2 and 1, respectively, and we use this representation of the molecule to explain the role of d functions.

In the absence of d functions, significant probability amplitudes of the bond orbitals centered on the sulfur atom (right panels in Figure 2a,b) lie outside the region of the bond, but when d functions are included, these parts of the orbitals are considerably reduced (Figure 1a,b). The other orbital in the bond, centered on the oxygen atom, changes to a much smaller extent. The angularly correlated lone-pair orbitals also show major changes in shape when d functions are included. In Figure 1c, the lone pair is significantly separated angularly and effectively "wraps around" the sulfur core, but in Figure 2c, the orbitals are not so effective in producing angular separation.

In these calculations[1] and those of others who have considered this question,[2] the orbitals of different electron pairs are required to be mutually orthogonal.[3] Thus in the present description of sulfur dioxide there are six lobes (or "hybrid" orbitals) which must be orthogonal to one another[4] *about the sulfur atom.* When only a double-ζ sp basis set is used, the functions centered on the sulfur atom can only remain orthogonal by having large proportions of each orbital on the opposite side of a node at the nucleus to the lobe forming a bond. This raises the kinetic energy of the electrons because of the increased curvature of the one-electron wave functions and increases their potential energy because they cannot achieve their optimum electron-nuclear attraction. We conclude that the importance of d basis functions on sulfur is in providing sufficient flexibility in the basis set to allow for six *angularly distinct* orbitals when required. Further evidence for this conclusion is provided by the following observation: a population analysis shows that the coefficients of the d basis functions in these calculations are small even though they play an important energetic role. In addition, when the coefficients *of* the d basis functions are set to 0 after the wave function has already been converged (contour plots in Figure 1) and the renormalized orbitals are plotted, there is essentially *no change* in the contours. This contrasts with the situation when the wave function is converged without the d functions and major changes in the contours are observed (Figure 2). The double-ζ sp basis functions therefore can reasonably describe the six sulfur orbitals, but the role of the d functions is to enable these orbitals to remain orthogonal to one another while attaining their optimum shapes, i.e., when d functions are included, charge is moved into the bond region *and* orbitals more effectively "wrap around" cores, thereby increasing their electron-nuclear attraction.[5]

The GVB description of hydrogen sulfide, on the other hand, has only four angularly distinct lobes centered on the sulfur atom because there are two bonds and two radially correlated lone pairs. A double-ζ sp basis set has adequate flexibility to describe this arrangement of electron pairs, and so there is no anomalously large energy lowering (0.98 eV for H_2S compared to 5.52 eV for SO_2) when d functions[6] are included in the O and S basis sets. However, when the electron-withdrawing potential of the hydrogen atoms is artificially increased, the GVB description of the electron pairs in hydrogen sulfide changes from tetrahedral to octahedral.[1] Lone pairs are now angularly correlated, and this description can only be obtained when at least two sets of d functions are included in the S atom basis set.[1] Hence it is the arrangement of electron pairs about an atomic site that dictates whether calculations on a molecule will have anomalously large energy lowerings when d functions are included in the basis set.

In summary, a combination of orthogonality constraints and an arrangement of electron pairs with more than four angularly distinct lobes in molecules containing second-row atoms require that calculations performed on these molecules should have better than double-ζ sp basis set flexibility. The importance of d functions in sulfur dioxide arises from this ability to provide additional spatial flexibility so that orbitals can remain orthogonal *and* attain their optimum shapes.

Acknowledgment. This work was supported in part by the Office of Naval Research and the National Science Foundation under MRL Grant No. DMR 8519059.

(2) (a) Guest, M. F.; Hillier, I. H.; Saunders, V. R. *J. Chem. Soc., Faraday Trans. 2* **1972**, *68*, 114. (b) Schmiedekamp, A.; Cruickshank, D. W. J.; Skaarup, S.; Pulay, P.; Hargittai, I.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2002. (c) Stromberg, A.; Wahlgren, U.; Pettersson, L.; Siegbahn, P. E. M. *Chem. Phys.* **1984**, *89*, 323. (d) Kutzelnigg, W. *Angew. Chem.* **1984**, *96*, 262; *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 272. (e) Cruickshank, D. W. J. *J. Mol. Struct.* **1985**, *130*, 177. (f) Cruickshank, D. W. J.; Eisenstein, M. *J. Mol. Struct.* **1985**, *130*, 143. (g) Cruickshank, D. W. J.; Eisenstein, M. *J. Comput. Chem.* **1987**, *8*, 6. (h) Magnusson, E.; Schaeffer, H. F. *J. Chem. Phys.* **1985**, *83*, 5721. (i) Yadav, A.; Surjan, P. R.; Poirier, R. A. *J. Mol. Struct.* (*THEOCHEM*) **1988**, *165*, 297.

(3) This is enforced by the strong orthogonality approximation to the GVB wave function. Relaxation of the strong orthogonality constraint can be achieved within the GVB model while independent particle interpretability is retained: Schultz, P. A.; Messmer, R. P. *J. Am. Chem. Soc.* **1988**, *110*, 8258. Schultz, P. A. Ph.D. Thesis, University of Pennsylvania, 1988. This is a major task for sulfur dioxide and and has not been carried out so far.

(4) GVB-SOPP orbitals of the angularly correlated lone pair overlap because they belong to the same electron pair. Although this pair of orbitals may overlap, both must be orthogonal to all other orbitals.

(5) There may be a parallel between the latter observation and the "penetration effects" introduced by d functions of second-row atoms referred to in the recent literature.[2e,g]

(6) d function exponents were $\zeta = 0.53$ for S and $\zeta = 0.80$ for O in each case where single-ζ d functions were employed.

Photodehalogenation of 4-Haloindoles

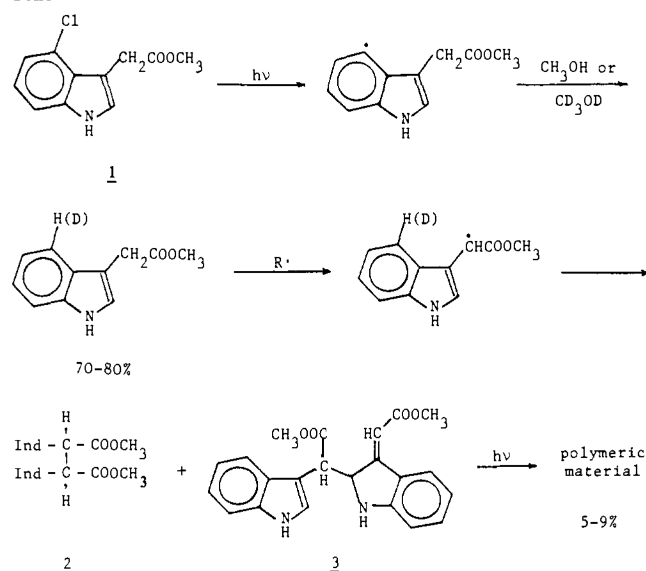
Nien-chu C. Yang,* Alex Huang, and Ding-Djung H. Yang

Department of Chemistry, University of Chicago Chicago, Illinois 60637

Received June 13, 1989

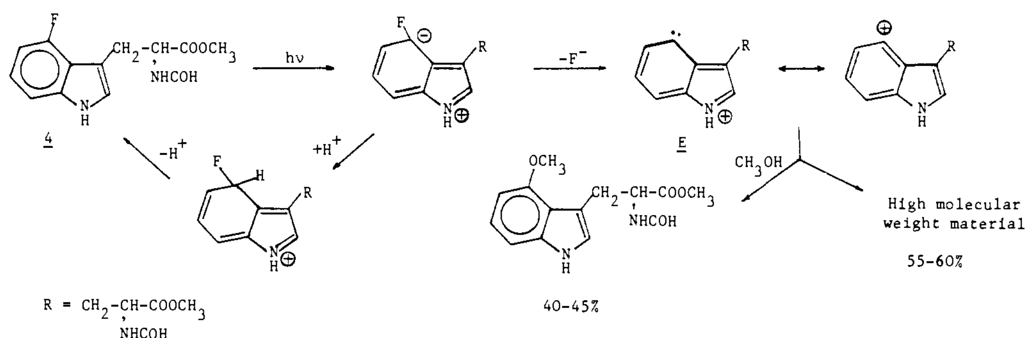
L-Tryptophan is an essential amino acid containing a bicyclic aromatic group, indole, which is highly hydrophobic in nature and often occupies critical positions in proteins.[1] Tryptophan also exhibits unique spectroscopic properties, which enables it to serve as an intrinsic probe for protein structures, protein dynamics, and intermolecular interactions between proteins and other molecules.[2] Fluorotryptophans may be incorporated into proteins via methods of molecular biology.[3] Tryptophan, 5- and 6-fluorotryptophan, and their derivatives exhibit a strong fluorescence with a quantum efficiency up to 0.25;[2,4] however, 4-haloindoles, including 4-fluorotryptophan, proteins containing 4-fluorotryptophan, and methyl 4-chloroindole-3-acetate, do not exhibit a measurable fluorescence.[5] This observation suggests that there is a unique nonradiative decay pathway present in 4-haloindoles that is not

Scheme I



0002-7863/89/1511-8060$01.50/0 © 1989 American Chemical Society

Scheme II



present in their unsubstituted, 5-halo, and 6-halo analogues. The C-4 in indole is a highly reactive position in photochemistry. Saito, Matsuura, and co-workers have shown that tryptophan undergoes a highly selective H–D exchange at the same position.[6] Witkop and Yonemitsu have demonstrated that *N*-chloroacetyl derivatives of tryptophan undergo an intramolecular cyclization at that position.[7] Both observations are in agreement with the theoretical calculation that the C-4 atom displays a considerable increase in electron density in its excited state.[8] Localization of electron density at the C-4 in the LUMO may exert a profound influence on the reactivity of the C–halogen bond at that position. We report two entirely different novel photochemical dehalogenations of *N*-acyl-4-fluorotryptophan methyl ester and methyl 4-chloro-3-indoleacetate.

Methyl 4-chloro-3-indoleacetate[9] (**1**) undergoes a remarkable facile dehalogenation under the influence of light. When a solution of **1** (80 mL of a 2.8 mM solution in methanol) was irradiated with a 200-W Hanovia Hg arc through a Vycor filter, the reaction was over in about 10 min.[10] Methyl 3-indoleacetate is the major product (70–80%). Two dimeric products, **2** and **3** (10–17% total), and a small amount of polymeric material (5–9%) are also formed. Dimers **2**, mp 119 °C, and **3**, mp 141 °C, were separated by HPLC. Dimer **3** rearranges slowly on standing to **2** and is decomposed by light (the reaction condition) to the polymeric material. When the irradiation was carried out in CD_3OD, all products were labeled at the 4-position. The results are illustrated in Scheme I.[11] (±)-*N*-Formyl-4-fluorotryptophan methyl ester (**4**) undergoes an entirely different dehalogenation in methanol to yield (±)-*N*-formyl-4-methoxytryptophan methyl ester as the only isolable and identifiable product (40–45%), and the balance of consumed starting material is converted into an unidentifiable polymeric material (Scheme II). A more intense light source was needed for this conversion (400-W Hg arc for 30 min). The *N*-acetyl derivative gives a similar result. The reaction is formally a photoinduced nucleophilic displacement of an aryl fluoride, which is unprecedented in the literature.

The dramatic difference between the photochemical behaviors of 4-chloro- and 4-fluoroindole derivatives may be related to the difference in the energies of Ar–Cl and Ar–F bonds. The excitation energy of 3-alkyl-4-haloindoles is approximately 98 kcal/mol,[12] which is substantially higher than the energy of an Ar–Cl bond of 86 kcal/mol but is less than that of an Ar–F bond of 125 kcal/mol.[13] Homolysis of 4-fluoroindole derivatives thus cannot occur under photolytic conditions. However, heterolysis in a polar medium may occur at a low quantum efficiency and may be rationalized to proceed via an electrophilic intermediate, E. Although the major pathway of the nonradiative decay is not certain at this moment, it may involve the same reversible protonation as that found in the H–D exchange.[6] The detailed mechanisms of these reactions will be investigated.

Photoaffinity labeling is a powerful method to explore the interaction of substrates and macromolecules including the protein–receptor interaction.[14] Azidotryptophans are known chemical compounds,[15] but they cannot be incorporated into peptides and proteins. Since 4-fluorotryptophan may be incorporated into proteins via the methods of molecular biology,[3] and 4-haloindole derivatives are expected to withstand the experimental conditions of solid-phase peptide synthesis, the photochemistry of 4-haloindoles may provide a unique way of photoaffinity labeling in protein chemistry. The potential applications of these reactions as photoaffinity labels are being actively pursued.

Acknowledgment. We acknowledge the National Science Foundation (Chem-85-40987) and the National Institutes of Health (GM-34957 and HL-19577) for partial support of this work. The NMR instruments used in this investigation were provided by the grants from the National Science Foundation and the National Cancer Institute through the University of Chicago Cancer Research Center.

Supplementary Material Available: The ^{19}F NMR spectrum of a cytosolic retinol binding protein (CRBP-II) labeled with 4-fluorotryptophan prepared by the technique of recombinant DNA from *E. coli* and the ^1H NMR spectra of dimer **2** in CD_2Cl_2 and of dimer **3** in CD_2Cl_2 and in CD_3CN (4 pages). Ordering information is given on any current masthead page.

(1) For examples, see: (a) Ali, S.; Clark, A. J. *J. Mol. Biol.* **1988**, *199*, 415–426. (b) Matarese, V.; Bernhoffer, D. A. *J. Biol. Chem.* **1988**, *263*, 14544–14551. (c) Sacchetini, J. C.; Said, B.; Shulz, H.; Gordon, J. I. *J. Biol. Chem.* **1986**, *261*, 8218–8223. (d) Sundelin, J.; Das, S. R.; Erikson, U.; Rask, L.; Peterson, P. A. *J. Biol. Chem.* **1985**, *260*, 6494–6499.

(2) For reviews on the fluorescence of tryptophan in proteins, see: (a) Beecham, J. M.; Brand, L. *Annu. Rev. Biochem.* **1985**, *54*, 43–71. (b) Creed, D. *Photochem. Photobiol.* **1984**, *39*, 537–562.

(3) Rule, G. S.; Pratt, E. A.; Simplaceanu, V.; Ho, C. *Biochemistry* **1987**, *26*, 549–556.

(4) Li, E., et al., unpublished results.

(5) (a) Bronskill, P. M.; Wong, J. T. *Biochem. J.* **1988**, *249*, 305–308. (b) Yang, N. C.; Locke, B., unpublished results.

(6) Saito, I.; Sugiyama, A.; Yamamoto, Y.; Muramatsu, S.; Matsuura, T. *J. Am. Chem. Soc.* **1984**, *106*, 4286–4287.

(7) Yonemitsu, O.; Cerruti, P.; Witkop, B. *J. Am. Chem. Soc.* **1966**, *88*, 3941–3945.

(8) Evleth, E. M.; Chalvet, O.; Bamiere, P. *J. Phys. Chem.* **1977**, *81*, 1913–1917.

(9) Hansch, J.; Godfrey, J. C. *J. Am. Chem. Soc.* **1951**, *73*, 3518.

(10) Due to experimental difficulties, the quantum yield of this reaction cannot be readily determined, but it is estimated to be much higher than 0.1.

(11) All new products have been characterized by mass spectrometry for molecular formula, high-resolution ^1H NMR, and UV and IR spectroscopy.

(12) This value is estimated from the 0 → 0 band of their absorption spectra.

(13) Benson, S. W. *Thermochemical Kinetics*, 2nd ed.; Wiley: New York, 1976; p 309. Morrison, R. T.; Boyd, R. N. *Organic Chemistry*, 5th ed.; Allyn and Bacon: Boston, 1987; inside cover.

(14) (a) Chowdhry, V.; Westheimer, F. H. *Annu. Rev. Biochem.* **1979**, *48*, 293–325. (b) Bayley, H.; Knowles, J. R. *Methods Enzymol.* **1977**, *46*, 69–114. (c) Tometsko, A. M.; Richards, F. M. *Applications of Photochemistry in Probing Biological Targets*; New York Academy of Sciences: New York, 1980. (d) Eberly, A. N.; De Graan, P. W. E. *Methods Enzymol.* **1985**, *109*, 129–156. (e) Kelly, P. A.; Katoh, M.; Dijane, J.; Sasaki, S. *Ibid.* **1985**, *109*, 156–170. (f) Yip, C. C.; Yeung, C. W. T. *Ibid.* **1985**, *109*, 170–179.

(15) (a) Saito, A.; Relling, H. C. *Prep. Biochem.* **1981**, *11*, 535–546. (b) Melhado, L. L.; Leonard, N. J. *J. Org. Chem.* **1983**, *48*, 5130–5133. (c) Miles, E. W.; Phillips, R. S. *Biochemistry* **1985**, *24*, 4694–4703.