrene. Note that no functional derivatives of styrene were chosen, since this would have provided an extra step and complication during the attempted synthesis.

As a means of comparing the changes in orbital energies, bromoindoles were also ran through the models. The bromoindoles correlate with the analogs, hence 6-bromoindole would be the reagent for 6-styryl-*mono*-indole. Phenol analogs were built and based off of the UAAs from Cheruku's study as another means of comparison, if not also to examine their electronic structures. Likewise, phenol and para-cresol were to be ran through the models.

All molecules were built within Spartan Student. The energies were minimized and stable conformers computed using Molecular Mechanics, an easy model with a short computational time intended for energy minimization. This model was used before energy calculations. Finding the stable conformer and minimizing energy proved to be of vital importance for molecules with multiple conformers and geometries. As will be discussed later, molecule geometry can very much influence molecular orbital energies and the electron structure of the HOMO and LUMO. All molecules were obtained in their *trans* configuration when possible.

Three models (EDFT, B3LYP, and MP2) were ran with the basis set 6-31G*. MP2 was ran twice when possible using 6-311+G**, a more extensive basis set with p-polarization functions, diffuse functions, and an additional function for valence orbital electrons. The 6-311+G** best resembled the 6-311++G(d,p) basis set used in the study by Chrostowska et al (2014). Running the energy models proved time consuming, with computational times ranging from thirty-minutes to about two hours.

Electrostatic potential maps were obtained for all molecules from the EDF2/6-31G* model due to unexpected graphics errors that occurred with the MP2 model. Overall, potential maps did not vary much between models.

To check the accuracy of the models, sample molecules of indole and the BN "external" indole were built in Spartan and compared to empirical values, and MP2/6-

311+G** was found to agree excellently with empirical and computational values from Chrostowska et al. BN "fused" indole was another compound utilized in the study, yet Spartan Student was unable to construct the desired geometry of the molecule, resulting in faulty molecular orbitals; it is possible that Spartan did not recognize resonance effects for BN "fused" indole.

Unfortunately, MP2/6-311+G** proved unusable with the styryl analogs due to Spartan Student's atom limitations. EDF2 and B3LYP were used with the 6-311+G** basis set as an alternative. However, EDF2 and B3LYP models proved to be relatively inaccurate at calculating HOMO and LUMO energies.

Data from each model was then compiled and organized into a table.

Heck Coupling Reaction

The standard Heck coupling procedure was followed according to Ciu et al, 2007. It is necessary, nonetheless, to elaborate on the specific steps taking in order to complete the experimental portion of this exploration as well as to shed some light as to the short-comings of the results. The overall set-up required a Schlenk line.

In consideration of the possible reactivity of the reaction mixture with water and oxygen, the solvent and reaction flasks were prepared under a home-built glove box that utilized positive-pressure N_2 gas in order to maintain an inert atmosphere and to prevent contamination of the reaction flasks with water and oxygen. This was also necessary considering DMF's hydrophilicity.

Overall, two reaction flasks were prepared. Into Flask 1 was added 0.9821 g 5-bromoindole, 0.7870 g styrene (~10 ml), 1.3866 g K_2CO_3 and an excess of Pd(quinoline-8-carboxylate)$_2$ (>0.0002 g). Flask 2: 0.9807 g 5-bromoindole, 0.7812 g styrene, and 1.3838 g K_2CO_3, and an excess of catalyst. The anhydrous K_2CO_3 serves as a desiccant. It is important to note that the catalyst's bi-dentate ligand, quinolone-8-carboxylate, also contains the oxidant that restores the catalyst.

The flasks consisted of double-necked, round-bottom flasks sealed with two rubber stoppers (one fitted with a thermostat, the other with a gas-connector) and made air-tight by applying a tiny bead of Vaseline around the stoppers before sealing.

The flasks were attached to the vacuum manifold, into which flowed the nitrogen gas provided by a large tank with a reservoir of 83-ft^3. There was no need to run the vacuum pump for this reaction, so during the entire reaction the flasks were simply allowed to have the nitrogen gas flow over them. The rate of gas flow was controlled by monitoring the number of bubbles per second that passed through the oil bubbler, which was filled with mineral oil. Due to mechanical issues with the gas-tank regulator, a flow-rate of 2-3 bubbles per second was the slowest maintainable rate.

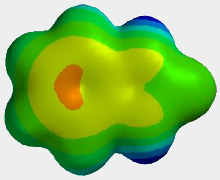
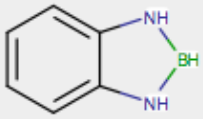
The flasks were heated to 130°C, yet the temperature fluctuated between 110°C and 145°C. This simply resulted from the uncontrollability of the power mites of the heating mantles; consequently, the temperature had to be monitored every 15 minutes in order to prevent over- or under-heating. The flasks were then stirred while heated for 9 hours. The reaction mixture turned a dark brown within an hour of heating the mixture.

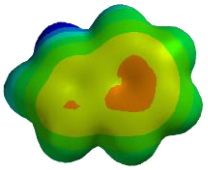
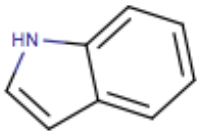
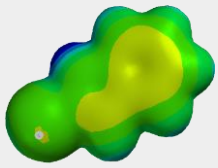
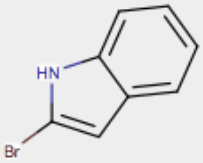
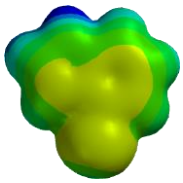
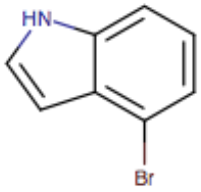
The crude reaction mixture was washed with water and extracted with ether. Using a rotary evaporator set to 30°C, the solution was evaporated off to afford a dark red mixture with excess styrene (detectable by a faint smell). This was further allowed to evaporate overnight under a vented hood to afford a light-brown solid. TLCs (EtOAc:Heptane=1:5) produced two spots: R_f= 0.32 (product) and R_f=0.79 (styrene). As the product further dried, the styrene had completely evaporated and only one spot was observed. The product's melting point was recorded. Absorption spectra were then obtained for 5-bromoindole, styrene, and the product; heptane was used as the solvent. 5-bromoindole's spectrum suggests the presence of dimerization due to the disappearance of a second peak with subsequent dilutions.

III. **Results**

QM Calculations

Table 1: Indole analogs

Analog/Structural Derivative	**EDF2/6-31G***	**B3LYP/6-31G***	**MP2/6-31G***	**MP2/6-311+G****
  **"external" BN Indole**	E-HOMO(eV): **-5.42** E-LUMO(eV): **+0.19** F.O. Gap(eV): **5.23** E-LUMO(+1)(eV): **+0.80** E-HOMO(-1)(eV): **-5.80** Molecular Energy(au): **-367.01**	E-HOMO(eV): **-5.53** E-LUMO(eV): **0.29** F.O. Gap(eV): **5.24** E-LUMO(+1)(eV): **0.90** E-HOMO(-1)(eV): **-6.0** Molecular Energy(au): **-367.29**	E-HOMO(eV): **-7.62** E-LUMO(eV): **+4.11** F.O. Gap(eV): **11.73** E-LUMO(+1)(eV): **+4.90** E-HOMO(-1)(eV): **-8.40** Molecular Energy(au): **-366.1**	E-HOMO(eV): **-7.85** (EXP: -7.9) E-LUMO(eV): **+1.57** F.O. Gap(eV): **9.47** E-LUMO(+1)(eV): **+1.60** E-HOMO(-1)(eV): **-8.60** (EXP: -8.5) Molecular Energy(au): **-366.3**

  **Indole**	E-HOMO(eV): **-5.20** E-LUMO(eV): **-0.24** F.O. Gap(eV): **4.96** E-LUMO(+1)(eV): +**0.90** E-HOMO(-1)(eV): **-5.80** Aq. Molecular Energy(au): **-363.54**	E-HOMO(eV): **-5.32** E-LUMO(eV): **-0.13** F.O. Gap(eV): **5.19** E-LUMO(+1)(eV): +**1.00** E-HOMO(-1)(eV): **-6.20** Aq. Molecular Energy(au): **-363.82**	E-HOMO(eV): **-7.43** E-LUMO(eV): +**3.54** F.O. Gap(eV): **10.97** E-LUMO(+1)(eV): +**5.00** E-HOMO(-1)(eV): **-8.10** Aq. Molecular Energy(au): **-362.65**	E-HOMO(eV): **-7.66** E-LUMO(eV): **+1.60** F.O. Gap(eV): **9.26** E-LUMO(+1)(eV): +**2.00** E-HOMO(-1)(eV): **-8.30** Aq. Molecular Energy(au): **-362.83**
  **2-Bromoindole**	E-HOMO(eV): **-5.50** E-LUMO(eV): **-0.70** F.O. Gap(eV): **4.80** E-LUMO(+1(eV)): **-0.60** E-HOMO(-1)(eV): **-6.10** Aq. Molecular Energy(au): **-2937.33**	E-HOMO(eV): **-5.62** E-LUMO(eV): **-0.58** F.O. Gap(eV): **5.04** E-LUMO(+1)(eV): **-0.50** E-HOMO(-1)(eV): **-6.20** Aq. Molecular Energy(au): **-2937.10**	E-HOMO(eV): **-7.74** E-LUMO(eV): +**3.16** F.O. Gap(eV): **10.90** E-LUMO(+1)(eV): +**3.70** E-HOMO(-1)(eV): **-8.40** Aq. Molecular Energy(au): **-2934.28**	E-HOMO(eV): **-7.93** E-LUMO(eV): **+1.57** F.O. Gap(eV): **9.50** E-LUMO(+1)(eV): **+1.90** E-HOMO(-1)(eV): **-8.60** Molecular Energy(au): **-2934.73**
  **4-Bromoindole**	E-HOMO(eV): **-5.55** E-LUMO(eV): **-0.47** F.O. Gap(eV): **5.08** E-LUMO(+1)(eV): **+0.30** E-HOMO(-1)(eV): **-6.10** Aq. Molecular Energy(au): **-2937.34**	E-HOMO(eV): **-5.56** E-LUMO(eV): **-0.55** F.O. Gap(eV): **5.01** E-LUMO(+1)(eV): **+0.10** E-HOMO(-1)(eV): **-6.30** Aq. Molecular Energy(au): **-2937.11**	E-HOMO(eV): **-7.89** E-LUMO(eV): **+3.27** F.O. Gap(eV): **+11.16** E-LUMO(+1)(eV): **+4.50** E-HOMO(-1)(eV): **-8.40** Aq. Molecular Energy(au): **-2934.28**	E-HOMO(eV): **-8.08** E-LUMO(eV): **+1.45** F.O. Gap(eV): **9.53** E-LUMO(+1)(eV): **+1.90** E-HOMO(-1)(eV): **-8.60** Molecular Energy(au): **-2934.73**

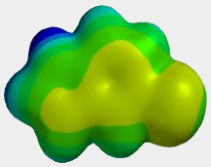
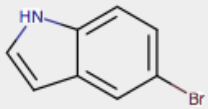
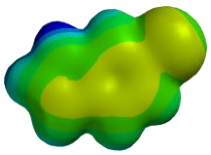
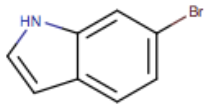
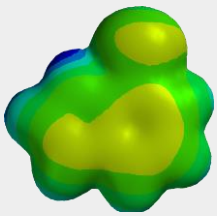
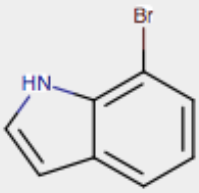
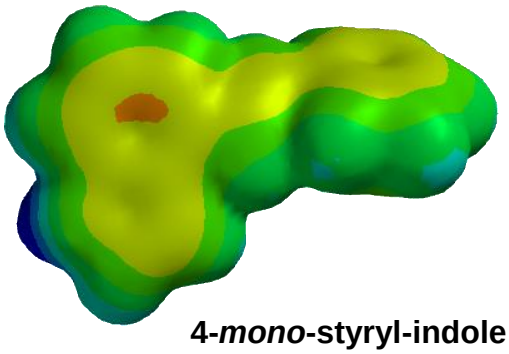
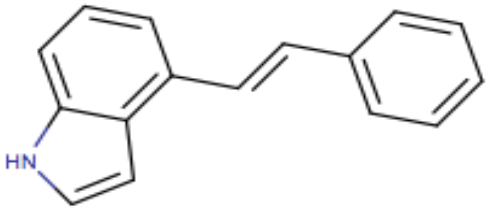
  **5-Bromoindole**	E-HOMO(eV): **-5.53** E-LUMO(eV): **-0.64** F.O. Gap(eV): **4.89** E-LUMO(+1)(eV): **+0.30** E-HOMO(-1)(eV): **-5.90** Aq. Molecular Energy(au): **-2937.34**	E-HOMO(eV): **-5.65** E-LUMO(eV): **-0.54** F.O. Gap(eV): **5.11** E-LUMO(+1)(eV): **+0.4** E-HOMO(-1)(eV): **-6.0** Aq. Molecular Energy(au): **-2937.11**	E-HOMO(eV): **-7.82** E-LUMO(eV): **+3.11** F.O. Gap(eV): **10.93** E-LUMO(+1)(eV): **+4.40** E-HOMO(-1)(eV): **-8.20** Aq. Molecular Energy(au): **-2934.29**	E-HOMO(eV): **-8.01** E-LUMO(eV): **+1.41** F.O. Gap(eV): **9.42** E-LUMO(+1)(eV): **+1.90** E-HOMO(-1)(eV): **-8.40** Aq. Molecular Energy(au): **-2934.74**
  **6-Bromoindole**	E-HOMO(eV): **-5.58** E-LUMO(eV): **-0.48** F.O. Gap(eV): **5.10** E-LUMO(+1)(eV): **+0.10** E-HOMO(-1)(eV): **-6.00** Aq. Molecular Energy(au): **-2937.34**	E-HOMO(eV): **-5.60** E-LUMO(eV): **-0.56** F.O. Gap(eV): **5.04** E-LUMO(+1)(eV): **-0.00** E-HOMO(-1)(eV): **-6.30** Aq. Molecular Energy(au): **-2937.11**	E-HOMO(eV): **-7.88** E-LUMO(eV): **+3.27** F.O. Gap(eV): **11.15** E-LUMO(+1)(eV): **+4.40** E-LUMO(+2): **+4.50** E-HOMO(-1)(eV): **-8.40** Aq. Molecular Energy(au): **-2934.29**	E-HOMO(eV): **-8.06** E-LUMO(eV): **+1.41** F.O. Gap(eV): **9.47** E-LUMO(+1)(eV): **+1.90** E-HOMO(-1)(eV): **-8.60** Aq. Molecular Energy(au): **-2934.74**
  **7-Bromoindole**	E-HOMO(eV): **-5.60** E-LUMO(eV): **-0.48** F.O. Gap(eV): **5.12** E-LUMO(+1)(eV): **0.00** E-HOMO(-1)(eV): **-6.10** Aq. Molecular Energy(au): **-2937.34**	E-HOMO(eV): **-5.60** E-LUMO(eV): **-0.55** F.O. Gap(eV): **5.05** E-LUMO(+1)(eV): **-0.20** E-HOMO(-1)(eV): **-6.30** Aq. Molecular Energy(au): **-2937.11**	E-HOMO(eV): **-7.95** E-LUMO(eV): **+3.27** F.O. Gap(eV): **11.22** E-LUMO(+1)(eV): **+4.40** E-HOMO(-1)(eV): **-8.40** Aq. Molecular Energy(au): **-2934.29**	E-HOMO(eV): **-8.14** E-LUMO(eV): **+1.55** F.O. Gap(eV): **9.69** E-LUMO(+1)(eV): E-HOMO(-1)(eV): Aq. Molecular Energy(au): **-2934.74**

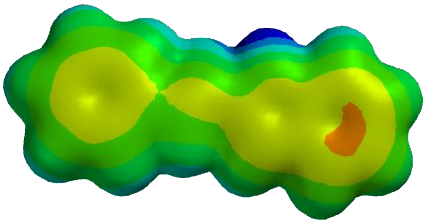
Table 2: Styryl Indole Derivatives (equilibrium geometries).

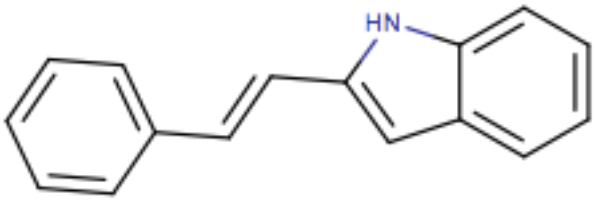




4-*mono*-styryl-indole

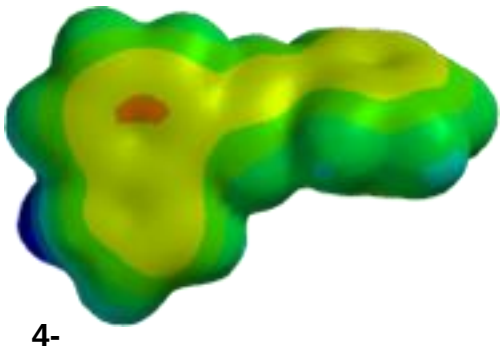
EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.11**	E-HOMO(eV): **-5.13**	E-HOMO(eV): **-5.51**	E-HOMO(eV): **-5.40**
E-LUMO(eV): **-1.01**	E-LUMO(eV): **-0.91**	E-LUMO(eV): **-1.39**	E-LUMO(eV): **-1.48**
F.O. Gap(eV): **4.1**	F.O. Gap(eV): **4.22**	F.O. Gap(eV): **4.12**	F.O. Gap(eV): **3.92**
E-LUMO(+1)(eV): **+0.10**	E-LUMO(+1)(eV): **-0.10**	E-LUMO(+1)(eV): **-0.40**	E-LUMO(+1)(eV): **-0.50**
E-HOMO(-1)(eV): **-5.80**	E-HOMO(-1)(eV): **-6.00**	E-HOMO(-1)(eV): **-6.30**	E-HOMO(-1)(eV): **-6.20**
Aq. Molecular Energy(au): **-671.76**	Aq. Molecular Energy(au): **-672.28**	Aq. Molecular Energy(au): **-672.45**	Aq. Molecular Energy(au): **-671.93**

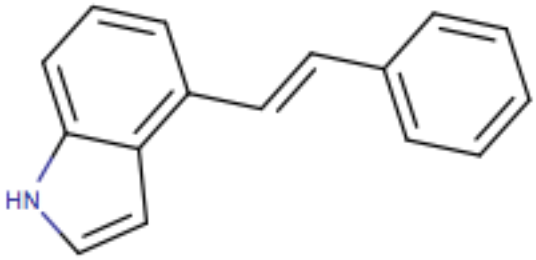




2-*mono*-styryl-indole

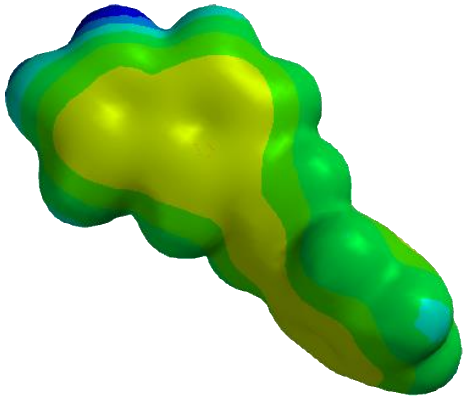
EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.04** E-LUMO(eV): **-1.34** F.O. Gap(eV): **3.70** E-LUMO(+1)(eV): **-0.10** E-HOMO(-1)(eV): **-5.60** Aq. Molecular Energy(au): **-671.56**	E-HOMO(eV): **-5.14** E-LUMO(eV): **-1.24** F.O. Gap(eV): **3.90** E-LUMO(+1)(eV): **0.00** E-HOMO(-1)(eV): **-5.10** Aq. Molecular Energy(au): **-672.28**	E-HOMO(eV): **-5.50** E-LUMO(eV): **-1.70** F.O. Gap(eV): **3.80** E-LUMO(+1)(eV): **-0.50** E-HOMO(-1)(eV): **-6.10** Aq. Molecular Energy(au): **-672.46**	E-HOMO(eV): **-5.56** E-LUMO(eV): **-1.66** F.O. Gap(eV): **3.90** E-LUMO(+1)(eV): **-0.70** E-HOMO(-1)(eV): **-6.20** Aq. Molecular Energy(au): **-671.92**

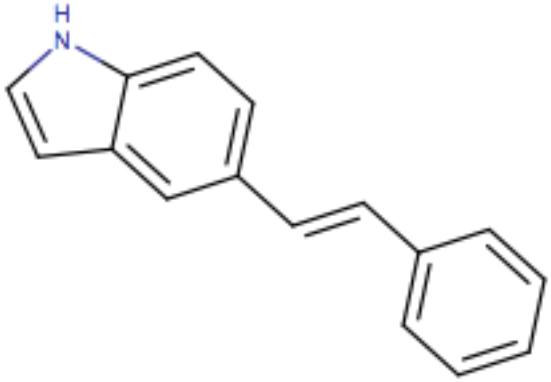




4- ***mono*-styryl-indole**

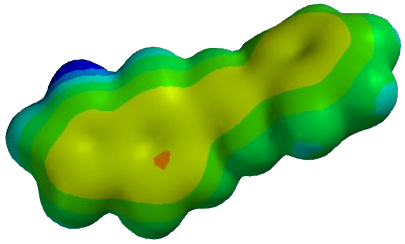
EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.11** E-LUMO(eV): **-1.01** F.O. Gap(eV): **4.1** E-LUMO(+1)(eV): **+0.10** E-HOMO(-1)(eV): **-5.80** Aq. Molecular Energy(au): **-671.76**	E-HOMO(eV): **-5.13** E-LUMO(eV): **-0.91** F.O. Gap(eV): **4.22** E-LUMO(+1)(eV): **-0.10** E-HOMO(-1)(eV): **-6.00** Aq. Molecular Energy(au): **-672.28**	E-HOMO(eV): **-5.51** E-LUMO(eV): **-1.39** F.O. Gap(eV): **4.12** E-LUMO(+1)(eV): **-0.40** E-HOMO(-1)(eV): **-6.30** Aq. Molecular Energy(au): **-672.45**	E-HOMO(eV): **-5.40** E-LUMO(eV): **-1.48** F.O. Gap(eV): **3.92** E-LUMO(+1)(eV): **-0.50** E-HOMO(-1)(eV): **-6.20** Aq. Molecular Energy(au): **-671.93**

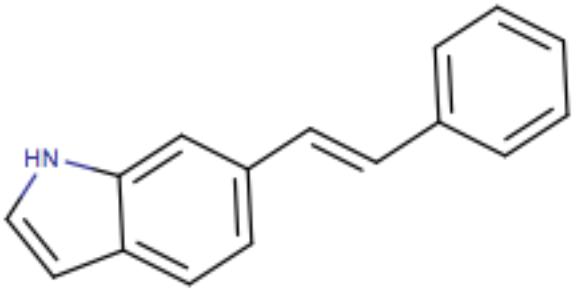




5-*mono*-styryl-indole

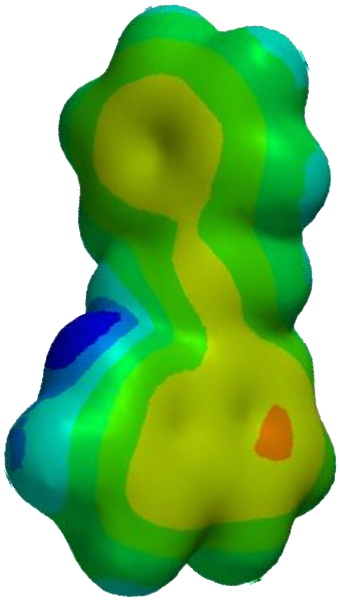
EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.17** E-LUMO(eV): **-0.82** F.O. Gap(eV): **4.35** E-LUMO(+1)(eV): **-0.20** E-HOMO(-1)(eV): **-5.40** Aq. Molecular Energy(au): **-671.75**	E-HOMO(eV): **-5.28** E-LUMO(eV): **-0.72** F.O. Gap(eV): **4.56** E-LUMO(+1)(eV): **-0.10** E-HOMO(-1)(eV): **-5.50** Aq. Molecular Energy(au): **-672.28**	E-HOMO(eV): **-5.65** E-LUMO(eV): **-1.21** F.O. Gap(eV): **4.44** E-LUMO(+1)(eV): **-0.70** E-HOMO(-1)(eV): **-5.90** Aq. Molecular Energy(au): **-672.45**	E-HOMO(eV): **-5.54** E-LUMO(eV): **-1.30** F.O. Gap(eV): **4.24** E-LUMO(+1)(eV): **-0.80** E-HOMO(-1)(eV): **-5.80** Aq. Molecular Energy(au): **-671.93**

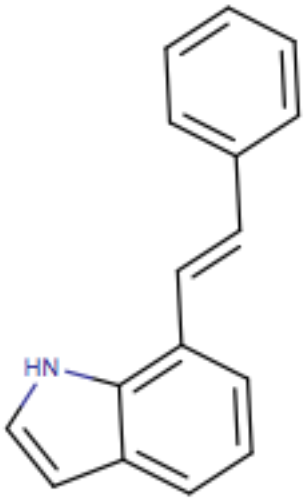




6-*mono*-styryl-indole

EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.09** E-LUMO(eV): **-1.00** F.O. Gap(eV): **4.09** E-LUMO(+1)(eV): **0.00** E-HOMO(-1)(eV): **-5.70** Aq. Molecular Energy(au): **-671.76**	E-HOMO(eV): **-5.20** E-LUMO(eV): **-0.80** F.O. Gap(eV): **4.40** E-LUMO(+1)(eV): **0.00** E-HOMO(-1)(eV): **-5.80** Aq. Molecular Energy(au): **-672.28**	E-HOMO(eV): **-5.54** E-LUMO(eV): **-1.30** F.O. Gap(eV): **4.24** E-LUMO(+1)(eV): **-0.60** E-HOMO(-1)(eV): **-6.20** Aq. Molecular Energy(au): **-672.45**	E-HOMO(eV): **-5.44** E-LUMO(eV): **-1.40** F.O. Gap(eV): **4.04** E-LUMO(+1)(eV): **-0.70** E-HOMO(-1)(eV): **-6.10** Aq. Molecular Energy(au): **-671.93**

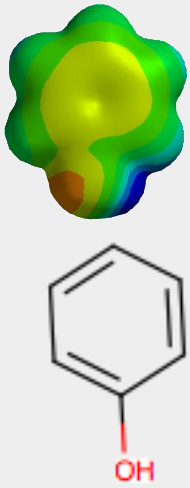
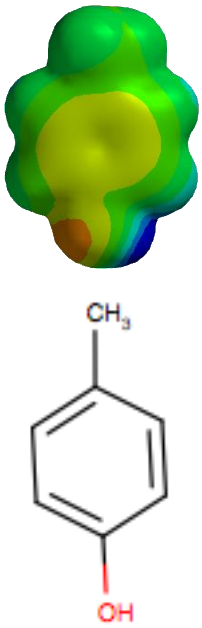


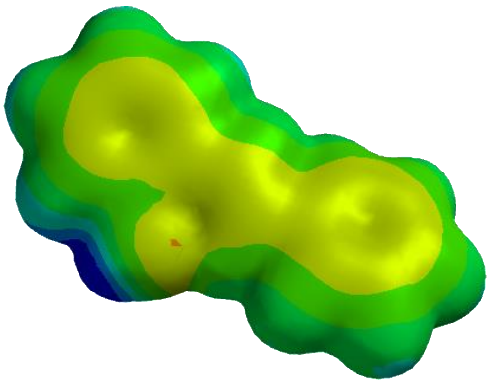


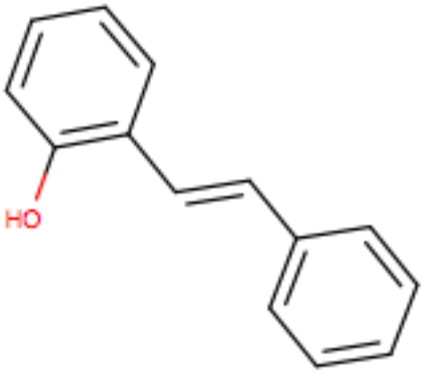
7-*mono*-styryl-indole

EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.21**	E-HOMO(eV): **-5.23**	E-HOMO(eV): **-5.61**	E-HOMO(eV): **-5.50**
E-LUMO(eV): **-1.09**	E-LUMO(eV): **-0.96**	E-LUMO(eV): **-1.44**	E-LUMO(eV): **-1.54**
F.O. Gap(eV): **4.12**	F.O. Gap(eV): **4.27**	F.O. Gap(eV): **4.17**	F.O. Gap(eV): **3.96**
E-LUMO(+1)(eV): **-0.00**	E-LUMO(+1)(eV): **0.00**	E-LUMO(+1)(eV): **-0.60**	E-LUMO(+1)(eV): **-0.70**
E-HOMO(-1)(eV): **-5.7**	E-HOMO(-1)(eV): **-5.90**	E-HOMO(-1)(eV): **-6.30**	E-HOMO(-1)(eV): **-6.10**
Aq. Molecular Energy(au): **-671.75**	Aq. Molecular Energy(au): **-672.28**	Aq. Molecular Energy(au): **-672.45**	Aq. Molecular Energy(au): **-671.93**

Table 3: Phenol Derivatives (from Cheruku)

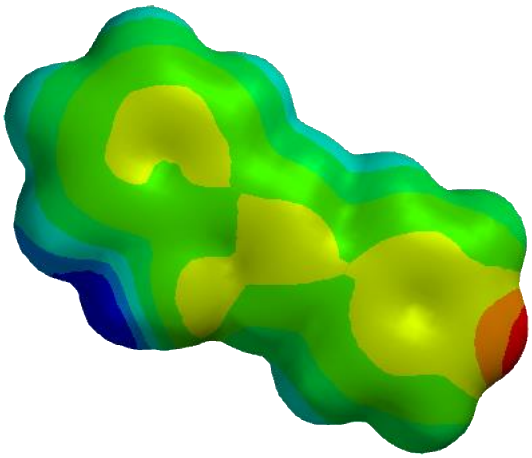
Analog/Structural Derivative	**EDF2/6-31G***	**B3LYP/6-31G***	**MP2/6-31G***	**MP2/6-311+G****
 **Phenol**	E-HOMO(eV): **-5.82** E-LUMO(eV): **-0.09** F.O. Gap(eV): **5.73** E-LUMO(+1)(eV): **+0.40** E-HOMO(-1)(eV): **-6.60** Molecular Energy(au): **-307.25**	E-HOMO(eV): **-5.95** E-LUMO(eV): **0.02** F.O. Gap(eV): **5.97** E-LUMO(+1)(eV): **0.50** E-HOMO(-1)(eV): **-6.80** Molecular Energy(au): **-307.47**	E-HOMO(eV): **-8.42** E-LUMO(eV): **+3.80** F.O. Gap(eV): **12.22** E-LUMO(+1)(eV): **+4.40** E-HOMO(-1)(eV): **-9.20** Molecular Energy(au): **-306.50**	E-HOMO(eV): **-8.65** E-LUMO(eV): **+1.79** F.O. Gap(eV): **10.44** E-LUMO(+1)(eV): **+2.00** E-HOMO(-1)(eV): **-9.40** Molecular Energy(au): **-306.67**
 ***p*-cresol**	E-HOMO(eV): **-5.64** E-LUMO(eV): **+0.05** F.O. Gap(eV): **5.69** E-LUMO(+1)(eV): **+0.50** E-HOMO(-1)(eV): **-6.60** Molecular Energy(au): **-346.53**	E-HOMO(eV): **-5.78** E-LUMO(eV): **0.15** F.O. Gap(eV): **5.93** E-LUMO(+1)(eV): **0.60** E-HOMO(-1)(eV): **-6.70** Molecular Energy(au): **-346.79**	E-HOMO(eV): **-8.21** E-LUMO(eV): **+3.92** F.O. Gap(eV): **12.13** E-LUMO(+1)(eV): **+4.50** E-HOMO(-1)(eV): **-9.20** Molecular Energy(au): **-345.67**	E-HOMO(eV): **-8.41** E-LUMO(eV): **+1.74** F.O. Gap(eV): **10.15** E-LUMO(+1)(eV): **2.00** E-HOMO(-1)(eV): **-9.30** Molecular Energy(au): **-345.87**

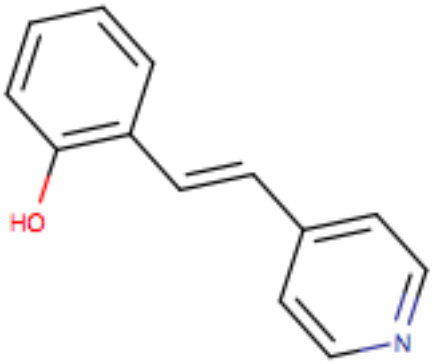




(4a) 2-styrylphenol

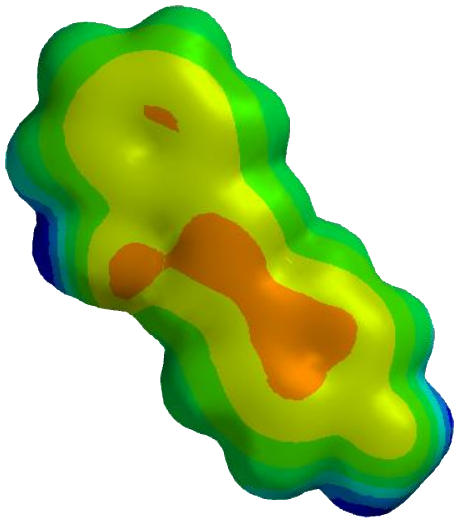
EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.30** E-LUMO(eV): **-1.05** F.O. Gap(eV): **4.25** E-LUMO(+1)(eV): **-0.10** E-HOMO(-1)(eV): **-6.20** Aq. Molecular Energy(au): **-615.46**	E-HOMO(eV): **-5.41** E-LUMO(eV): **-0.95** F.O. Gap(eV): **4.46** E-LUMO(+1)(eV): **+0.30** E-HOMO(-1)(eV): **-6.30** Aq. Molecular Energy(au): **-615.93**	E-HOMO(eV): **-5.79** E-LUMO(eV): **-1.41** F.O. Gap(eV): **4.38** E-LUMO(+1)(eV): **-0.40** E-HOMO(-1)(eV): **-6.70** Aq. Molecular Energy(au): **-616.10**	E-HOMO(eV): **-5.69** E-LUMO(eV): **-1.50** F.O. Gap(eV): **4.19** E-LUMO(+1)(eV): **-0.50** E-HOMO(-1)(eV): **-6.60** Aq. Molecular Energy(au): **-615.63**

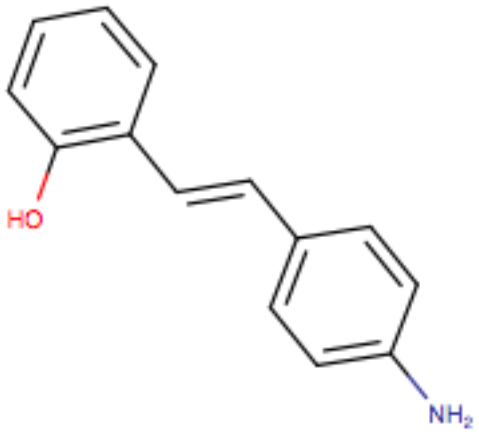




(4f) 2-(4-vinylpyridine)-phenol

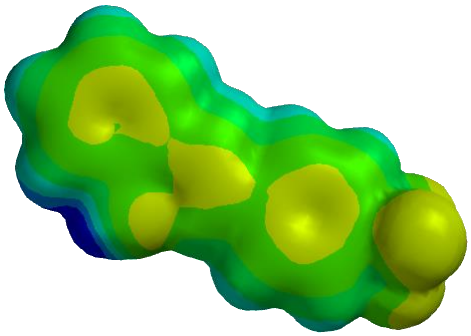
EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.59**	E-HOMO(eV): **-5.70**	E-HOMO(eV): **-6.03**	E-HOMO(eV): **-5.98**
E-LUMO(eV): **-1.67**	E-LUMO(eV): **-1.57**	E-LUMO(eV): **-2.02**	E-LUMO(eV): **-2.11**
F.O. Gap(eV): **3.92**	F.O. Gap(eV): **4.13**	F.O. Gap(eV): **4.01**	F.O. Gap(eV): **3.87**
E-LUMO(+1)(eV): **-0.10**	E-LUMO(+1)(eV): **0.00**	E-LUMO(+1)(eV): **-0.70**	E-LUMO(+1)(eV): **-0.70**
E-HOMO(-1)(eV): **-6.40**	E-HOMO(-1)(eV): **-6.50**	E-HOMO(-1)(eV): **-6.90**	E-HOMO(-1)(eV): **-6.80**
Aq. Molecular Energy(au): **-631.51**	Aq. Molecular Energy(au): **-631.98**	Aq. Molecular Energy(au): **-632.15**	Aq. Molecular Energy(au): **-631.68**

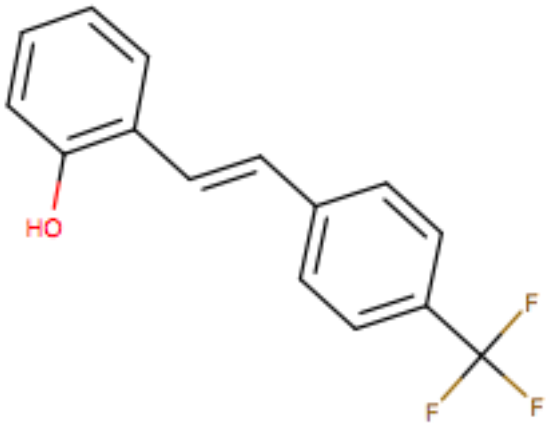




(4c) 2-(4-vinylaniline)-phenol

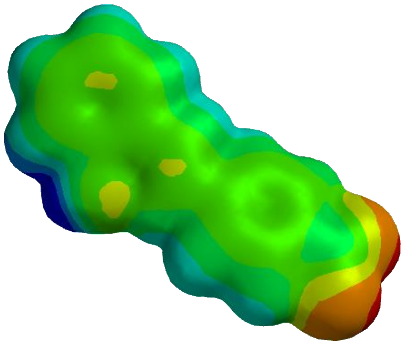
EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-4.55** E-LUMO(eV): **-0.71** F.O. Gap(eV): **3.84** E-LUMO(+1)(eV): **+0.30** E-HOMO(-1)(eV): **-5.80** Aq. Molecular Energy(au): **-670.79**	E-HOMO(eV): **-4.66** E-LUMO(eV): **-0.62** F.O. Gap(eV): **4.04** E-LUMO(+1)(eV): **+0.40** E-HOMO(-1)(eV): **-5.90** Aq. Molecular Energy(au): **-671.29**	E-HOMO(eV): **-5.07** E-LUMO(eV): **-1.10** F.O. Gap(eV): **3.97** E-LUMO(+1)(eV): **-0.30** E-HOMO(-1)(eV): **-6.30** Aq. Molecular Energy(au): **-671.48**	E-HOMO(eV): **-4.97** E-LUMO(eV): **-1.19** F.O. Gap(eV): **3.78** E-LUMO(+1)(eV): **-0.30** E-HOMO(-1)(eV): **-6.20** Aq. Molecular Energy(au): **-670.98**

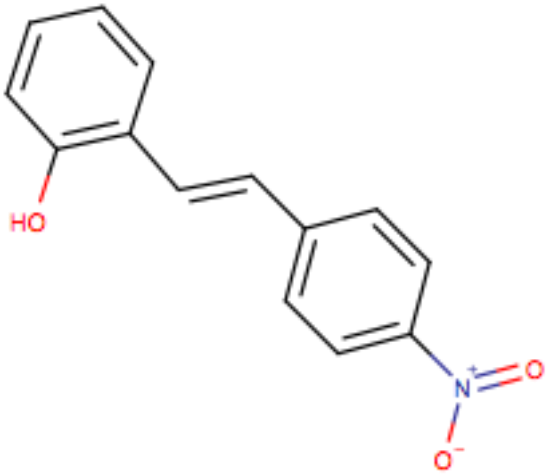




(4e) 2-(4-vinyltrifluoromethylbenzene)-phenol

EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.49** E-LUMO(eV): **-1.66** F.O. Gap(eV): **3.83** E-LUMO(+1)(eV): **-1.20** E-HOMO(-1)(eV): **-6.60** Aq. Molecular Energy(au): **-819.88**	E-HOMO(eV): **-5.60** E-LUMO(eV): **-1.56** F.O. Gap(eV): **4.04** E-LUMO(+1)(eV): **-0.20** E-HOMO(-1)(eV): **-6.50** Aq. Molecular Energy(au): **-952.97**	E-HOMO(eV): **-6.02** E-LUMO(eV): **-2.05** F.O. Gap(eV): **3.97** E-LUMO(+1)(eV): **-0.80** E-HOMO(-1)(eV): **-6.90** Aq. Molecular Energy(au): **-953.25**	E-HOMO(eV): **-5.92** E-LUMO(eV): **-2.15** F.O. Gap(eV): **3.77** E-LUMO(+1)(eV): **-0.90** E-HOMO(-1)(eV): **-6.80** Aq. Molecular Energy(au): **-952.67**





(4d) 2-(4-nitrobenzene)-phenol

EDF2/6-31G*	**B3LYP/6-31G***	**B3LYP/6-311+G****	**EDF2/6-311+G****
E-HOMO(eV): **-5.87** E-LUMO(eV): **-2.54** F.O. Gap(eV): **3.33** E-LUMO(+1)(eV): **-0.30** E-HOMO(-1)(eV): **-6.40** Aq. Molecular Energy(au): **-952.39**	E-HOMO(eV): **-5.97** E-LUMO(eV): **-2.43** F.O. Gap(eV): **3.54** E-LUMO(+1)(eV): **-1.10** E-HOMO(-1)(eV): **-6.70** Aq. Molecular Energy(au): **-820.43**	E-HOMO(eV): **-6.37** E-LUMO(eV): **-2.95** F.O. Gap(eV): **3.42** E-LUMO(+1)(eV): **-1.60** E-HOMO(-1)(eV): **-7.10** Aq. Molecular Energy(au): **-820.66**	E-HOMO(eV): **-6.27** E-LUMO(eV): **-3.05** F.O. Gap(eV): **3.22** E-LUMO(+1)(eV): **-1.70** E-HOMO(-1)(eV): **-7.00** Aq. Molecular Energy(au): **-820.11**

Table 4: Dipole Moments of Reactants and Predicted Product

5-bromoindole	Styrene	(E)-5-*mono*-styryl-indole	(E)-21-*mono*-styryl indole
4.11 debeye	0 debeye	2.71 debeye	1.58 debeye

HOMO and LUMO Maps for Qualitative Comparison and FO Theory Considerations

*Figure 1: (EDF2/6-311+G**) LUMO of 5-mono-styryl-indole (top left), 2-styryl-mono-indole (top right), 7-mono-styryl-indole (bottom). A quick look shows the effects of the styryl substituent on the relative energies of the LUMO. Compare the lobe sizes of LUMO lobes near substituent.*

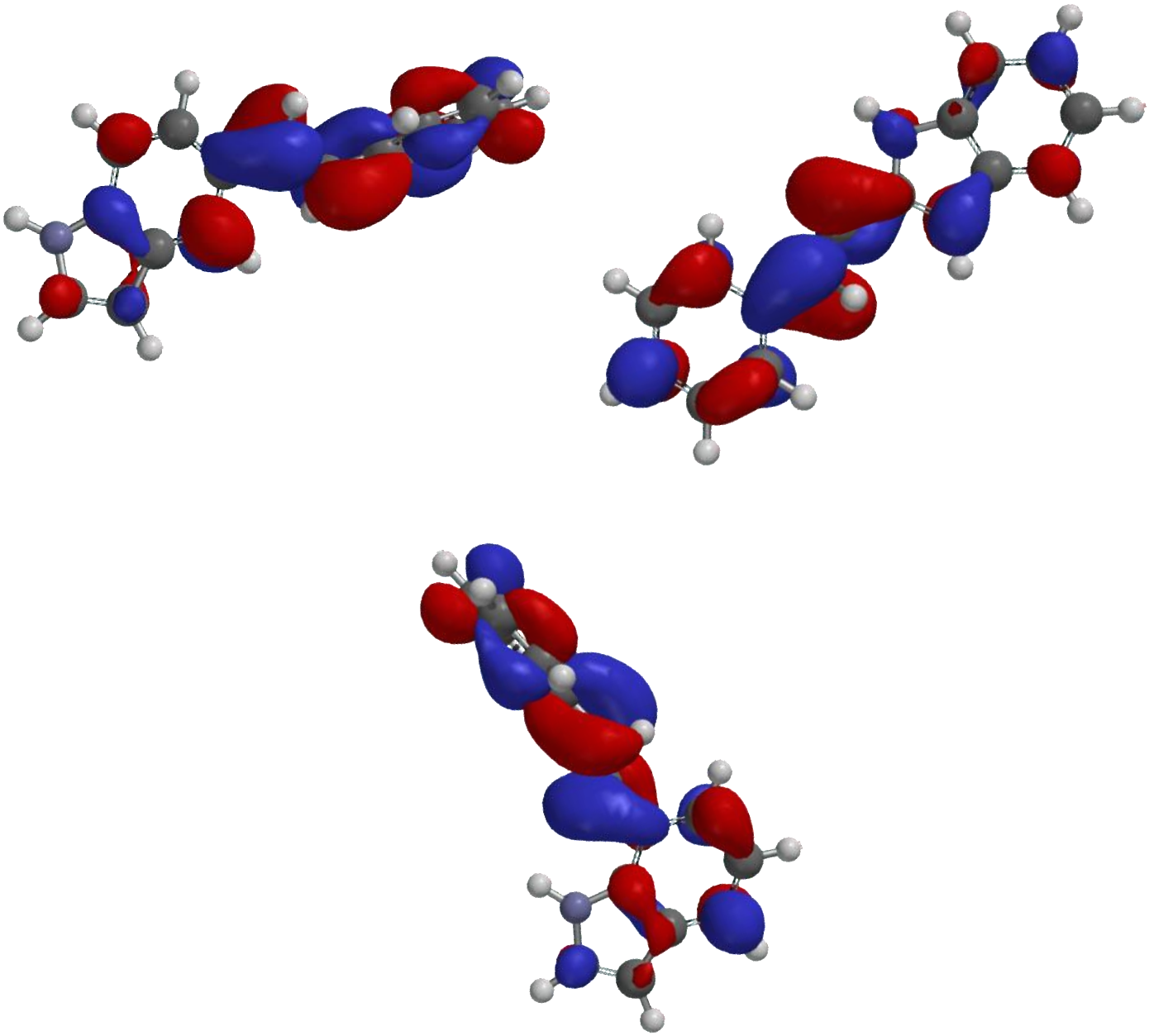


*Figure 2: (EDF2/6-311+G**) HOMO of* 5-mono-*styryl-indole (top left),* 2-mono-*styryl-indole (top-right),* 7-mono-*styryl-indole (bottom); Notice that the 5-styryl analog's HOMO lobes over the indole structure is larger than the 7-styryl analog due to the greater electron-releasing effects of the substituent owing to its closer position to the electrophilic N-1 site. The HOMO structure is completely changed for the 2-styryl analog.*

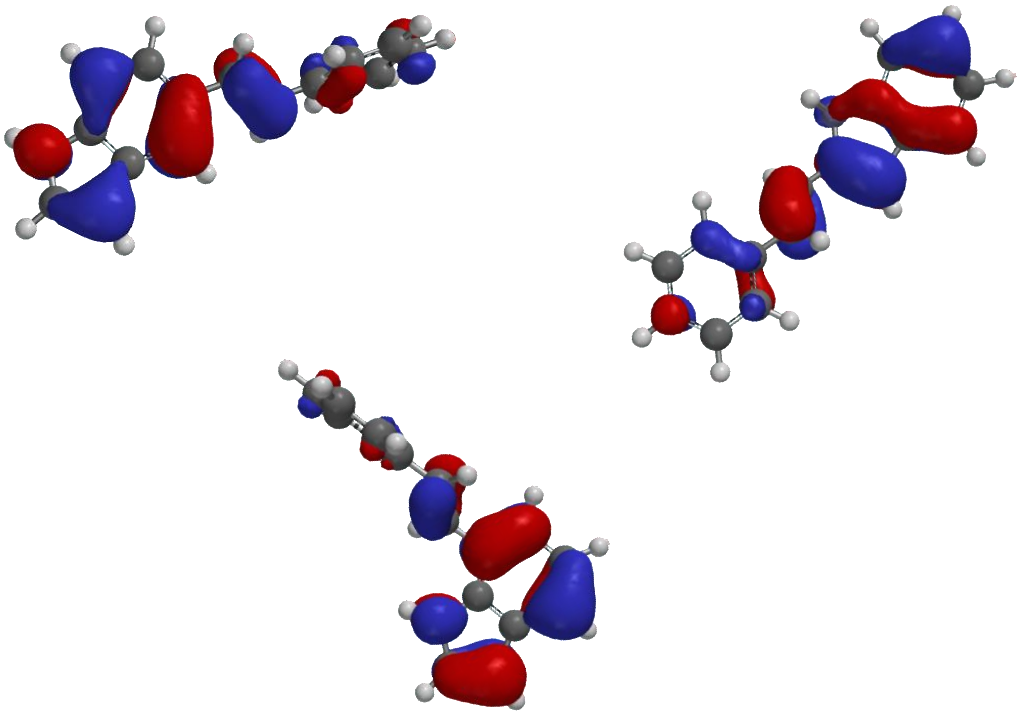
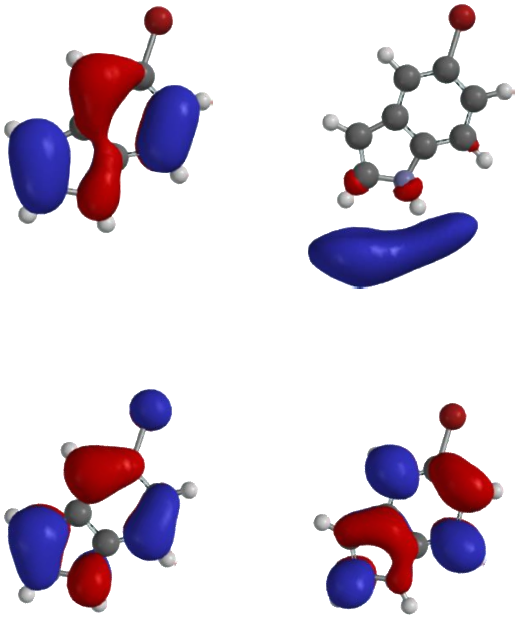


*Figure 3: Using MP2/6-311+G** 5-bromoindole HOMO (top left) and LUMO (top right); EDF2/6-311+G**, HOMO (bottom left) and LUMO (bottom right).*



Synthesis Results

Figure 5: 5-bromoindole absorption spectra, in heptane, λ_{max} 226 nm. Note the evidence of either possible formations of dimers, trimers, and tetramers: first measurement (top-left); first dilution with two peaks (top-right), second dilution with reduced peak (bottom-left), third dilution with monomer peak (bottom-right).

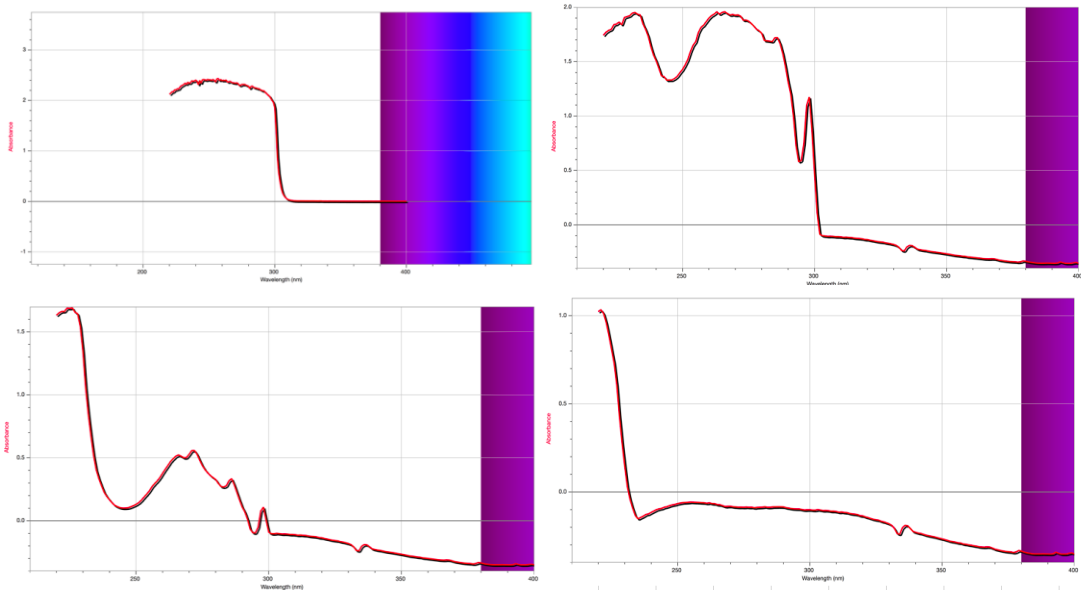


Figure 6: Styrene absorption spectra, in heptane, λ_{max} 246 nm.

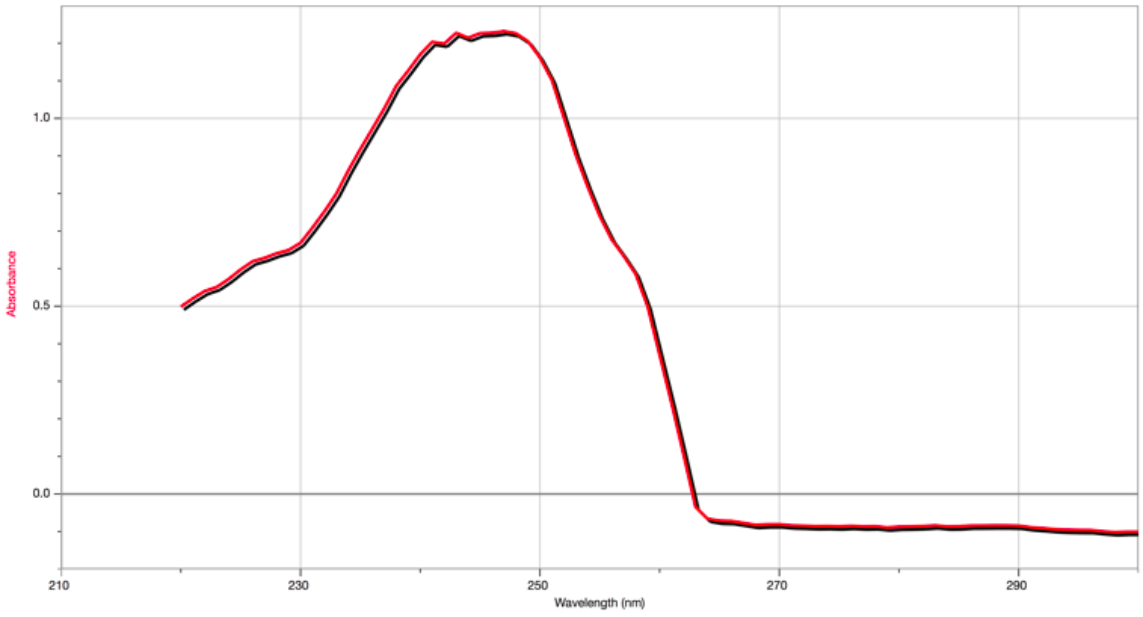


Figure 7: Product absorption spectra, in heptane, λ_{max} 256 nm.

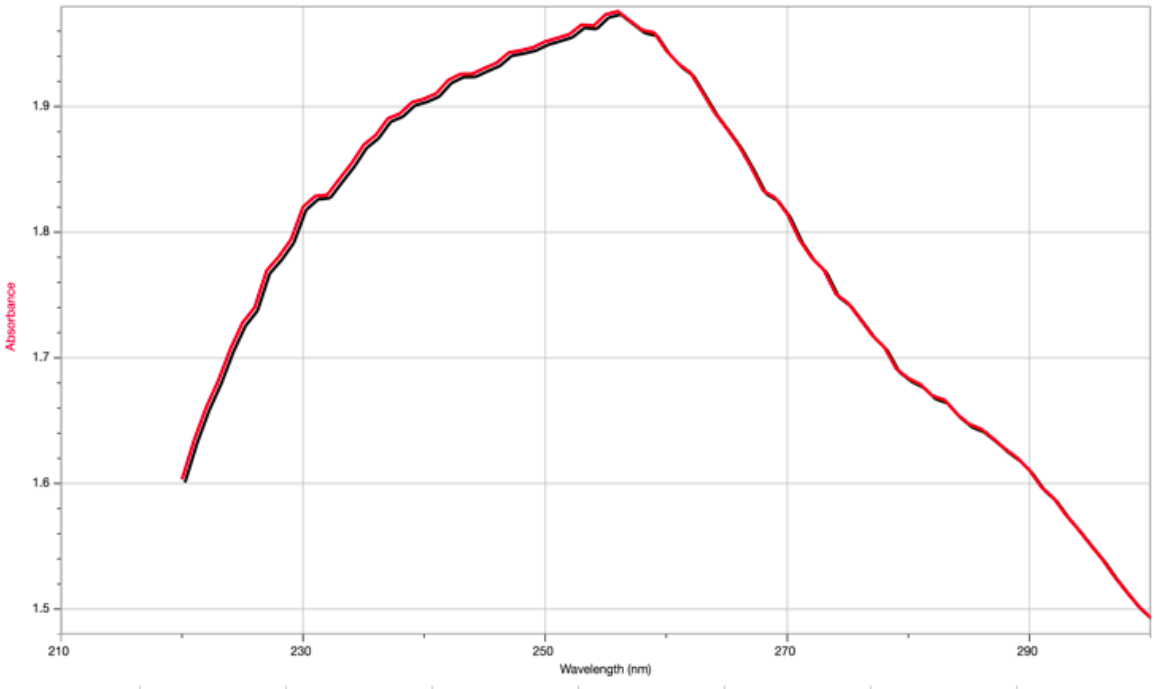


Table 5: Melting temperatures and retention times. TLC silica gel 60 with EtOAc:Heptane (1:5).

	5-bromoindole	Unknown Product	Styrene
Melting Point	92 C	84 C	---
R_f	0.27	0.32	0.79

IV. Discussion and Conclusion

It is important to remember that this study serves more as an exploration of styryl indole analogs rather than tryptophan derivatives, stemming from the initial applications towards possible uses as biochemical labels. Throughout the study, several limitations have influenced the direction taken.

Returning to the initial concept—unnatural amino acids—I proposed that tryptophan analogs could possibly serve as unique labels for fluorescent studies, owing to its extended π-conjugated system centered around its indole sidechain, a concept similar to Cheruku (2015). Unfortunately, faced with limited resources and tools, I made an approximation of sorts, focusing instead on indole. This approximation can work given the indole (and indole analog) sidechain serving as the fluorophore of said amino acid.

A computational analysis was planned in order to predict an absorption spectrum of a styryl-indole compound. Because of the Spartan Students limitations, and lack of a time-dependent model, atom systems had to be reduced. This was the first roadblock. However, the computations were still followed out for two main reasons: to make

predictions based off of HOMO/LUMO energy gaps and to examine electron structures to plan a synthesis.

While I was able to find the ideal basis set and model (MP2-6-311+G**) for the QM calculations of the indole reagents, the theoretically expected product was too large to use the MP2. The MP2 model matched experimental values from Chrostowska et al (2014) excellently; in spite of that, Chrostowska et al (2014) used a time-dependent model for their study and were able to create a theoretical UV-Vis spectrum for their BN indoles, which were smaller molecules than the styryl indoles of this study. Of course, a ground-state calculation would produce more accurate molecular orbital energies than a time-dependent calculation. However, the accuracy of MP2/6-311+G** fell greatly after HOMO-1 and HOMO-2, -3 and so forth did not match empirically calculated values as well. I suspect this has to do with my basis set including fewer diffuse functions than those used in Chrostowska et al (2014). Also, MP2 uses second-order corrections whereas Chrostowska et al (2014) used third-order corrections. Again, Spartan Student provides everything through a GUI with a limited selection of basis sets.

Because the styryl indole analogs were too large for the most accurate model (MP2), I had to carry out the calculations using EDF2 and B3LYP with basis sets 6-311+G** and 6-31G*. There was little difference between the four combinations of model/basis set.

Indeed, extending the π-conjugated system showed a decrease in the frontier orbital gap (HOMO/LUMO energy gap) as expected. However, the true value of that gap cannot be inferred due to the inaccuracies of B3LYP and EDF2 relative to MP2 for these molecules.

The fluorophores of the various unnatural tyrosine utilized in Cheruku (2015) were created in Spartan Student in hopes of correlating the HOMO/LUMO energy gap with maximum absorbed wavelengths. However, this cannot be done; the actual fluorophores would have to be synthesized, with their substituents varied, after which the variations of these substituents would allow for the correlation between absorbed wavelength and the frontier orbital gap. The frontier orbital gap, though it can be used to make inferences about the molecule's color and the energy needed to excite an electron to various molecular orbitals, cannot necessarily be used to infer fluorescent properties, especially given the low accuracy of the models and the fact that these were ground state calculations. Nonetheless, Table 3 displays the fluorophores that relate to the unnatural tyrosine from Cheruku (2015).

Substituent Effect of Styrene on HOMO/LUMO energy gap (ground state)

Though molecular energies of the styryl analogs stayed relatively the same (about -672 eV), there is one interesting effect of molecular orbital energy reduction, related to the frontier orbital theory, that depends on the position of the styryl group relative to the nitrogen in the indole analogs, displayed in Table 2. Styrene, acting as an electron releasing group, is shown to reduce the LUMO and raise slightly the HOMO energies, with the energy-lowering effect observed in the LUMO being more pronounced than the energy-raising effect seen in the HOMO as it is gradually moved closer to the

nitrogen atom. This reduction in energy of the LUMO that is greater than the increase in the HOMO results in a reduction of the frontier orbital gap (and hence the energy required to excite the first electron of that molecule). There may be two ways of looking at this.

First, the nitrogen, being quite electronegative (see electrostatic maps in Table 2), exerts a stronger 'pull' on the electrons of the styryl group. This translates to an increase in the electron releasing effect of the styryl group as it is moved closer to nitrogen (compare 2-*mono*-styryl indole energy gap to that of 5-*mono*-styryl indole). Notice, also, that 5-*mono*-indole's frontier orbital gap is the largest, owing to the C-5 position being the farthest from N-1. This energy reduction (reflected as a red-shift in the absorption spectrum) can be pinpointed to stabilization via charge distribution i.e. pi electrons contributed by the styryl group can on average interact more with the nitrogen nucleus. There is also slight change in geometry, with the phenyl group being planar relative to indole in the 2-styryl analog, compared to a near 90-degree rotation found in the 5-styrylindole; this may affect the molecular orbitals and nitrogen's relative contribution. One can suspect that should such compounds be used as co-factors or protein labels, geometry conformations may vary.

The dipole moment (See Table 4) of the 2-*mono*-styrylindole is also lower (1.58 debeye) than that of 5-*mono*-styrylindole (2.71 debeye), another reflection of the charges being more evenly distributed. To provide the contrasting effect, take the brominated indoles displayed in Table 1 (MP2/6-311+G**): 2-bromoindole and 5-bromoindole. The change in the frontier orbital gap is reversed for an electron withdrawing group, such as any halide. 2-bromoindole's LUMO energy (+1.57 eV) is higher relative to 5-bromoindole (+1.41 eV), owing to some destabilization of an electron-withdrawing substituent closer to the electronegative nitrogen. Thus, an overall decrease in the LUMO molecular orbital (MO) energies observed in the isomers of styryl-indole seem to result from charge stabilizations. It is then important to offer a suitable explanation for the lowering of the LUMO energies and the raising of the HOMO energies as the styryl substituent is moved from the C-5 position to C-2. Because the total number of electrons for the 2- and 5-substitude indoles are the same, the effects involved relate solely to electron distribution about the molecule (and a change in the orbital coefficients of various AOs). Note, also, that B3LYP/6-311++G is taken to be the most accurate model for the styryl indoles, offering the lowest MO energies.

Nitrogen, being electronegative, is expected to have lower energy atomic orbitals (AO). The MO energies depend on the AO energies that construct them. Of course, the AOs of the nitrogen atom would have a greater contribution to the MO the closer it is to the substituent. Essentially, moving the styryl group closer to nitrogen allows the AOs of nitrogen translates to an increase in the coefficient of the nitrogen AO in the linear combination of AOs that form the MO, and, particularly, the LUMO. The decrease in the LUMO energy is more pronounced than the slight increase in the HOMO energy allows for the FO gap to be smaller for the 2-styryl analog.

The slight increase of the HOMO energies (related to the ionization energies) may be attributed to the increase in electron density about the nitrogen atom (compare

electrostatic maps of the 2- and 5-styryl indole). The increase of the HOMO energy (going from C-5 to C-2) can then be explained in terms of increased electron-electron repulsion, though actually pinpointing the exact reasons for these differences in MO energy levels between these isomers can be rather ambiguous. The energies of the HOMO and LUMO relate to their lobe sizes (See Fig. 2 and 3 for comparisons).

Synthesis

After the product was allowed to dry, the TLC analysis consistently showed two spots: styrene and the unknown product. After styrene further evaporated, the plates showed only one spot, with some streaking involved. A quick look at the melting point of the unknown product (84 C) rules out the possibility of a successful synthesis of 5-*mono*-styrylindole. However, two inferences can be made from initial observations, absorption spectra, retention times, and melting points.

Based off of my initial observations after the ether was evaporated from the product, I suspect that the styrene did not react at all—my product was dissolved in excess styrene, which amounted close to the starting amount of ~10 mL. The aqueous phase that was separated was light orange in color. It is possible that the 5-bromoindole reacted with itself.

Before delving into some conjectures as to what may have resulted during the synthesis, I find it interesting to note that the final compound had an absorbed wavelength peak 10 nm greater than styrene and 30 nm greater than 5-bromoindole. This suggests that some sort of extension of the pi-conjugated system resulted from this synthesis. The higher retention time and lower melting point of the product (relative to 5-bromoindole) hints at a more nonpolar compound. Thus, the red-shift in the absorption spectra and the decrease in polarity of the product relative to 5-bromoindole suggests that some sort of coupling reaction or pi-electron contribution occurred.

It is possible that the 5-bromoindole reacted with itself at the C-3 or C-2 position, the most and second-most reactive positions, respectively. This is understandable, since the C-3 position is the position of the backbone for tryptophan. Fig. 3 displays the HOMO and LUMO of 5-bromoindole using two models (MP2 and EDF2); both models were displayed since the MP2 model did not properly portray the LUMO, an anomaly that could be a result of my own computer system used for these calculations. I had thought that an examination of the orbitals would hint as to how 5-bromoindole may have reacted with itself, yet this analysis might be more appropriate after the product has been identified through spectroscopic techniques.

A complete indole-indole coupling is possible, though I would have expected to find a much higher melting point for a bi-indole compound. Reactions at the N-H position are also possible, though C-5 functionalization with Br was expected to direct the coupling to the C-5 position.

If anything, it can be concluded that this procedure for a Heck coupling reaction between 5-bromoindole and styrene did not produce the target molecule of 5-*mono*-styrylindole. As for styryl-tryptophan analogs, there may be possible routes of synthesis

through the Wittig reaction or synthesizing the respective UAA from 5-styrylindole by formation of the amino acid backbone at the C-3 position. Though excess catalyst was used, I doubt that any coordination compound is largely responsible for the observed physical properties of the product.

The next step would be to characterize my unknown product via NMR, for this synthesis might yield some interesting compound or additional insight into coupling reactions with indole.

Acknowledgements

I would like to thank Dr. Feza Ozturk for guiding me through this study; Dr. Brian Frederking, director of the honors program; and David Dressel, who is in charge of all the lab materials, for providing me with the equipment and reagents to complete the experimental portion of this study.

V. References

Atkins, Peter, and De Paula, Julio. *Physical Chemistry*. New York: W.H. Freeman and Company. 2006. Print.

Bartolucci, Silvia, et al. "Direct, Regioselective, and Chemoselective Preparation of Novel Boronated Tryptophans by Friedel–Crafts Alkylation."*Organic letters* 14.2 (2011): 600-603.

Bittner, S., R. Scherzer, and E. Harlev. "The five bromotryptophans." *Amino Acids* 33.1 (2007): 19-42.

Broggini, Gianluigi, et al. "Palladium-catalyzed dual C–H or N–H functionalization of unfunctionalized indole derivatives with alkenes and arenes." *Beilstein journal of organic chemistry* 8.1 (2012): 1730-1746.

Chrostowska, Anna et al. “UV-Photoelectron Spectroscopy of BN Indoles: Experimental and Computational Electronic Structure Analysis.” Journal of the American Chemical Society 136.33 (2014): 11813–11820. Web.

Cheruku, Pradeep, et al. "Tyrosine-derived stimuli responsive, fluorescent amino acids." *Chemical Science* 6.2 (2015): 1150-1158.

Cui, Xin, et al. "Pd (quinoline-8-carboxylate) 2 as a low-priced, phosphine-free catalyst for Heck and Suzuki reactions." The Journal of organic chemistry 72.24 (2007): 9342-9345.

de Jesus, Armando J., and Toby W. Allen. "The role of tryptophan side chains in membrane protein anchoring and hydrophobic mis match." *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1828.2 (2013): 864-876.

Gustavsson, Thomas, et al. "Singlet excited-state behavior of uracil and thymine in aqueous solution: a combined experimental and computa tional study of 11 uracil derivatives." *Journal of the American Chemical Society* 128.2 (2006): 607-619.

Hehre, Warren, and Ohlinger, Sean. *Spartan '10 Tutorial and User Guide*. Wave function, Inc. (2011)

Heller, Stephen T., Erica E. Schultz, and Richmond Sarpong. "Chemoselective N-Acylation of Indoles and Oxazolidinones with Carbonylazoles." *Ange wandte Chemie International Edition* 51.33 (2012): 8304-8308.

Jacquemin, Denis, Benedetta Mennucci, and Carlo Adamo. "Excited-State Cal culations with TD-DFT: From Benchmarks to Simulations in Complex Envi ronments." Phys. Chem. Chem. Phys. 13.38 (2011): 16987–16998. Web.

Janczuk, Adam, et al. "Ytterbium triflate catalyzed electrophilic substitution of indoles: the synthesis of unnatural tryptophan derivatives." *Tetrahe dron letters* 43.23 (2002): 4271-4274.

Roeser, Julien, et al. "Oxidative protein labeling in mass-spectrometry-based proteomics." *Analytical and bioanalytical chemistry* 397.8 (2010): 3441-3455.

Kitamatsu, Mizuki, Midori Futami, and Masahiko Sisido. "A novel method for screening peptides that bind to proteins by using multiple fluorescent amino acids as fluorescent tags." *Chem. Commun.* 46.5 (2010): 761-763.

Li, Min, Z. H. A. O. L. A. N. Lin, and Michael E. Johnson. "Photoaffinity labelling of cyanomethaemoglobin with derivatives of tryptophan and 5-bromo-tryptophan." *Biochem. J* 308 (1995): 251-260.

Mansoor, Steven E., Mark A. DeWitt, and David L. Farrens. "Distance Mapping in Proteins Using Fluorescence Spectroscopy: The Tryptophan-Induced Quenching (TrIQ) Method." *Biochemistry* 49.45 (2010): 9722–9731. *PMC*. Web. 14 Sept. 2015.

Masayuki Matsushita, * et al. "A Blue Fluorescent Antibody−Cofactor Sensor for Mercury." Organic Letters 7.22 (2005): 4943–4946. Web.

Mollica, Adriano, et al. "Synthesis and bioactivity of secondary metabolites from marine sponges containing dibrominated indolic systems." *Mole--cules* 17.5 (2012): 6083-6099.

Sainio, E-L., Pulkki, K., and Young, S. N. "L-Tryptophan: Biochemical, nutrition al and pharmacological aspects." Amino Acids 10.1 (1996): 21-47.

Spartan Student Edition v. 6 Topics: Theoretical Models. Wavefunction, Inc.

Yang, Yang, Hong-Mei Jia, and Bo-Li Liu. "(E)-5-styryl-1H-indole and (E)-6-styrylquinoline derivatives serve as probes for ß-amyloid plaques." *Molecules*17.4 (2012): 4252-4265.