2860 (CH), 2250 (CN), 1780 (C=O), and 1200-1100 cm^{-1} (CF); NMR (CCl_4) ^1H δ 3.98 (s, OCH_3); ^{19}F ϕ -107.6 (t, J_{FF} = 5.7 Hz, 2 F, CF_2), -119.5 (t, J_{FF} = 5.7 Hz, 2 F, CF_2). Anal. ($C_5H_3F_4NO_2$) C, H, N.

Methyl 3-Azido-2,2,3,3-tetrafluoropropionate (5). A 400-mL tube charged with 26.0 g (0.40 mol) of sodium azide, 150 ml of Me_2SO, 33 g (0.75 mol) of carbon dioxide, and 50 g (0.50 mol) of tetrafluoroethylene was agitated for 4 h at 50 °C. Further heating at 100 °C gave no additional pressure drop. The mixture was stirred with 56.7 g (0.45 mol) of dimethyl sulfate until the mildly exothermic reaction had subsided. The volatiles were then removed under reduced pressure and fractionated to give 73.2 g (91%) of **5**: bp 54 °C (50 mm); IR (CCl_4) 3010, 2960 and 2855 (CH), 2160 (N_3), 1785 (C=O), and 1300-1100 cm^{-1} (CF, C-O); NMR (CCl_4) ^1H δ 3.99 (s, OCH_3); ^{19}F ϕ -91.7 (t, J_{FF} = 4.7 Hz, 2 F, CF_2N_3), -120.5 (t, J_{FF} = 4.7 Hz, 2 F, CF_2C=O). Anal. ($C_4H_3F_4N_3O_2$) C, H, N.

Caution: Azidoester **5**, the smallest azide-containing molecule prepared in this work, was handled without incident and appears not to be unduly sensitive to mechanical, thermal, and electrical shock. However, when heated in a closed system to 230-240 °C, **5** detonated *violently*. Appropriate precautions are recommended in handling this and related azides.

Methyl 3-phenoxy-2,2,3,3-tetrafluoropropionate (3); 81% yield; bp 70 °C (1.8 mm); IR (neat) 3070, 3040 (aromatic CH), 2960 (saturated CH), 1780 (C=O), 1590 and 1492 (C=C), 1200-1100 cm^{-1} (CF, C-O); NMR (CCl_4) ^1H δ 7.30 (m, 5 H, aromatic CH), 3.95 (s, 3 H, OCH_3); ^{19}F ϕ -85.8 (t, J_{FF} = 4.6 Hz, 2 F, OCF_2), -121.6 (t, J_{FF} = 4.6 Hz, 2 F, CF_2C=O). Anal. ($C_{10}H_8F_4O_3$) C, H.

Methyl 3-(methylthio)-2,2,3,3-tetrafluoropropionate (4): 67% yield; bp 53 °C (10 mm) [lit.[22] bp 83 °C (50 mm)]. Anal. ($C_5H_6F_4O_2S$) C, H, F.

Methyl 2,2,3,3,6,6,6-heptafluoro-4-oxahexanoate (6): 51% yield; bp 68-69 °C (50 mm); IR (neat) 3020, 2980, 2920, 2870 (CH), 1785 (C=O), 1250-1100 cm^{-1} (CF, C-O); NMR (CCl_4) ^1H δ 4.21 (q, J_{HF} = 7.7 Hz, 2 H, CH_2), 3.84 (s, 3 H, OCH_3); ^{19}F ϕ -75.6 (t of t, J_{HF} = 7.7 Hz, J_{FF} = 2.3 Hz, 3 F, CF_3), -89.9 (t of q, J_{FF} = 4.3, 2.3 Hz, 2 F, CF_2O), -121.7 (t, J_{FF} = 4.3 Hz, 2 F, CF_2C=O). Anal. ($C_6H_5F_7O_3$) C, H, F.

(22) Kimoto et al. (Kimoto, K.; Miyauchi, H.; Ohmura, J.; Ebisawa, M.; Hane, J. U.K. Patent Appl. GB 2051831, 1981) report the preparation of ester **4** by a related route.

3-(Dichlorophosphonyl)-2,2,3,3-tetrafluoropropionyl chloride (7): 11% yield (contaminated with ca. 10% CH_3SCCl_3); bp 59-69 °C (20 mm). For **7**: IR (CCl_4) 1800 (C=O), 1310 (P=O), 1250-1100 cm^{-1} (CF); NMR (CCl_4) ^{19}F ϕ -110.6 (m, 2 F, CF_2C=O), -116.0 (d of t, J_{PF} = 113 Hz, J_{FF} = 2.3 Hz, 2 F, CF_2-P).

Methyl 3-cyano-2-chloro-2,3,3-trifluoropropionate (8): 64% yield; bp 58 °C (50 mm); IR (CCl_4) 3020, 2960, 2950 (CH), 2250 (CN), 1770 (C=O), 1250-1100 cm^{-1} (CF); NMR (CCl_4) ^1H δ 3.95 (s, OCH_3); ^{19}F ϕ -132.9 (t, J_{FF} = 13.6 Hz, 1 F, CF); -100.2, -103.9 (AB m of m, J_{AB} = 287 Hz; A: d., J = 13.6 Hz; B: d, J = 13.6 Hz, 2 F, CF_2). Anal. ($C_5H_3ClF_3NO_2$) C, H, N.

Methyl 3-azido-2-(trifluoromethoxy)-2,3,3-trifluoropropionate (9): 84% yield, bp 38-41 °C (24 mm); IR (CCl_4) 3010, 2960, 2850 (CH), 2150 (N_3), 1780 (C=O), 1250-1100 cm^{-1} (CF); NMR (CCl_4) ^1H δ 3.89 (s, OCH_3); ^{19}F ϕ -56.0 (d, J_{FF} = 8.6 Hz, 3 F, OCF_3), -131.6 (q of t, J_{FF} = 8.6, 5.0 Hz, 1 F, CF); -89.6, -91.8 (AB m of m, J_{AB} = 188 Hz; A: , J_{FF} = 5.0 Hz; B: d, J_{FF} = 5.6 Hz, 2 F, CF_2N_3). Anal. ($C_5H_3F_6N_3O_3$) C, H, N.

Methyl 3-azido-2-(*n*-heptafluoropropoxy)-2,3,3-trifluoropropionate (10): 79% yield; bp 44-45 °C (10 mm); IR (CCl_4) 3010, 2970, 2850 (CH), 2150 (N_3), 1780 (C=O), 1250-1100 cm^{-1} (CF); NMR (CCl_4) ^1H δ 3.92 (s, OCH_3); ^{19}F ϕ -82.1 (t, J_{FF} = 7.3 Hz, 3 F, CF_3), -130.3 (d of d of t, J_{FF} = 19.6, 6, 6 Hz, 1 F, CF), -130.4 (s, 2 F, CF_2); -79.9, -87.8 (AB m of m, J_{AB} = 150 Hz; A: d of q, J_{FF} = 19.6, 7.3 Hz; B: q of d, J_{FF} = 7.3, 6 Hz, 2 F, OCF_2); -87.8, -89.9 (AB m of m, J_{AB} = 189 Hz; A: d, J_{FF} = 6 Hz; B: d, J_{FF} = 6 Hz, 2 F, CF_2N_3). Anal. ($C_7H_3F_{10}N_3O_3$) C, H, N.

Methyl 3-phenoxy-2-(trifluoromethoxy)-2,3,3-trifluoropropionate: 58% yield; bp 63 °C (0.5 mm); IR (neat) 3070 (aromatic CH), 2960 (saturated CH), 1785 (C=O), 1590, 1485 (C=C), 1250-1100 cm^{-1} (CF, C-O); NMR (CCl_4) ^1H δ 7.23 (m, 5 H, aromatic CH), 3.93 (s, 3 H, OCH_3); ^{19}F ϕ -55.8 (d, J_{FF} = 8.6 Hz, 3 F, OCF_3), -132.8 (q of d, J_{FF} = 8.6, 8.2 Hz, 1 F, CF); -83.3, -85.8 (AB m of m, J_{AB} = 138 Hz; B: d, J_{FF} = 8.2 Hz, 2 F, CF_2). Anal. ($C_{11}H_8F_6O_4$) C, H, F.

Registry No. 2b, 86414-22-4; **3**, 91312-77-5; **4**, 77705-91-0; **5**, 86414-01-9; **6**, 91312-78-6; **7**, 86414-15-5; **8**, 86413-99-2; **9**, 86414-03-1; **10**, 86414-05-3; $NCCF_2CF_2^-$, 91312-80-0; $N_3CF_2CF_2^-$, 91312-81-1; $NCCF_2CF(OCF_3)^-$, 91312-82-2; $N_3CF_2CF(OCF_3)^-$, 91312-83-3; methyl 3-phenoxy-2-(trifluoromethoxy)-2,3,3-trifluoropropionate, 91312-79-7; sodium cyanide, 143-33-9; tetrafluoroethylene, 116-14-3; sodium azide, 26628-22-8.

Three-Center (Bifurcated) Hydrogen Bonding in the Crystal Structures of Amino Acids

G. A. Jeffrey* and J. Mitra

Contribution from the Department of Crystallography, University of Pittsburgh, Pittsburgh, Pennsylvania 15260. *Received December 5, 1983*

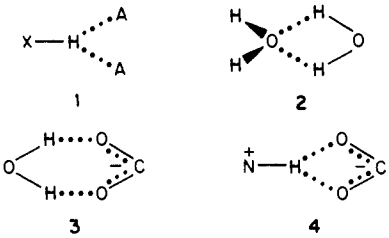
Abstract: The hydrogen bonding of the NH_3^+ groups in the crystal structures of the amino acids falls into three distinct classes. Of the 56 structures examined in this analysis, the distribution is as follows. In 10 structures, the NH_3^+ group forms three two-center (linear) hydrogen bonds, in 25, the NH_3^+ group forms two two-center bonds and one three-center bond, and in 14, the NH_3^+ group forms one two-center bond and two three-center bonds. There were no examples of three three-center bonds being formed. There are three examples of four-center (trifurcated) bonds. Two of the four structures with NH_2^+ and NH^+ groups also had three-center hydrogen bonds. This relatively high proportion of three-center bonds is a consequence of a deficiency in the number of functional protons necessary to satisfy the normal acceptor coordination of the carboxylate oxygens, which is two per oxygen, and the chloride ions, which is four.

Three-centered or bifurcated hydrogen bonding of the type **1** was recognized in one of the first amino acid crystal structures to be determined, that of α-glycine.[1] Although 12 well-documented examples of crystal structures containing three-centered hydrogen bonds of the type **1** were identified in 1968,[2] it is only comparatively recently that the concept has become generally accepted.[3-5] A recent metrical survey of 1509 N—H···O=C hydrogen bonds identified 304 as being three-centered.[4]

(1) Albrecht, G.; Corey, R. B. *J. Am. Chem. Soc.* **1939**, *61*, 1087-1103.
(2) Donohue, J. In "Structural Chemistry and Molecular Biology"; Rich, A., Davidson, N., Eds.; W. H. Freeman: San Francisco, 1968; pp 443-465.
(3) Newton, M. D.; Jeffrey, G. A.; Takagi, S. *J. Am. Chem. Soc.* **1979**, *101*, 1997-2002.
(4) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1984**, *106*, 244-248.
(5) Newton, M. D. *J. Phys. Chem.* **1983**, *87*, 4288-4292.

0002-7863/84/1506-5546$01.50/0

In the first major text on hydrogen bonding,[6] the term is applied to describe the configuration **2**, which is of interest to theoreticians[7]



1 **2**

3 **4**

as a possible structure for the water dimer. This configuration has never been observed in the crystalline state. The term has also been applied to the configuration, **3**, which is uncommon in the crystalline state.[8] A few examples of configuration **4**, which is a special case of **1**, have been noted in crystal structures of the amino acids.[9] This ambiguity in terminology can be avoided by using the descriptor "three-center"[5,9] hydrogen bond for **1** and **4**, to distinguish them from **2** and **3**, in which the hydrogen atom is only involved in two bonds. The term linear hydrogen bond, which is used to describe a configuration that is seldom linear,[10] could then be referred to as a "two-center" hydrogen bond.

Criteria for Three-Center Hydrogen Bonds

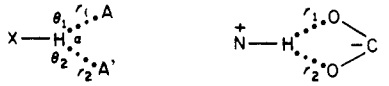
The difficulty of locating hydrogen atoms in pre-diffractometer crystal structure analyses, or for that matter in current X-ray analyses of large molecules, led to the identification of hydrogen bonds in crystals from the X···A distances alone. Since these distances are a function of two distances, the X–H and H···A covalent and hydrogen bond lengths, and the X-H···A angle, which is rarely 180°, such identification is frequently ambiguous. Hydrogen atoms covalently bonded to electronegative atoms, such as oxygens, are well-known to be especially elusive in X-ray structure analyses. It is not surprising therefore that the compelling evidence of three-center hydrogen bonding came first from the neutron diffraction analyses of relatively simple molecules.[11]

Some of the reluctance to accept the concept of three-center bonds may arise from a common misconception of the meaning of the frequently used Pauling van der Waals radii.[12] These radii are derived from the *minimum* interatomic nonbonding distances observed in crystals. They correspond therefore to distances between atoms that are compressed, relative to the minima in the atom-pair potential energy curves. This distinction is particularly important in the derivation of appropriate empirical constants for use in molecular mechanics calculations and is clearly explained in relation to that subject.[13]

Using the criterion for the existence of a hydrogen bond proposed by Hamilton and Ibers,[14] $X-H + H\cdots A < W_X + W_A$, the longest N–H···O and N–H···Cl distances to qualify as hydrogen bonds would be 1.9 and 2.3 Å, using Pauling values for W_X and W_A. From the van der Waals constants derived from estimates of van der Waals potential energy minima, such as those of Allinger,[15] the longest distances are 2.35 and 2.65 Å.

The distribution of two-center hydrogen bond lengths observed by neutron diffraction in the amino acids[9] ranges from 1.58 to 2.05 Å for N–H···O bonds and 1.70 to 1.96 Å for N^+–H···O bonds. For N^+–H···Cl^- bonds, the observed range is much narrower, 2.101–2.161 Å. Thus use of the Pauling van der Waals radii is too restrictive for N–H···O bonds by about 0.2 Å, while use of the Allinger radii appears to be unnecessarily permissive.

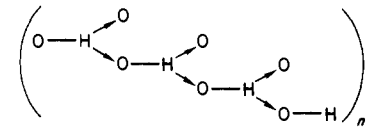
A different criterion is necessary for describing the distribution of three-center bonds. In a study of the geometry of the hydrogen bonding in the neutron diffraction analyses of carbohydrate crystal structures,[16] it was observed that in the configuration **1** the hydrogen atoms lie in, or close to, the plane of the three atoms to which they are bonded; thus in **5**, $\theta_1 + \theta_2 + \alpha \sim 360°$. Similar



5 **6**

configurations are observed in the amino acids,[9] except that the observed geometry is more often unsymmetrical; r_1 and r_2, at the extreme, may differ by 1 Å, but the displacement of the proton from linearity with respect to X–H···A is always such as to imply an attractive force from the more distant acceptor atom. The distribution of r_1 and r_2 in the amino acids in which this planar geometry is observed ranges from 2.119 and 2.364 Å, with $\theta_1 + \theta_2 + \alpha = 359°$ in the most symmetrical case, to 1.710 and 2.711 Å, with $\theta_1 + \theta_2 + \alpha = 359°$ in the most unsymmetrical case. It was interesting, and unexpected, to note that the most unsymmetrical three-centered bonds observed occur in the chelating configuration with the carboxylate group shown in **6**. In α-glycine, L-lysine hydrochloride hydrate, and L-cysteine, r_1 = 1.728, 1.740, and 1.710 Å, while r_2 = 2.648, 2.710, and 2.711 Å. The values of $\theta_1 + \theta_2 + \alpha$ are 358°, 357°, and 359°.[9]

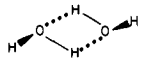
In the study of hydrogen bonding in the neutron diffraction crystal structure analyses of carbohydrates,[16] it was found that about 20% of the O–H···O hydrogen bonds were three centered. This corresponds rather closely with the excess of acceptor oxygen atoms over hydroxyl protons available to form bonds. Studies of the hydrogen-bonding patterns[17-19] in these carbohydrate crystal structures revealed that the dominant theme was the *cooperative effect*.[20,21] This favored the formation of infinite or long finite chains of bonds with a regular donor-to-acceptor direction throughout, i.e., $(\rightarrow O-H\rightarrow O-H\rightarrow O-H\rightarrow O-H\rightarrow)_n$. The role of the three-centered interaction was to include acetal oxygens, which lack protons, in the hydrogen-bonding scheme without disrupting the cooperative effect, as in **7**.



7

(6) Pimental, G. C.; McClelland, A. L. "The Hydrogen Bond"; W. H. Freeman: San Francisco, 1960; p 100.

(7) Umeyana, H.; Morohuma, K. *J. Am. Chem. Soc.* **1977**, *99*, 1316–1332. In this ab initio SCF-MO treatment, three-center bonds of type **2** and cyclic bonds with the configuration



are discussed, but not the more commonly observed configu ration **1**.

(8) Alagona, G.; Ghio, C.; Kollman, P. *J. Am. Chem. Soc.* **1983**, *105*, 5226–5229.

(9) Jeffrey, G. A.; Maluszynska, H. *Int. J. Macromol.* **1982**, *4*, 173–185.

(10) Kroon, J.; Kanters, J. A.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B.; Vliegenthart, J. A. *J. Mol. Struct.* **1975**, *24*, 109–129.

(11) E.g.: α-glyine (Burns, J. H.; Levy, H. A. *American Crystallographic Association Meeting*, June, 1958; Abstr. H9. Jonsson, P. G.; Kvick, A. *Acta Crystallogr., Sect. B* **1972**, *B28*, 1827–1833), methyl α-D-altropyranoside (Poppleton, B. J.; Jeffrey, G. A.; Williams, G. J. B. *Acta Crystallogr., Sect. B* **1975**, *B31*, 2400–2404), methyl α-D-glucopyranoside (Jeffrey, G. A.; McMullan, R. K.; Takagi, S. *Acta Crystallogr., Sect. B* **1977**, *B33*, 728–737).

(12) Pauling, L. "Nature of the Chemical Bond"; Cornell University Press: Ithaca, New York, 1939.

(13) Allinger, N. L. *ACS Monogr.* **1982**, *No.* 177.

(14) Hamilton, W. C.; Ibers, J. A. "Hydrogen Bonding in Solids"; W. A. Benjamin: New York, 1968, p 16.

(15) Allinger, N. L. *Adv. Phys. Org. Chem.* **1976**, *13*, 1–82.

(16) Ceccarelli, C.; Jeffrey, G. A.; Taylor, R. *J. Mol. Struct.* **1981**, *70*, 255–271. The range for O–H(···O)$_2$ bonds is from symmetrical with r_1 and r_2 equal to 2.085 and 2.140 Å and $\theta_1 + \theta_2 + \alpha = 358°$ to unsymmetrical with r_1 and r_2 equal to 1.836 and 2.414 Å and $\theta_1 + \theta_2 + \alpha = 352°$. In a quantum mechanical treatment of three-center O–H(···O)$_2$ bonds, at the SCF.HF/4-31G level of approximating, binding energies of 10.3 and 10.9 kcal/mol were calculated, respectively, for water trimers with a symmetrical bond ($r_1 = r_2$ = 2.15 Å, $\theta_1 = \theta_2 = 135°$, $\alpha = 90°$) and an unsymmetrical bond (r_1 = 1.95, r_2 = 2.50 Å, θ_1 = 150, θ_2 = 110°, α = 100°). These bonds were calculated to be ~1–2 kcal/mol more stable than an equilibrium linear dimer with H···O = 1.882 Å and $\alpha = 180°$.[3]

(17) Jeffrey, G. A.; Lewis, L. *Carbohydr. Res.* **1978**, *60*, 179–182.

(18) Jeffrey, G. A.; Takagi, S. *Acc. Chem. Res.* **1978**, *11*, 264–270.

(19) Jeffrey, G. A.; Mitra, J. *Acta Crystallogr., Sect. B* **1983**, *B39*, 469–480.

(20) Del Bene, J. E.; Pople, J. A. *J. Chem. Phys.* **1970**, *52*, 4858–4866; **1973**, *58*, 3605–3608.

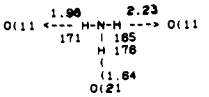
(21) Tse, Y. C.; Newton, M. D. *J. Am. Chem. Soc.* **1977**, *99*, 611–613.

Table I. Patterns of Hydrogen Bonding in Amino Acid Crystal Structures[a]

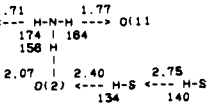
CLASS 1 TYPE : A <---- H-N-H ----> A

A = CARBOXYLATE OR HYDROXYL OXYGENS

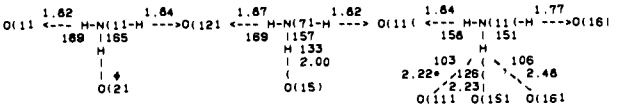
1. O,L-LEUCINE DLLEUC



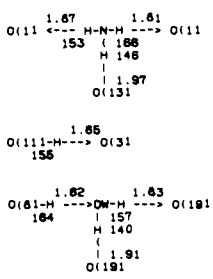
2. L-CYSTEINE (N) LCYSTN12



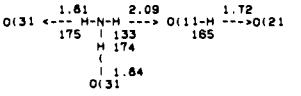
3. L-LYSINE-L-ASPARTATE LYSASP



4. L-GLUTAMIC AC1O-L-PYROGLUTAMIC AC1O H_2O LGPYRG

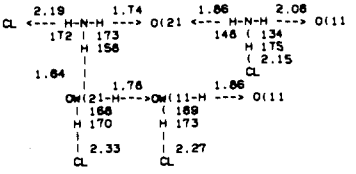


5. O,L-TYROSINE DLTYRS

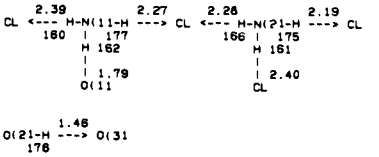


A = CARBOXYLATE OXYGEN,WATER OXYGEN OR CHLORIDE ION

6. L-LYSINE HCL $2H_2O$ (N) LYSCLH11

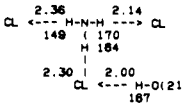


7. *DI*-L-LEUCINE HCL DLEUHC

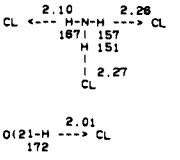


A = CHLORIDE IONS ONLY

8. L-PHENYLALANINE HCL (N1 PHALNC01



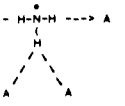
9. L-CYSTINE 2HCL (N1 CYSTCL02



10. OI-L-LEUCINE HCL DLEUHC

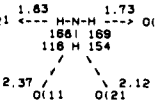
SAME AS 7.

CLASS ((TYPE : A<--- H-N-H ---> A

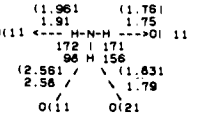


A = OXYGEN ATOMS,CARBOXYLATE,CATION(C,HYDROXYL OR WATER,NITROGEN

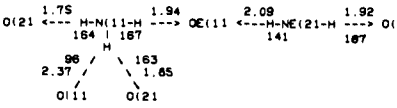
11. A-GLYCINE (N1 GLYC1N01



12. G-GLYCINE 83K (298K1 (N) GLYC1N16



13. L-GLUTAMINE (N1 GLUTAM01



14. O,L-ASPARTIC AC1O DLASPA10

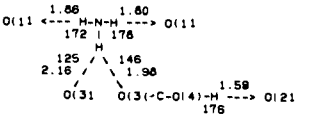
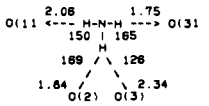
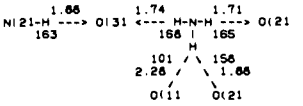


Table I (Continued)

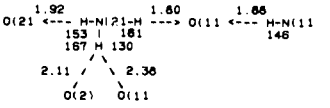
15. D,L ALANYL-L,O-METHIONINE ALAMET10



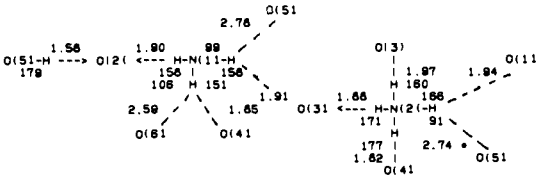
16. GLYCYL-L-LEUCINE GLYLEU10



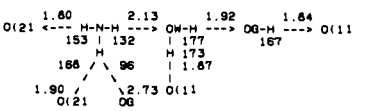
17. L-ALANYL GLYCINE ALAGLY



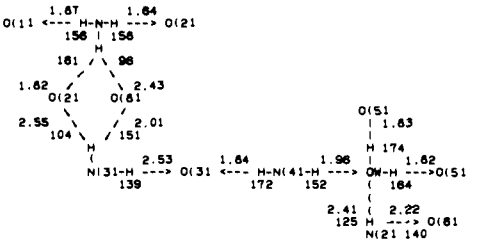
18. AMMONIUM GLYCINIUM SULFATE (N1 AGLYSL01



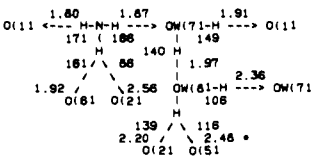
19. L-SERINE H_2O (N1 LSERMH10



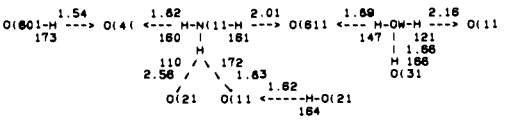
20. L-ARGININE PHOSPHATE H_2O LARGPH01



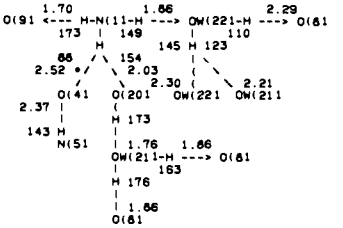
21. POTASSIUM-L-TYROSINE SULFATE $2H_2O$ KTYSUM10



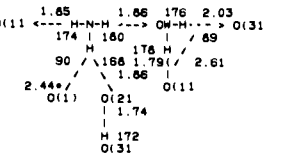
22. GLYCYLGLYCINE PHOSPHATE H_2O GLYGLP



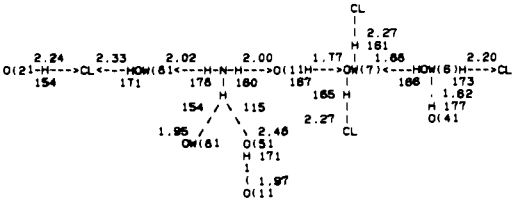
23. GLYCYL-L-TYROSINE $2H_2O$ GLTLYR10



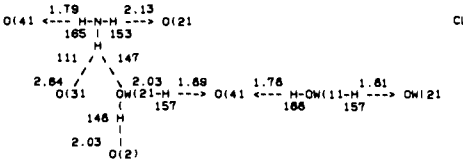
24. THIENYL-O,L-SERINE H_2O THOSER



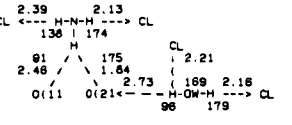
25. O,L-THREO-3,4 -DIHYDROXYPHENYL SERINE HCL $3H_2O$ XPSERC



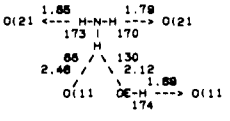
26. CA-DI-L-GLUTAMATE $4H_2O$ LGLUCA



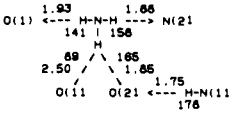
27. CA-L-GLUTAMATE CHLORIDE H_2O CAGLCL10



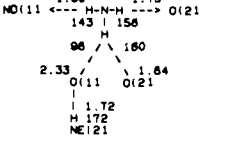
28. L-TYROSINE (N1 LTYROS11



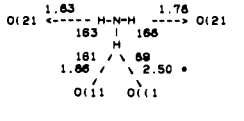
29. D,L-HISTIDINE DLHIST



30. L-HISTIDINE (N1 LHISTD13

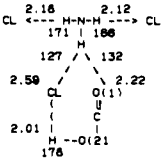


31. L-ALANINE (N1 LALN(N12

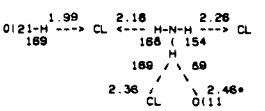


A = CHLORIDE ION OR OXYGEN ATOM

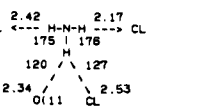
32. GLYCINE HCL (N1 GLYHCL



33. L-VALINE HCL (N) VALEHC11



34. D,L-VALINE HCL DLVALC



35. O-ALLOISOLEUCINE HCL H_2O OALILU10

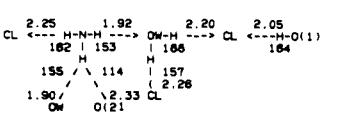
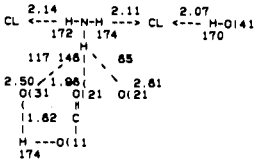


Table I (Continued)

CLASS 11' TYPE :

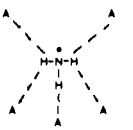
36. L-GLUTAMIC AC1D HCL (N1 LGLUTA



36'. L-LYSINE-L-ASPARTATE LYSASP

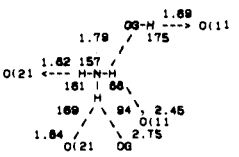
SAME AS 3 .

CLASS 111 TYPE :

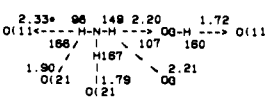


A = CARBOXYLATE OXYGENS

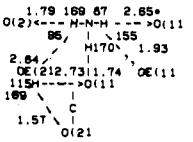
37. O,L-SERINE (N1 DLSERN11



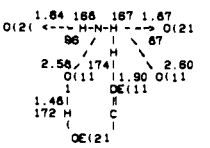
38. L_S - THREONINE (N1 LTHREO01



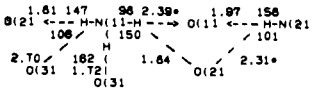
39. L-GLUTAMIC AC1D (A FORM1 (N1 LGLUAC03



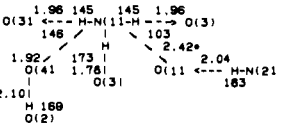
40. L-GLUTAMIC AC1D (B FORM1 (N1 LGLUAC11



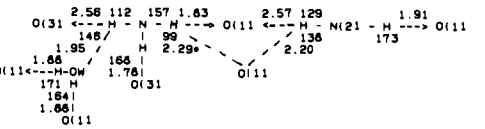
41. A-GLYCYLGLYCINE (N1 GLYGLY04



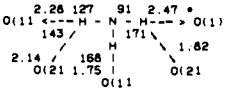
42. L-ALANYL-L-SERINE LALLSE



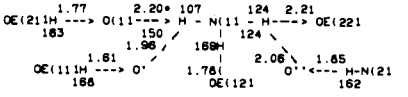
43. L-ASPARAGINE H_2O (N1 ASPARM02



44. B-GLYCINE GLYC(N

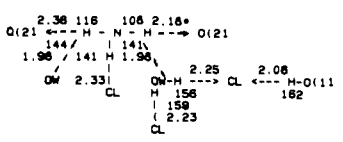


45. A-L-GLUTAMYL-L-GLUTAMIC AC1D

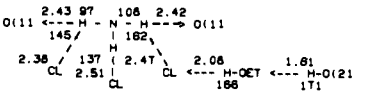


A = CHLORIDE IONS OR CARBOXYLATE OR WATER OXYGENS

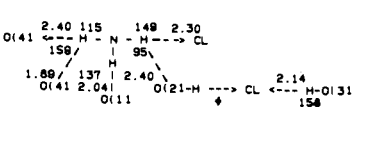
46. L-ISOLEUCINE HCL H_2O LILEUC10



47. L-TYROSINE HCL (N1 LTYRHC10



48. D,L-ASPARTIC AC1D HCL ASPART10



49. GLYCYLGLYCINE HCL H_2O (N1 GLC1CH01

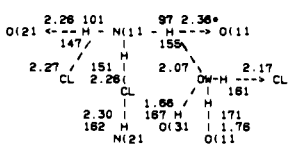
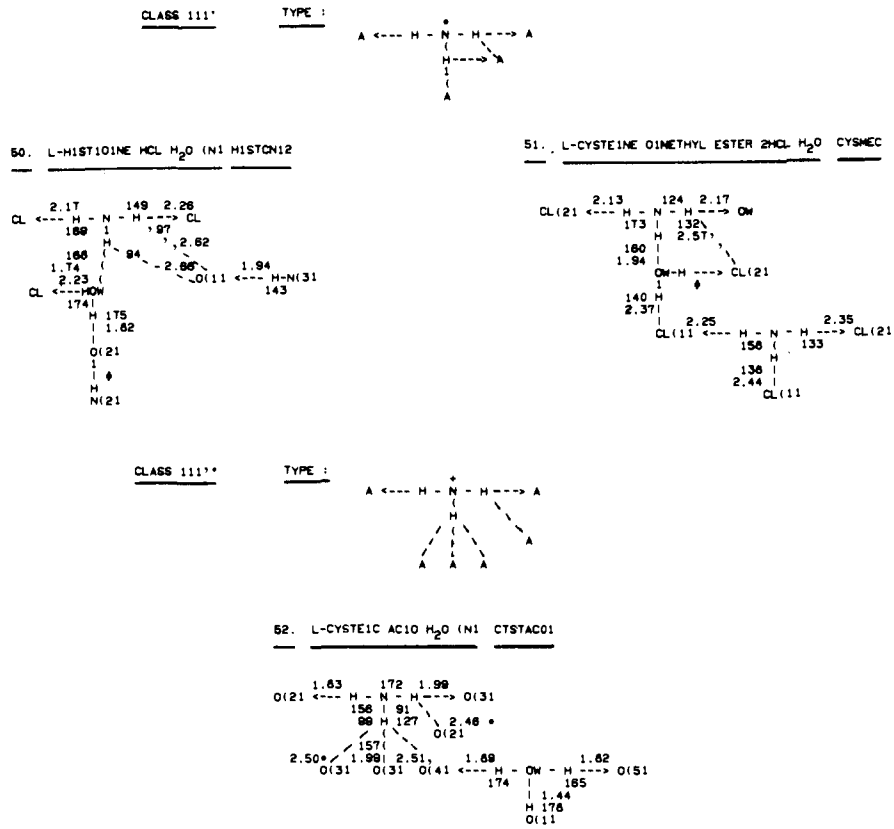


Table I (Continued)



[a] Schematic diagrams of the hydrogen bonding in the amino acid crystal structures. Neutron diffraction analyses are indicated by (N). In the X-ray analyses, the covalent N^+–H distances were normalized to 1.02 Å and the O–H distances to 0.97 Å, by extension, where necessary, in the direction of the X–X bond. The REFCODES are those in the Cambridge Crystallographic Data File.[23] * refers to an intramolecular hydrogen bond. + indicates no distance is reported due to error or omission in the reported hydrogen coordinates.

The metrical survey of hydrogen-bond geometry in 32 neutron diffraction analyses of amino acids[9] revealed an even greater proportion of three-centered bonds than in the carbohydrates. Of the 64 N^+–H···O bonds in this survey, a majority, 46, were three centered. In the amino acids, which are zwitterions in the crystalline state, the cooperative effect must play only a minor role, if any. The principal hydrogen-bond donor group, NH_3^+, cannot, for steric reasons, accept hydrogen bonds, and the principal acceptors, the carboxylate or carbonyl oxygens or halide ions, have no protons for hydrogen-bond donation. As shown in previous surveys,[9,22] carboxylate oxygens tend to accept two hydrogen bonds per acceptor and chloride ions four. Since the principal source of hydrogen-bonding protons is the NH_3^+ group, and there is an equal number of cationic and anionic groups or atoms, there will be an even greater proton deficiency for optimum hydrogen bonding in the amino acids than in the carbohydrates. This qualitatively explains the greater number of three-centered bonds.

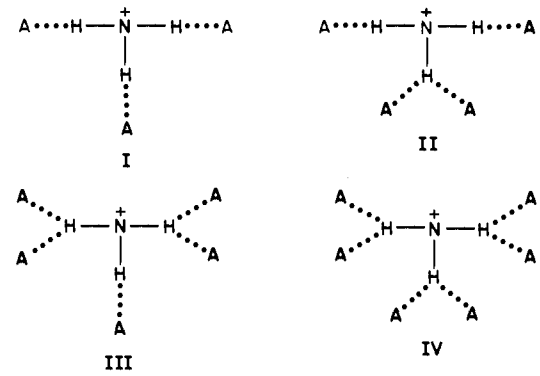
We now discuss this question in more detail by examining the hydrogen-bonding patterns in the vicinity of the NH_3^+ groups in a representative group of amino acid crystal structures.

Amino Acid Hydrogen-Bonding Patterns

Since we are concerned with the recognition of patterns of bonding rather than metrical details, the accuracy of the hydrogen atomic positions is less critical. For this reason we have included in this survey crystal structures determined by X-ray diffraction in addition to the more limited number of neutron diffraction analyses.

It was soon clear that patterns common to significant groups of structures could only be observed in the vicinity of the NH_3^+ groups. This is in contrast to the carbohydrate analysis,[19] for which a meaningful survey could be restricted to molecules containing only the C–OH and C–O–C functional groups. In the amino acids, any significant survey must, by necessity, include a larger variety of hydrogen-bonding functional groups, i.e., NH^+, NH, OH, and CH as donors and O^-, Cl^-, —N═, N<, OH, and OH_2 as acceptors.

If we consider only two- and three-center hydrogen bonds and assume that all the NH_3^+ protons are utilized in hydrogen-bonding, there are only four basic patterns possible within the first NH_3^+ coordination shell. These are I, II, III, and IV. In the 52 crystal



structures examined the distribution was 10 for class I, 25 for class II, 14 for class III, and none for class IV. There were three examples of four-center hydrogen bonds.

The distribution of the amino acid crystal structures into these three classes is given below. The crystal structure analyses are identified by their REFCODES.[23] Schematic diagrams of the

(22) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1983**, *105*, 5761–5766.

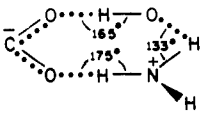
(23) The six-letter codes are the REFCODES in the Cambridge Crystallographic Data File (Allen, F. H.; et al. *Acta Crystallogr., Sect. B* **1979**, *B35*, 2331–2339), from which these structural data were retrieved.

Table II. Classification of Amino Acid Crystal Structures according to Hydrogen-Bonding Types

Class I	
A = Carboxylate or Hydroxyl Oxygens	
D,L-leucine	DLLEUC
L-cysteine (N)	LCYSTN12
L-lysine–L-aspartate (2 examples)	LYSASP
L-glutamic acid–L-pyroglutamic acid hydrate	LGPYRG
D,L-tyrosine	DLTYRS
A = Carboxylate Oxygen, Water Oxygen or Chloride Ion	
L-lysine hydrochloride dihydrate (N) (2 examples)	LYSCLH11
di-L-leucine hydrochloride (1 example)	DLEUHC
A = Chloride Ions Only	
L-phenylalanine hydrochloride (N)	PHALNC01
L-cystine dihydrochloride (N)	CYSTCL02
di-L-leucine hydrochloride (1 example)	DLEUHC
Class II	
A = Oxygen Atoms, Carboxylate, Cationic Hydroxyl or Water, or Nitrogen Atoms	
α-glycine (N)	GLYCIN01
γ-glycine (N)	GLYCIN16
L-glutamine (N)	GLUTAM01
D,L-aspartic acid	DLASPA10
D,L-alanyl-L,D-methionine	ALAMET10
glycyl-L-leucine	GLYLEU10
L-alanyl glycine	ALAGLY
ammonium glycinium sulfate (N)	AGLYSL01
L-serine hydrate (N)	LSERMH10
L-arginine phosphate hydrate	LARGPH01
potassium L-tyrosine sulfate dihydrate	KTYSUH10
glycyl glycine phosphate hydrate	GLYGLP
glycyl-L-tyrosine dihydrate	GLTLYR10
thienyl-D,L-serine hydrate	THDSER
D,L-*threo*-β-(3,4-dihydroxyphenyl)serine hydrochloride trihydrate	XPSERC
calcium di-L-glutamate tetrahydrate	LGLUCA
calcium L-glutamate hydrochloride hydrate	CAGLCL10
L-tyrosine (N)	LTYROS11
D,L-histidine (A = nitrogen)	DLHIST
L-histidine (N) (A = nitrogen)	LHISTD13
L-alanine (N)	LALNIN12
A = Chloride Ion or Oxygen Atom	
glycine hydrochloride (N)	GLYHCL
L-valine hydrochloride	VALEHC11
D,L-valine hydrochloride	DLVALC
D-alloisoleucine hydrochloride hydrate	DALILU10
Class III	
A = Carboxylate Oxygens	
D,L-serine (N)	DLSERN11
L_s-threonine (N)	LTHREO01
L-glutamic acid (A form) (N)	LGLUAC03
L-glutamic acid (B form) (N)	LGLUAC11
α-glycylglycine (N)	GLYCLY04
L-alanyl-L-serine	LALLSE
L-asparagine hydrate (N)	ASPARM02
β-glycine	GLYCIN
α-L-glutamyl-L-glutamic acid[27]	
A = Chloride Ions, Carboxylate, or Water Oxygens	
L-isoleucine hydrochloride hydrate	LILEUC10
L-tyrosine hydrochloride (N)	LTYRHC10
D,L-aspartic acid hydrochloride	ASPART10
glycylglycine hydrochloride hydrate (N)	GLCICN01

hydrogen bonding, which contain the H···A distances and X–H···A angles, are given in Table I. The neutron diffraction analyses are indicated by (N). An interesting overall observation from these data is a definite trend for the formation of stronger two-center bonds when the other hydrogen bonds are three centered. The mean values for the two-center N^+–H···O distances in classes I, II, and III (Table II) are 1.899, 1.880 and 1.789 Å, respectively. The formation of a three-center bond has an apparent deshielding effect on the protons involved in the two-center bonding.

Rather surprisingly, class I, is the smallest class, with 19% of the structures. In the four where the acceptor atoms, A, are carboxylate oxygens, the H···O distances range from 1.64 (in D,L-leucine) to 2.23 Å (also D,L-leucine),[24] with a mean value of 1.90 Å. The N^+–H···O angles range from 133° to 176°, with a mean value of 169°. In D,L-tyrosine, where one of the acceptor oxygens is the hydroxyl, H···O = 2.09 Å, N^+–H···O = 133°. This abnormal angle may be a consequence of the particular closed trimer configuration **8**, which occurs in this structure.



8

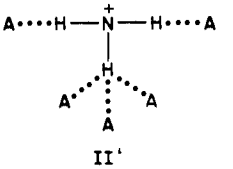
Another unusually "bent" N–H···O angle of 134° occurs in L-lysine hydrochloride hydrate. It is also associated with a weak N–H···O=C bond of 2.08 Å. In di-L-leucine hydrochloride, a very short O–H···O hydrogen bond of 1.46 Å is reported in the dimeric cation. Both of these, however, are X-ray analyses.

In the three examples where the acceptor atoms are chloride ions only, the H···Cl distances range from 2.100 to 2.398 Å, with a mean of 2.255 Å, and the O–H···Cl angles range from 149° to 175°, with a mean of 162°.[25]

Class II is the largest group with 48% of the structures. There were no examples where all the acceptor atoms were chloride ions. The two-center H···O bond lengths range from 1.694 to 2.131 Å with a mean of 1.880 Å. The two-center N–H···O angles range from 132° to 180° with a mean of 160°. The N^+–H···N bond lengths in D,L-histidine and L-histidine are notably short; 1.88, 1.90 Å. The angles are both 158°. Abnormally bent two-center bonds occur in L-glutamine, with N–H···O angle of 167° and H···O = 1.92 Å, L-serine hydrate with N^+–H···O_W angle of 132° and H···O_W = 2.13 Å, L-arginine phosphate hydrate, with N–H···O angle of 139° and H···O = 2.53 Å, glycylglycine phosphate hydrate, with O_W–H···O angle of 121° and H···O = 2.16 Å, and glycycl-L-tyrosine dihydrate, with O_W–H···O angle of 110° and H···O = 2.29 Å. Of these, only the L-glutamine and L-serine hydrate are neutron analyses. When the X–H···A angle is smaller than 140°, the hydrogen bond is always long, i.e., H···A > 1.9 Å, but the reverse is not observed.

The three-center interactions were mostly unsymmetrical with the shorter component ranging from H···O = 1.816 to 2.034 Å and the longer from 2.443 to 2.730 Å.

There are related patterns, with a trifurcated bond, in L-glutamic hydrochloride (N) and L-lysine–L-asparate. In the former, the N^+–H···O distances are 1.98, 2.50, and 2.61 Å; in the latter they are 2.22, 2.23, and 2.48 Å.[26]



II'

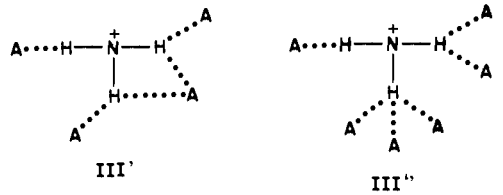
(24) These corrected X-ray values lies beyond the extremes of the distribution of N–H···O distances observed by neutron diffraction; a neutron diffraction refinement of the X-ray crystal structure would be valuable to obtain more reliable values.

(25) This distribution is 0.1 Å wider than that observed using the neutron data only. This corresponds to the average systematic error in H···O distances in N—H···O=C bonds (Taylor, R.; Kennard, O. *Acta Crystallogr., Sect. B* **1983**, *B39*, 133–138).

(26) A similar four-center bond occurs in the crystal structure of sucrose with O–H···O distances of 1.91, 2.26, and 2.52 Å.[9]

(27) Eggleston, D. E.; Hodgson, D. J. *Acta Crystallogr., Sect. B* **1982**, *B38*, 1216–1220.

In addition to the class III examples listed in Table II, there are two related patterns:



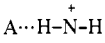
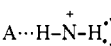
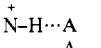
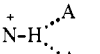
III' III''

where A are chloride or oxygen atoms. III' is observed in L-histidine hydrochloride hydrate (N) (HISTCN12) and L-cysteine dimethyl ester dihydrochloride hydrate (CYSMEC); III'' is observed in, L-cysteic acid hydrate (N) (CYSTAC01). The four-center H···O distances are 2.51, 1.99, and 2.50 Å.

In this group, the two-center bond lengths range from N^+–H···O = 1.72 Å with an angle of 162° in α-glycylglycine to 2.04 Å with an angle of 137° in D,L-aspartic acid HCl (an X-ray analysis). The three-center bonds range from the unsymmetrical in α-glycylglycine with N^+–H···O = 1.806 and 2.701 Å, 1.840 and 2.391 Å, to symmetrical configurations in glycylglycine hydrochloride hydrate with N^+–H···O = 2.069 and 2.360 Å and N^+–H···O and N^+–H···Cl^- = 2.256 and 2.269 Å.

N^+–H_2 and N^+–H Hydrogen Bonding

The data available for examining these cases was more limited. Such as it is, it suggests that both linear and three-center bonding will occur. For example:

	*N*-acetyl-L-glutamine	AGLUAM10
	4-hydroxy-L-proline (N)	HOPROL12
	allo-4-hydroxy-L-proline dihydrate	AHLPRO
	L-prolylsarcosine hydrate	PRSARH
	*N*-acetyl-L-histidine-*N*-methylamide	AHISMA
	*N*-acetylglycine (N)	ACYGLY11

Acknowledgment. This research was supported by the U.S. Public Health Service, Grant GM-24526.

Registry No. DLLEUC, 328-39-2; LCYSTN12, 52-90-4; LYSASP, 20556-18-7; LGPYRG, 91491-98-4; DLTYRS, 556-03-6; LYSCLH11, 39041-33-3; DLEUHC, 81344-48-1; PHALNC01, 17585-69-2; CYSTCL02, 30925-07-6; GLYCIN01, 56-40-6; GLUTAM01, 56-85-9; DLASPA10, 617-45-8; ALAMET10, 1999-43-5; GLYLEU10, 869-19-2; ALAGLY, 687-69-4; AGLYSL01, 14313-00-9; LSERMH10, 41195-60-2; LARGPH01, 91491-99-5; KTYSUH10, 91492-00-1; GLYGLP, 38724-17-3; GLTLYR10, 39630-46-1; THDSER, 91492-01-2; XPSERC, 70384-37-1; LGLUCA, 91492-02-3; CAGLCL10, 91492-03-4; LTYROS11, 60-18-4; DLHIST, 4998-57-6; LHISTD13, 71-00-1; LALNIN12, 56-41-7; GLYHCL, 6000-43-7; VALEHC11, 17498-50-9; DLVALC, 25616-14-2; DALILU10, 53999-01-2; LGLUTA, 138-15-8; DLSERN11, 302-84-1; LTHREO01, 72-19-5; LGLUAC03, 56-86-0; GLYGLY04, 556-50-3; LALLSE, 3303-41-1; ASPARM02, 5794-13-8; LILEUC10, 53999-00-1; LTYRHC10, 16870-43-2; ASPART10, 40149-75-5; GLCICH01, 23273-91-8; HISTCN12, 5934-29-2; CTSTAC01, 23537-25-9; A-L-glutamyl-L-glutamic acid, 3929-61-1.

Direct Observation of Metastable Intermediates in the Photochemical Ring Closure of 2-Naphthyl Vinyl Sulfides

William G. Herkstroeter*† and Arthur G. Schultz‡

Contribution from the Research Laboratories, Eastman Kodak Company, Rochester, New York 14650, and Rensselaer Polytechnic Institute, Troy, New York 12181.
Received October 26, 1983

Abstract: The mechanism for the photochemical cyclization of 2-naphthyl vinyl sulfides was the subject of flash photolysis investigations. Five such sulfides with structural variations each formed phototransient species with absorption maxima near 650 nm. Kinetic analysis of the disappearance of these transients required the presence of two species decaying independently at different rates. We identified these transients as reaction intermediates having zwitterionic thiocarbonyl ylide structures. All of the sulfides showed singlet-state reactivity. Triplet-state reactivity was shown only by those starting molecules whose vinyl substituents were restricted from free rotation in the excited state.

In 1976, one of us presented a detailed study of the photocyclization of 2-naphthyl vinyl sulfides.[1] We provided chemical evidence for the intermediacy of thiocarbonyl ylides in these photoreactions and noted that flash studies designed to detect and characterize reaction intermediates were in progress; these investigations turned out to be more complex than originally anticipated. In the meantime, Wolff reported flash-photolytic studies of some of these same aryl vinyl sulfides.[2] In this paper we present our flash-photolysis results and our interpretation of the reaction mechanism.

In our work, flash-photolysis experiments were initiated with 2-naphthyl 1-indenyl sulfide (**1**). Our previous studies have demonstrated that irradiation of a degassed benzene solution of **1** gives *trans*-dihydrothiophene **3** in 78% isolated yield; on the other hand, irradiation of **1** in the presence of the dipolarophile *N*-phenylmaleimide (NPMI, 2 equiv) results in isolation of a single cycloadduct **4** in 90% yield.[1] A consideration of stereochemistry in **3** and **4** suggests that cyclization of **1** is conrotatory to give the intermediate thiocarbonyl ylide **2** and that hydrogen migration in **2** is suprafacial to give *trans*-dihydrothiophene **3**.

Results and Discussion

Flash excitation of 2-naphthyl 1-indenyl sulfide (**1**) in degassed benzene solution effected transient production with a lifetime exceeding 5 ms. The transient absorption spectrum (Figure 1)

† Eastman Kodak Company.
‡ Rensselaer Polytechnic Institute.

(1) Schultz, A. G.; DeTar, M. B. *J. Am. Chem. Soc.* **1974**, *96*, 296. Schultz, A. G.; DeTar, M. B. *J. Am. Chem. Soc.* **1976**, *98*, 3564. For an account of our involvement in the area of "Photochemical Six-Electron Heterocyclization Reactions", see: Schultz, A. G. *Acc. Chem. Res.* **1983**, *16*, 210. For a more comprehensive review, see: Schultz, A. G.; Motyka, L. "Organic Photochemistry"; Padwa, A., Ed.; Marcel-Dekker: New York, 1983; Vol. 6, p 1.
(2) Wolff, T. *J. Am. Chem. Soc.* **1978**, *100*, 6157.

0002-7863/84/1506-5553$01.50/0