Eisinger, *Proc. Natl. Acad. Sci. U.S.A.*, **64**, 352 (1969); (e) T. G. Scott, R. D. Spencer, N. J. Leonard, and G. Weber, *J. Am. Chem. Soc.*, **92**, 687 (1970); (f) N. J. Leonard and R. F. Lambert, *J. Org. Chem.*, **34**, 3240 (1969); (g) H. Iwamura, N. J. Leonard, and J. Eisinger, *Proc. Natl. Acad. Sci. U.S.A.*, **65**, 1025 (1970); (h) J. A. Secrist III and N. J. Leonard, *J. Am. Chem. Soc.*, **94**, 1702 (1972); (i) M. W. Logue and N. J. Leonard, *ibid.*, **94**, 2842 (1972); (j) J. C. Craig, P. C. Huang, T. G. Scott, and N. J. Leonard, *ibid.*, **94**, 5872 (1972); (k) N. J. Leonard, R. S. McCredie, M. W. Logue, and R. L. Cundall, *ibid.*, **95**, 2320 (1973); (l) N. J. Leonard and K. Ito, *ibid.*, **95**, 4010 (1973); (m) N. J. Leonard and R. L. Cundall, *ibid.*, **96**, 5904 (1974); (n) K. Mutai, B. A. Gruber, and N. J. Leonard, *ibid.*, **97**, 4095 (1975).
(40) K. Shimada, Y. Nagase, and U. Matsumoto, *Yakugaku Zasshi*, **89**, 436 (1969).
(41) D. B. McCormick and J. A. Roth, *Anal. Biochem.*, **34**, 226 (1970).
(42) J. B. Brown, H. B. Henbest, and E. R. H. Jones, *J. Chem. Soc.*, 3172 (1952).
(43) V. duVigneaud, K. Hofmann, D. B. Melville, and P. György, *J. Biol. Chem.*, **140**, 643 (1941).
(44) L. H. Sternbach, *Compr. Biochem.*, **11**, 78 (1963).
(45) M. W. Goldberg and L. H. Sternbach, U.S. Patent 2 489 237 (Nov 22, 1949); *Chem. Abstr.*, **45**, 185*g* (1951).
(46) D. B. Melville, *J. Biol. Chem.*, **208**, 495 (1954).
(47) E. L. Jackson, *Org. React.*, **2**, 341 (1944).
(48) N. M. Green, *Biochem. J.*, **94**, 23C (1965).
(49) N. M. Green, *Methods Enzymol.*, **18A**, 418 (1970).
(50) D. A. Chignell and W. B. Gratzer, *J. Phys. Chem.*, **72**, 2934 (1968).
(51) E. H. Strickland, J. Horwitz, and C. Billups, *Biochemistry*, **9**, 4914 (1970).
(52) E. H. Strickland, C. Billups, and E. Kay, *Biochemistry*, **11**, 3657 (1972).
(53) L. J. Andrews and L. S. Forster, *Biochemistry*, **11**, 1875 (1972).
(54) L. J. Andrews and L. S. Forster, *Photochem. Photobiol.*, **19**, 353 (1974).
(55) G. Barth, R. E. Linder, E. Bunnenberg, and C. Djerassi, *Helv. Chim. Acta*, **55**, 2168 (1972).
(56) Y. Yamomoto and J. Tanaka, *Bull. Chem. Soc. Jpn.*, **45**, 1362 (1972).
(57) B. Valeur and G. Weber, *Photochem. Photobiol.*, **25**, 441 (1977).
(58) G. Weber, *Biochem. J.*, **75**, 335 (1960).
(59) B. L. Van Duuren, *J. Org. Chem.*, **26**, 2954 (1961).
(60) E. Lippert, W. Luder, and H. Boos, *Adv. Mol. Spectrosc.*, **1**, 443 (1962).
(61) N. Matago, Y. Torihashi, and K. Ezumi, *Theor. Chim. Acta*, **2**, 158 (1964).
(62) M. S. Walker, T. W. Bednar, and R. Lumry, *J. Chem. Phys.*, **45**, 3455 (1966).
(63) M. S. Walker, T. W. Bednar, and R. Lumry, *J. Chem. Phys.*, **47**, 1020 (1967).
(64) J. Eisinger and G. Navon, *J. Chem. Phys.*, **50**, 2069 (1969).
(65) E. V. Donckt, *Bull. Soc. Chim. Belg.*, **78**, 69 (1969).
(66) P. S. Song and W. E. Kurtin, *J. Am. Chem. Soc.*, **91**, 4892 (1969).
(67) M. Sun and P. S. Song, *Photochem. Photobiol.*, **25**, 3 (1977).
(68) C. T. Chang, C. Y. Wu, A. R. Muirhead, and J. R. Lombardi, *Photochem. Photobiol.*, **19**, 347 (1974).
(69) J. Longworth, *Photochem. Photobiol.*, **7**, 587 (1968).
(70) G. T. DeTitta, J. W. Edmonds, W. Stallings, and J. Donohue, *J. Am. Chem. Soc.*, **98**, 1920 (1976).
(71) R. Lett and A. Marquet, *Tetrahedron Lett.*, 2851, 2855 (1971).
(72) R. Lett and A. Marquet, *Tetrahedron*, **30**, 3365, 3379 (1974), and references cited therein.
(73) G. C. Levy and G. L. Nelson, "Carbon-13 Nuclear Magnetic Resonance for Organic Chemists", Wiley-Interscience, New York, N.Y., 1972, e.g., p 42.
(74) C. E. Bowen, E. Rauscher, and L. L. Ingraham, *Arch. Biochem. Biophys.*, **125**, 865 (1968).
(75) The ^{13}C NMR spectrum of biotin methyl ester has not been reported previously. In the reported spectrum for 1-*N*-methoxycarbonylbiotin methyl ester, there appear to be several questionable carbon assignments; see R. B. Guchhait, S. E. Polakis, D. Hollis, C. Fenslau, and M. D. Lane, *J. Biol. Chem.*, **249**, 6646 (1974). A paper on the ^{13}C NMR analysis of biotin has come to our attention after we submitted our manuscript (S. P. Singh, U. I. Stenberg, S. A. Farnum, and S. S. Parmar, *Spectrosc. Lett.*, **11**, 259 (1978)). The carbon assignments of Singh et al., which were made mainly on the basis of reported chemical shifts of simpler analogues, are in agreement with ours for the ester with the exception of those for C-7, -8, and -9 (our numbering).
(76) D. M. Frieze, P. F. Hughes, R. L. Merrill, and S. A. Evans, Jr., *J. Org. Chem.*, **42**, 2206 (1977), and references cited therein.
(77) G. W. Buchanan and T. Durst, *Tetrahedron Lett.*, 1683 (1975).
(78) N. J. Leonard and C. R. Johnson, *J. Org. Chem.*, **27**, 282 (1962).
(79) L. J. Dolby and D. L. Booth, *J. Am. Chem. Soc.*, **88**, 1049 (1966).
(80) R. Lett and A. Marquet[72] reported a ratio of 3:2.
(81) B. Witkop, *Adv. Protein Chem.*, **16**, 221 (1961), and references cited therein.
(82) W. Tagaki, K. Kikukawa, K. Ando, and S. Oae, *Chem. Ind. (London)*, 1624 (1964).
(83) R. Harville and S. F. Reed, Jr., *J. Org. Chem.*, **33**, 3976 (1968).
(84) C. R. Johnson and J. J. Rigau, *J. Am. Chem. Soc.*, **91**, 5398 (1969).
(85) C. R. Johnson and M. P. Jones, *J. Org. Chem.*, **32**, 2014 (1967).
(86) C. Walling and M. J. Mintz, *J. Org. Chem.*, **32**, 1286 (1967).
(87) L. Skattebøl, B. Boulette, and S. Solomon, *J. Org. Chem.*, **32**, 3111 (1967).
(88) J. Klein and H. Stollar, *Tetrahedron*, **30**, 2541 (1974).
(89) Autoclaving for longer than 15 min was necessary for denaturation of the complex. Cf. ref 1; C. H. Pal and H. C. Lichstein, *Proc. Soc. Exp. Biol. Med.*, **116**, 197 (1964); R. D. Wei and L. D. Wright, *ibid.*, **117**, 341 (1964).

Avidin–Biotin Interaction. Crystal and Molecular Structures of Two Linked Models[1]

William F. Paton, Fu-Tong Liu, and Iain C. Paul*

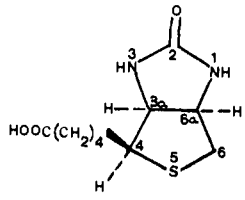
Contribution from the Department of Chemistry, School of Chemical Sciences, University of Illinois, Urbana, Illinois 61801. *Received August 30, 1978*

Abstract: The nature of the strong interaction between biotin (**1**) and the glycoprotein avidin has been examined by X-ray analysis of two model compounds containing the biotin ring system separated by a trimethylene bridge from an indole ring that represents the tryptophan molecules present in the binding site of avidin. Crystals of Biot-C_3-Ind (**2**) hemihydrate are orthorhombic ($P2_12_12_1$) with $a = 46.229$ (13), $b = 7.533$ (2), and $c = 9.012$ (2) Å, and the structure containing two molecules (**2** and **2'**) in the asymmetric unit was refined to an R factor of 0.067 on 2285 nonzero reflections. The crystals of Biot(SO)-C_3-Ind (**6**) are also orthorhombic ($P2_12_12_1$) with $a = 10.634$ (3), $b = 13.641$ (2), and $c = 10.847$ (3) Å, and the structure has been refined to an R factor of 0.045 on 1595 nonzero reflections. One of the molecules (**2**) has a gauche arrangement around the central bond of the trimethylene bridge, while **2'** and **6** have a fully extended arrangement. There is no evidence for a significant intramolecular interaction between the biotin and the indole rings in any of the three molecules, nor is there any evidence for a strong intermolecular hydrogen bonding interaction between the rings. In the crystal of Biot-C_3-Ind, one of the molecules, **2'**, fits snugly into a hydrophobic V-shaped crevice formed by two indole rings. It is postulated that such an interaction may be significant in the avidin–biotin complex.

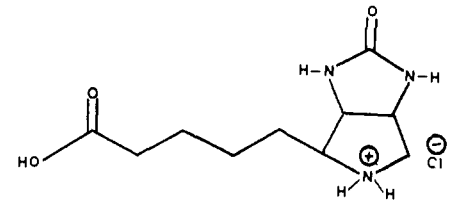
Introduction

The nature of the binding of biotin (**1**) and avidin, which together form one of the strongest biological complexes known, has not been established. It has been concluded, however, that spectroscopic changes and the protection of tryptophans from oxidation in the glycoprotein avidin which are brought about by complex formation with biotin are due to interaction between biotin and the tryptophans of each subunit of avidin.[2]

0002-7863/79/1501-1005$01.00/0

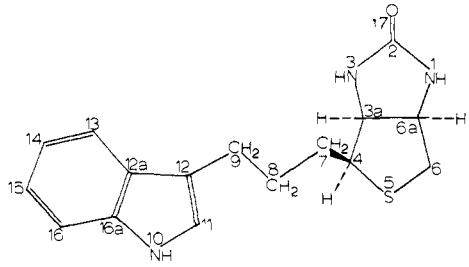


1



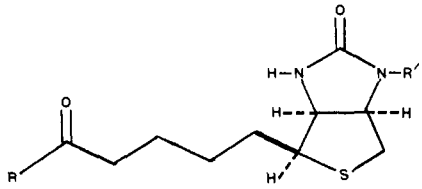
5

In order to identify any biotin-tryptophan interaction that might be contributory to the very strong binding observed in the avidin-biotin complex, $K_D \approx 10^{-15}$ M, model compounds were synthesized in this laboratory[3] having trimethylene (**2**) and tetramethylene chains between the biotin ring system and indole. The name of the former compound is abbreviated as Biot-C_3-Ind to indicate the essential features of the model.[4] Its structure and stereochemistry are established by its method of synthesis from the precursor (+)-biotin. In this model, the trimethylene bridge, which has been used to advantage in earlier photochemical[5] and structural[6] studies of interactions, allows the existence of folded conformations in which the indole can lie under, or endo to, the fused rings of the biotin moiety. Evidence has been gathered from ultraviolet absorption, fluorescence emission, and NMR spectroscopy for the existence of a weak intramolecular interaction between the two ring systems of Biot-C_3-Ind (**2**) in dilute solutions in hydrophobic

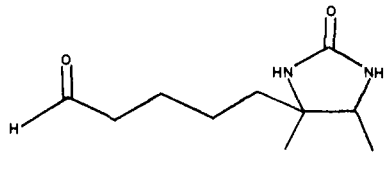


2

solvents.[3] A determination of the X-ray structure of compound **2** was undertaken to delineate potential intra- or intermolecular proximity, or association, of biotin with indole, and, by analogy, of biotin with the tryptophan units of avidin. The structure of Biot-C_3-Ind (**2**) may also be related to the X-ray structures of biotin (**1**)[7,8] and several derivatives of biotin (**3**,[9] **4**,[10] **5**[11]) re-



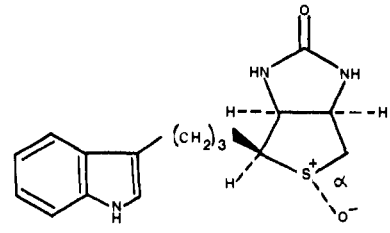
3 R = NHC_6H_4Br R' = $CONHC_6H_4Br$



4

ported previously. A prominent feature of all these structures is the longer than usual ureido carbonyl bond and the corresponding shortening of the carbonyl carbon-nitrogen bonds of the ureido ring. This feature has been attributed[8] to partial delocalization of electronic charge within the ureido group in response to participation of the carbonyl oxygen in a strong hydrogen bond.

We have also determined the X-ray structure of the major oxidation product of Biot-C_3-Ind (**2**) with hydrogen peroxide in glacial acetic acid, namely, the α-sulfoxide **6**. The stereo-



6

chemistry of this sulfoxide, which had been assigned on the basis of comparison of NMR spectra of Biot-C_3-Ind and its two diastereomeric sulfoxides,[3] was confirmed by this analysis. Moreover, a combination of the detailed structure obtained by X-ray analysis with the ^1H and ^{13}C NMR correlations for α-Biot(SO)-C_3-Ind (**6**)[3] would provide the basis for future structural and stereochemical assignments in asymmetric sulfoxides. The related α-sulfoxide of (+)-biotin has assumed importance because of the predominant formation of this diastereomer from NBS oxidation of bound biotin in the avidin-biotin complex.

Experimental Section

Crystal Data for 2, [3a*S*-(3aα,4β,6aα)]4-[3-(Indol-3-yl)propyl]-hexahydro-2-oxo-1*H*-thieno[3,4-*d*]imidazole (Biot-C_3-Ind) Hemihydrate.[3] Anal. Calcd for $C_{16}H_{19}N_3OS \cdot \frac{1}{2}H_2O$: C, 61.91; H, 6.49; N, 13.54. Found: C, 61.94; H, 6.30; N, 13.33. Crystallized from aqueous ethanol; mol wt 310.5, orthorhombic, a = 46.229 (13) Å, b = 7.533 (2) Å, c = 9.012 (2) Å, V = 3138.4 Å^3, Z = 8, ρ_c = 1.31 g cm^{-3}, μ(Cu Kα) = 18.4 cm^{-1}, F(000) = 1320, systematic absences for $h00$ when $h = 2n + 1$, for $0k0$ when $k = 2n + 1$, and for $00l$ when $l = 2n + 1$ establish the space group as $P2_12_12_1$. The cell dimensions were obtained by a least-squares fit to the settings for 15 reflections on a Syntex $P2_1$ diffractometer equipped with a graphite monochromator, λ(Cu Kα) = 1.541 78 Å.

Solution and Refinement of the Structure of 2. A crystal with dimensions ca. 0.10 × 0.05 × 0.03 mm was used for data collection. The data collection was performed in the 2θ-θ scan mode. The variable scan option was employed (2.0 to 12.0° min^{-1}) with the total background time equal to the scan time. Three standards were monitored every 57 reflections. Examination of these reflections showed no crystal deterioration. The hkl octant was collected out to 2θ = 140° (sin θ/λ = 0.609). Out of the possible 3387 unique reflections collected, 2285 were considered significant at the 2σ level based on counting statistics. The data were corrected for Lorentz and polarization effects, but not for absorption; the minimum and maximum transmission factors were estimated to be 0.95 to 0.91.

The structure was solved by direct methods using the version of the MULTAN program[12] supplied with the Syntex EXTL system.[13] Full-matrix, least-squares refinement of positional and anisotropic

Table I. Final Atomic Coordinates for the Two Molecules of **2**

atom	*x*	*y*	*z*	atom	*x*	*y*	*z*
N(1)	0.67363 (14)	0.6116 (9)	0.9095 (10)	O(W)	0.60121 (10)	0.2050 (8)	0.8665 (6)
C(2)	0.66079 (16)	0.4616 (12)	0.9616 (10)	H(1)[a]	0.667 (1)	0.724 (9)	0.921 (9)
N(3)	0.67782 (13)	0.3186 (9)	0.9324 (9)	H(3)	0.670 (1)	0.186 (9)	0.930 (8)
C(3A)	0.70360 (16)	0.3677 (10)	0.8461 (9)	H(3A)	0.703 (0)[b]	0.309 (0)	0.747 (0)
C(4)	0.73232 (14)	0.3119 (10)	0.9202 (9)	H(4)	0.749 (0)	0.315 (0)	0.847 (0)
S(5)	0.73808 (4)	0.4710 (3)	1.0701 (3)	HA(6)	0.719 (0)	0.751 (0)	1.008 (0)
C(6)	0.72665 (15)	0.6497 (9)	0.9484 (10)	HB(6)	0.743 (0)	0.692 (0)	0.887 (0)
C(6A)	0.70243 (16)	0.5753 (10)	0.8459 (9)	H(6A)	0.706 (0)	0.630 (0)	0.746 (0)
C(7)	0.73255 (14)	0.1175 (9)	0.9748 (8)	HA(7)	0.716 (0)	0.101 (0)	1.047 (0)
C(8)	0.76107 (15)	0.0642 (9)	1.0519 (9)	HB(7)	0.729 (0)	0.038 (0)	0.888 (0)
C(9)	0.78849 (15)	0.0942 (11)	0.9551 (10)	HA(8)	0.763 (0)	0.136 (0)	1.144 (0)
N(10)	0.85147 (16)	−0.0888 (13)	1.1626 (8)	HB(8)	0.760 (0)	−0.064 (0)	1.078 (0)
C(11)	0.83194 (16)	0.0529 (12)	1.1299 (10)	HA(9)	0.785 (0)	0.052 (0)	0.852 (0)
C(12)	0.81420 (15)	−0.0053 (11)	1.0170 (9)	HB(9)	0.793 (0)	0.224 (0)	0.953 (0)
C(12A)	0.82255 (15)	−0.1809 (12)	0.9812 (9)	H(10)	0.856 (2)	−0.072 (14)	1.243 (8)
C(13)	0.81316 (17)	−0.3006 (13)	0.8758 (10)	H(11)	0.831 (0)	0.172 (0)	1.179 (0)
C(14)	0.82623 (20)	−0.4670 (15)	0.8649 (11)	H(13)	0.797 (0)	−0.267 (0)	0.808 (0)
C(15)	0.84836 (20)	−0.5127 (13)	0.9629 (12)	H(14)	0.820 (0)	−0.551 (0)	0.786 (0)
C(16)	0.85836 (17)	−0.3977 (13)	1.0678 (12)	H(15)	0.857 (0)	−0.635 (0)	0.956 (0)
C(16A)	0.84567 (16)	−0.2271 (11)	1.0746 (10)	H(16)	0.874 (0)	−0.432 (0)	1.138 (0)
O(17)	0.63630 (9)	0.4531 (7)	1.0193 (6)	H(1′)	0.366 (2)	0.700 (9)	0.317 (9)
N(1′)	0.37013 (15)	0.5806 (10)	0.3026 (10)	H(3′)	0.375 (2)	0.222 (9)	0.414 (9)
C(2′)	0.36229 (17)	0.4507 (12)	0.3954 (10)	H(3A′)	0.398 (0)	0.244 (0)	0.170 (0)
N(3′)	0.37988 (15)	0.3099 (9)	0.3703 (10)	H(4′)	0.445 (0)	0.295 (0)	0.217 (0)
C(3A′)	0.40018 (16)	0.3382 (11)	0.2478 (9)	HA(6′)	0.414 (0)	0.776 (0)	0.228 (0)
C(4′)	0.43259 (15)	0.3329 (11)	0.2995 (9)	HB(6′)	0.431 (0)	0.648 (0)	0.114 (0)
S(5′)	0.44053 (5)	0.5610 (3)	0.3608 (3)	H(6A′)	0.386 (0)	0.540 (0)	0.090 (0)
C(6′)	0.41988 (17)	0.6514 (12)	0.2076 (10)	HA(7′)	0.428 (0)	0.223 (0)	0.507 (0)
C(6A′)	0.39257 (17)	0.5303 (12)	0.1959 (9)	HB(7′)	0.433 (0)	0.075 (0)	0.376 (0)
C(7′)	0.43941 (16)	0.1936 (11)	0.4150 (10)	HA(8′)	0.483 (0)	0.162 (0)	0.358 (0)
C(8′)	0.47133 (16)	0.1825 (12)	0.4525 (10)	HB(8′)	0.478 (0)	0.297 (0)	0.498 (0)
C(9′)	0.47782 (16)	0.0317 (12)	0.5599 (11)	HA(9′)	0.467 (0)	0.056 (0)	0.655 (0)
N(10′)	0.55386 (17)	0.0440 (15)	0.6887 (10)	HB(9′)	0.471 (0)	−0.081 (0)	0.515 (0)
C(11′)	0.52707 (20)	0.1228 (12)	0.6659 (11)	H(10′)	0.569 (1)	0.078 (12)	0.723 (9)
C(12′)	0.50899 (15)	0.0059 (11)	0.5982 (11)	H(11′)	0.522 (0)	0.249 (0)	0.692 (0)
C(12A′)	0.52581 (16)	−0.1513 (12)	0.5733 (11)	H(13′)	0.500 (0)	−0.335 (0)	0.466 (0)
C(13′)	0.51938 (16)	−0.3127 (14)	0.5090 (11)	H(14′)	0.536 (0)	−0.565 (0)	0.456 (0)
C(14′)	0.54034 (20)	−0.4482 (14)	0.5030 (12)	H(15′)	0.582 (0)	−0.512 (0)	0.565 (0)
C(15′)	0.56758 (20)	−0.4146 (16)	0.5653 (17)	H(16′)	0.595 (0)	−0.234 (0)	0.666 (0)
C(16′)	0.57489 (20)	−0.2534 (17)	0.6269 (14)	HA(W)	0.610 (0)	0.286 (0)	0.938 (0)
C(16A′)	0.55350 (18)	−0.1239 (14)	0.6294 (12)	HB(W)	0.593 (0)	0.127 (0)	0.933 (0)
O(17′)	0.34313 (10)	0.4586 (8)	0.4904 (7)				

[a] The hydrogen atoms have been assigned numbers corresponding to the atom to which they are bonded. [b] If the standard deviation is reported as zero, the parameter was not varied in the least-squares refinement.

thermal parameters for the nonhydrogen atoms converged with values for R and R_w of 0.105 and 0.076, respectively.[14] The hydrogen atoms bonded to nitrogen and oxygen were located from a difference map; the hydrogen atoms bonded to carbon were placed at ideal positions with C–H distances of 1.0 Å. The hydrogen atoms were assigned isotropic thermal parameters 10% larger than the equivalent isotropic B of the atoms to which they are bonded. Continued full-matrix, least-squares refinement of the positional and anisotropic thermal parameters for the nonhydrogen atoms and of the positional parameters for the hydrogen atoms bonded to nitrogen and oxygen converged with values for R and R_w of 0.068 and 0.044, respectively. However, the O–H bond lengths and H–O–H bond angles were chemically unreasonable. Subsequently, the hydrogen atoms of the water molecule were removed from the structure factor calculations and relocated from a difference map. In the final cycles of full-matrix, least-squares refinement, the positional and anisotropic thermal parameters of the nonhydrogen atoms and the positional parameters of the hydrogen atoms bonded to nitrogen were refined. The hydrogen atoms bonded to carbon and oxygen were included in the structure factor calculations but their parameters were not refined. The final values of R and R_w are 0.067 and 0.044, respectively; the final value of $[\Sigma w(|F_o| - |F_c|)^2/(m-n)]^{1/2}$, where m is the number of observations and n is the number of variables, is 1.75. The scattering curves were taken from the analytical expression given in the "International Tables for X-ray Crystallography".[15] The highest residual electron density in the final difference map is 0.27 Å^{-3}. The final values for the atomic coordinates are given in Table I. The thermal parameters and the structure factors have been deposited.[16]

Crystal Data for 6, [3a*S*-(3aα,4β,6aα)]4-[3-(Indol-3-yl)propyl]-hexahydro-2,5-α-dioxo-1*H*-thieno[3,4-*d*]imidazole (α-Biot(SO)-C$_3$-Ind).[3] $C_{16}H_{19}N_3O_2S$, mol wt 317.3, orthorhombic, a = 10.634 (3) Å, b = 13.641 (2) Å, c = 10.847 (3) Å, V = 1573.4 Å^3, Z = 4, ρ_c = 1.34 g cm^{-3}, μ(Cu Kα) = 18.7 cm^{-1}, F(000) = 672, systematic absences for $h00$ when $h = 2n+1$, for $0k0$ when $k = 2n+1$, and for $00l$ when $l = 2n+1$ establish the space group as $P2_12_12_1$. The cell dimensions were obtained by a least-squares fit to the settings for 15 reflections on a Syntex $P2_1$ diffractometer equipped with a graphite monochromator, λ(Cu Kα) = 1.541 78 Å.

Solution and Refinement of the Structure of 6. A crystal with dimensions ca. 0.20 × 0.10 × 0.08 mm was used for data collection. The data collection was performed as described above with a 2θ scan rate of 2.0 to 29.3° min^{-1}. Three standards were monitored every 47 reflections. The $\bar{h}kl$ octant was collected out to 2θ = 140° (sin θ/λ = 0.609). Out of the possible 1697 unique reflections collected, 1595 were above zero at the 2σ significance level based on counting statistics. The data were corrected for Lorentz and polarization effects, but not for absorption; the minimum and maximum transmission factors were estimated to be 0.86 to 0.83.

The structure was solved by direct methods using the programs described above. All of the hydrogen atoms were located from a difference map. Full-matrix, least-squares refinement of positional and anisotropic thermal parameters for the nonhydrogen atoms and of

Table II. Final Atomic Coordinates for **6**

atom	x	y	z
N(1)	0.4341 (4)	0.7527 (3)	0.6437 (4)
C(2)	0.3299 (5)	0.7469 (3)	0.5710 (4)
N(3)	0.3254 (4)	0.6572 (2)	0.5207 (4)
C(3A)	0.4422 (4)	0.6039 (3)	0.5351 (4)
C(4)	0.4234 (4)	0.5001 (3)	0.5850 (4)
S(5)	0.3861 (1)	0.5212 (1)	0.7474 (1)
C(6)	0.5165 (5)	0.6068 (3)	0.7533 (6)
C(6A)	0.5113 (4)	0.6664 (3)	0.6341 (5)
C(7)	0.3274 (5)	0.4396 (3)	0.5159 (4)
C(8)	0.3161 (5)	0.3337 (3)	0.5623 (5)
C(9)	0.2209 (7)	0.2761 (4)	0.4882 (6)
N(10)	0.2106 (4)	0.0149 (3)	0.5744 (5)
C(11)	0.2486 (5)	0.0894 (4)	0.5040 (6)
C(12)	0.1979 (4)	0.1747 (3)	0.5409 (4)
C(12A)	0.1241 (4)	0.1524 (3)	0.6448 (4)
C(13)	0.0493 (6)	0.2076 (5)	0.7240 (6)
C(14)	−0.0130 (7)	0.1626 (9)	0.8166 (7)
C(15)	−0.0033 (8)	0.0579 (9)	0.8339 (7)
C(16)	0.0673 (7)	0.0048 (6)	0.7588 (8)
C(16A)	0.1338 (4)	0.0512 (3)	0.6639 (4)
O(17)	0.2513 (3)	0.8125 (2)	0.5577 (4)
O(18)	0.4182 (4)	0.4325 (2)	0.8224 (3)
HN(1)	0.455 (4)	0.817 (4)	0.675 (4)
HN(3)	0.290 (4)	0.647 (3)	0.454 (4)
HC(3A)	0.480 (3)	0.599 (3)	0.461 (3)
HC(4)	0.493 (4)	0.474 (3)	0.599 (4)
HC(6′)	0.499 (4)	0.649 (3)	0.822 (4)
HC(6″)	0.594 (5)	0.564 (4)	0.766 (5)
HC(6A)	0.597 (5)	0.682 (4)	0.626 (5)
HC(7′)	0.355 (3)	0.444 (3)	0.423 (3)
HC(7″)	0.250 (4)	0.477 (3)	0.531 (4)
HC(8′)	0.287 (4)	0.331 (3)	0.657 (5)
HC(8″)	0.402 (4)	0.299 (3)	0.566 (4)
HC(9′)	0.139 (6)	0.311 (4)	0.481 (6)
HC(9″)	0.244 (6)	0.274 (4)	0.406 (6)
HN(10)	0.225 (6)	−0.054 (5)	0.587 (6)
HC(11)	0.279 (6)	0.071 (4)	0.435 (5)
HC(13)	0.056 (6)	0.289 (4)	0.700 (6)
HC(14)	−0.057 (5)	0.190 (4)	0.877 (6)
HC(15)	−0.044 (6)	0.017 (4)	0.905 (6)
HC(16)	0.084 (7)	−0.058 (6)	0.754 (7)

positional and isotropic thermal parameters for the hydrogen atoms converged with values for R and R_w of 0.045 and 0.049, respectively. The final value of $[\Sigma w(|F_o| - |F_c|)^2/(m-n)]^{1/2}$, where m is the number of observations and n is the number of variables, was 3.93. The weighting scheme and atomic scattering factors used were as described above. The highest residual electron density in the final difference map was 0.37 e Å^{-3}. The final values for the atomic coordinates are given in Table II. The thermal parameters and the structure factors have been deposited.[16]

Results

A stereoscopic view of the two crystallographically independent molecules of Biot-C_3-Ind (**2**) is shown in Figure 1. While the absolute configuration was not determined in this study, all coordinates and drawings correspond to the absolute configuration determined previously for a heavy atom derivative of biotin.[9b] The molecule of Biot (SO)-C_3-Ind (**6**) is shown in Figure 2. While the atom numbering of the biotin moiety in these molecules is consistent with the chemical convention, the atoms of the trimethylene bridge and the indole group are numbered sequentially following numbers assigned to the biotin moiety (see **2**). Bond lengths and angles in the two molecules of **2** and in **6** are given in Table III.

Unlike some previously studied examples of trimethylene linked molecules,[6b,d] the molecules of **2** and that of **6** are more or less extended with no intramolecular overlap. The two molecules of **2** differ most clearly by rotations about the C(7)-C(8) bond; one molecule, designated as unprimed (**2**), has C(4) and C(9) disposed gauche, while in the primed molecule (**2′**), C(4) and C(9) are trans. There is also a slight difference in the orientation of C(8) with respect to the indole ring about the C(9)-C(12) bond. The molecule **6** has the trimeth-

Table III. Bond Lengths (Å) and Angles (deg) in the Two Independent Molecules of Biot-C_3-Ind (**2**) and in the Molecule of Biot(SO)-C_3-Ind (**6**)[a]

	unprimed	primed	Biot(SO)-C_3-Ind
N(1)-C(2)	1.360 (11)	1.337 (12)	1.362 (7)
N(1)-C(6a)	1.475 (10)	1.465 (11)	1.439 (6)
C(2)-N(3)	1.360 (11)	1.356 (11)	1.340 (5)
C(2)-O(17)	1.248 (9)	1.233 (10)	1.233 (5)
N(3)-C(3a)	1.470 (10)	1.464 (11)	1.448 (6)
C(3a)-C(4)	1.545 (10)	1.570 (11)	1.529 (6)
C(3a)-C(6a)	1.564 (11)	1.561 (12)	1.556 (7)
C(4)-S(5)	1.825 (8)	1.842 (9)	1.829 (4)
C(4)-C(7)	1.545 (10)	1.511 (12)	1.512 (6)
S(5)-C(6)	1.815 (8)	1.812 (9)	1.815 (5)
S(5)-O(18)			1.498 (3)
C(6)-C(6a)	1.556 (11)	1.561 (12)	1.528 (8)
C(7)-C(8)	1.543 (10)	1.516 (11)	1.535 (6)
C(8)-C(9)	1.555 (11)	1.522 (13)	1.512 (9)
C(9)-C(12)	1.512 (11)	1.494 (10)	1.517 (7)
N(10)-C(11)	1.429 (12)	1.389 (13)	1.335 (7)
N(10)-C(16a)	1.336 (12)	1.373 (15)	1.362 (7)
C(11)-C(12)	1.379 (12)	1.359 (13)	1.343 (7)
C(12)-C(12a)	1.415 (12)	1.434 (12)	1.406 (6)
C(12a)-C(13)	1.380 (12)	1.379 (13)	1.392 (7)
C(12a)-C(16a)	1.404 (11)	1.392 (12)	1.400 (6)
C(13)-C(14)	1.395 (14)	1.409 (13)	1.351 (11)
C(14)-C(15)	1.395 (14)	1.402 (14)	1.444 (18)
C(15)-C(16)	1.363 (14)	1.377 (18)	1.324 (12)
C(16)-C(16a)	1.414 (12)	1.389 (15)	1.400 (9)
C(6a)-N(1)-C(2)	112.0 (7)	114.4 (7)	112.0 (4)
N(1)-C(2)-N(3)	109.8 (7)	107.8 (8)	108.5 (4)
N(1)-C(2)-O(17)	125.6 (8)	126.4 (8)	125.5 (4)
N(3)-C(2)-O(17)	124.4 (8)	125.8 (8)	126.2 (4)
C(2)-N(3)-C(3a)	111.8 (7)	113.4 (7)	112.5 (4)
N(3)-C(3a)-C(4)	113.6 (6)	112.6 (6)	113.0 (4)
N(3)-C(3a)-C(6a)	103.0 (6)	102.5 (6)	101.8 (4)
C(4)-C(3a)-C(6a)	107.6 (6)	109.1 (6)	109.0 (4)
C(3a)-C(4)-S(5)	105.5 (5)	104.8 (5)	101.9 (3)
C(3a)-C(4)-C(7)	113.7 (6)	114.9 (6)	114.7 (4)
S(5)-C(4)-C(7)	112.7 (5)	113.6 (6)	114.6 (3)
C(4)-S(5)-C(6)	89.8 (4)	91.0 (4)	88.2 (2)
C(4)-S(5)-O(18)			110.3 (2)
C(6)-S(5)-O(18)			109.0 (2)
S(5)-C(6)-C(6a)	107.5 (5)	105.0 (6)	106.6 (4)
N(1)-C(6a)-C(3a)	102.4 (6)	101.7 (6)	103.2 (4)
N(1)-C(6a)-C(6)	110.6 (6)	112.2 (7)	113.2 (4)
C(3a)-C(6a)-C(6)	109.5 (6)	109.8 (7)	108.0 (4)
C(4)-C(7)-C(8)	113.3 (6)	113.3 (7)	113.8 (4)
C(7)-C(8)-C(9)	113.9 (6)	112.0 (7)	111.5 (4)
C(8)-C(9)-C(12)	111.2 (6)	115.7 (7)	112.4 (5)
C(11)-N(10)-C(16a)	109.4 (7)	109.0 (8)	108.1 (4)
N(10)-C(11)-C(12)	106.9 (7)	109.8 (8)	111.6 (5)
C(9)-C(12)-C(11)	125.6 (7)	127.8 (8)	127.8 (5)
C(9)-C(12)-C(12a)	126.5 (7)	126.4 (8)	126.1 (4)
C(11)-C(12)-C(12a)	107.6 (7)	105.8 (8)	106.0 (4)
C(12)-C(12a)-C(13)	133.0 (8)	132.6 (8)	134.2 (4)
C(12)-C(12a)-C(16a)	107.6 (7)	108.6 (8)	106.9 (4)
C(13)-C(12a)-C(16a)	119.3 (8)	118.8 (8)	119.0 (4)
C(12a)-C(13)-C(14)	120.0 (8)	120.4 (9)	119.6 (6)
C(13)-C(14)-C(15)	119.6 (9)	118.2 (10)	120.7 (8)
C(14)-C(15)-C(16)	122.1 (9)	122.8 (12)	120.1 (9)
C(15)-C(16)-C(16a)	117.8 (9)	116.8 (10)	119.5 (7)
N(10)-C(16a)-C(12a)	108.4 (8)	106.8 (8)	107.4 (4)
N(10)-C(16a)-C(16)	130.6 (8)	130.1 (10)	131.4 (5)
C(12a)-C(16a)-C(16)	121.0 (8)	123.0 (9)	121.2 (5)

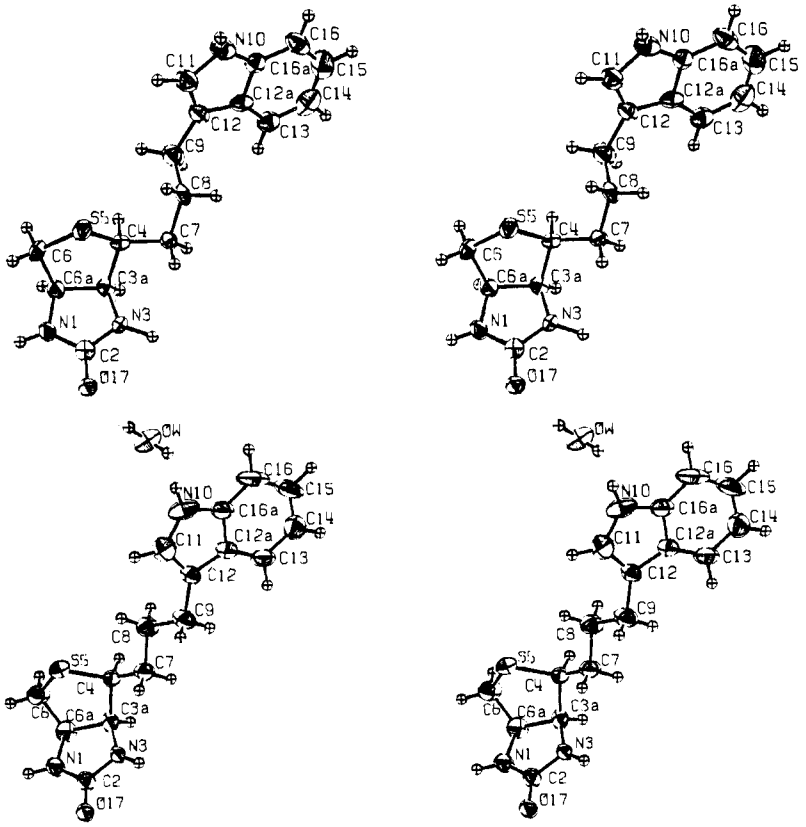


Figure 1. Stereoscopic view of the two molecules of **2**. The upper molecule is the unprimed one; the lower molecule is the primed one.

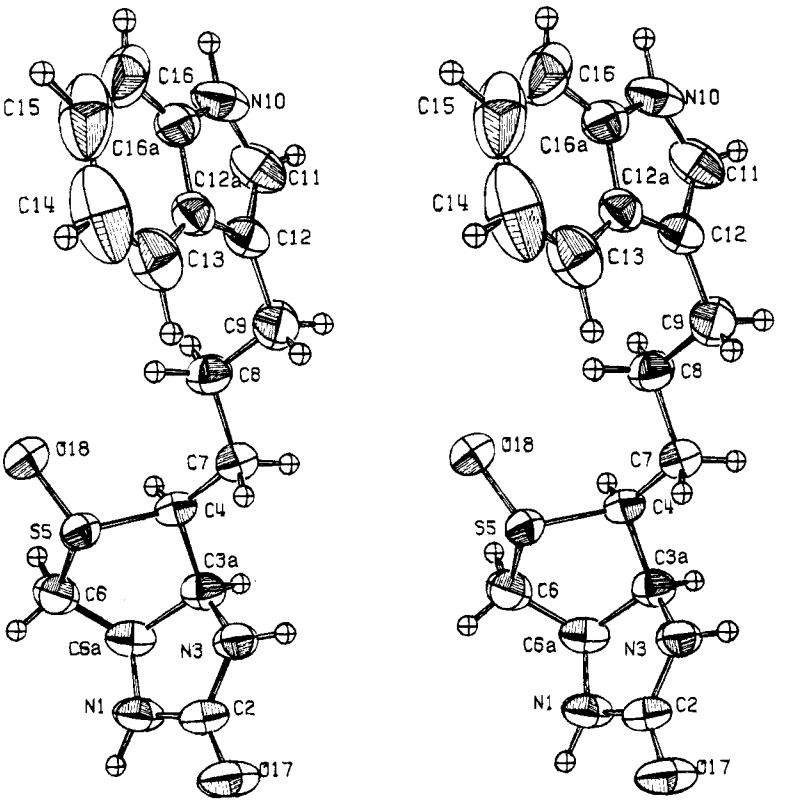


Figure 2. Stereoscopic view of the molecule of **6**.

ylene chain fully extended. The major difference in this molecule from **2** and **2′** is that the indole group is oriented around the C(9)-C(12) bond such that the six-membered ring of the indole is somewhat to the endo side of the biotin ring system.

Ureido Ring. The ureido rings in the two independent molecules of **2** and in **6** are approximately planar (Table IV); the carbonyl oxygen is displaced by 0.082, 0.056, and 0.172 Å in **2**, **2′**, and **6**, respectively, from the best plane of the atoms in the ring toward the exo side of the folded bicyclic biotin ring

Table IV. Some Details of Best Planes in the Molecules of **2**, **2′**, and **6**[b]

	2	2′	6
	Plane A		
N(1)	−0.034[a]	0.026[a]	0.002[a]
C(2)	−0.003[a]	−0.026[a]	−0.061[a]
N(3)	0.034[a]	0.006[a]	0.068[a]
C(3a)	−0.057[a]	0.008[a]	−0.072[a]
C(6a)	0.059[a]	−0.022[a]	0.050[a]
O(17)	−0.082	−0.056	−0.172
	Plane B		
C(3a)	0.067[a]	−0.010[a]	0.045[a]
C(4)	−0.038[a]	0.006[a]	−0.025[a]
C(6)	0.044[a]	−0.007[a]	0.045[a]
C(6a)	−0.072[a]	0.011[a]	−0.054[a]
S(5)	0.825	0.848	0.946
	Plane C		
N(10)	−0.022[a]	0.008[a]	−0.002[a]
C(11)	−0.004[a]	−0.004[a]	0.005[a]
C(12)	0.014[a]	0.016[a]	−0.011[a]
C(12a)	0.016[a]	−0.009[a]	0.005[a]
C(13)	−0.010[a]	−0.010[a]	0.005[a]
C(14)	−0.028[a]	0.002[a]	−0.002[a]
C(15)	0.009[a]	0.029[a]	−0.014[a]
C(16)	0.026[a]	−0.004[a]	−0.018[a]
C(16a)	−0.004[a]	−0.022[a]	0.011[a]
C(9)	0.158	0.080	0.038

[a] These distances involve atoms that were included in the best plane calculations. [b] The best planes were calculated by weighing the atoms according to $1/\sigma^2$.

system. The ureido carbonyl bonds are in the range 1.23–1.25 Å and the carbonyl carbon–nitrogen bond distances average 1.35 Å. These values may be compared with averages of 1.21 and 1.37 Å for the carbonyl and C–N bonds in 12 barbiturate structures.[17] The observed ureido carbonyl bond lengthening and carbonyl C–N shortening have also been reported in biotin,[7,8] dethiobiotin,[10] and azabiotin.[11] While the ring approaches planarity, the deviations from planarity impart considerable asymmetry to the ring about a plane that contains C(2) and bisects, and is normal to, the C(3a)–C(6a) bond. If we adopt the convention used by DeTitta et al.[8] for describing this asymmetry, the values of Φ_u are 12.6, 3.5, and 13.1° for **2**, **2′**, and **6**, respectively. The value of Φ_u in **1** is 2.5°.

Tetrahydrothiophene Ring. The tetrahydrothiophene rings in **2** and **6** are all nonplanar (Table IV). However, an examination of the torsion angles around the ring (Table V) indicates somewhat different patterns of nonplanarity. In **2′**, the ring is quite symmetrical about S(5) and the midpoint of the C(3a)–C(6a) bond, with the C(4)–C(3a)–C(6a)–C(6) torsion angle being 1.8°. In **2**, there is considerable deviation from symmetry, and the corresponding torsion angle is −12.2°. The ring in **6** is intermediate between those in **2** and **2′**, with a torsion angle of −9.1°. Applying the convention used by DeTitta et al.,[8] the values of Φ_T are 13,4, 4.1, and 10.0° for **2**, **2′**, and **6**, respectively; the value for **1** is 3.1°.[8] In all three molecules, the sulfur is displaced from the best plane of the four carbon atoms in the ring toward the concave or endo side of the folded bicyclic framework; these displacements are 0.825 and 0.848 Å in **2** and **2′** and 0.946 Å in **6**. A displacement of 0.862 Å was reported for biotin[8] and of 0.964 Å for **3**.[9] The dihedral angles between the four-atom plane C(3a)–C(4)–C(6)–C(6a) and the three-atom plane C(4)–S(5)–C(6) are 140.3 and 138.5° in **2** and **2′**, respectively, and 134.3° in **6**. These values may be compared with the corresponding angle, 137.6°, in biotin.[8]

Bicyclic Ring System. The bicyclic ring system is folded along the bond between C(3a) and C(6a) in both **2** and **6**. The dihedral angles between the best plane of the ureido ring and the best plane of the four carbon atoms of the tetrahydrothiophene ring are 118.5 and 120.1° in **2** and **2′** and 121.7° in **6**. These values are in agreement with the value of 122° in biotin.[8]

Table V. Torsion Angles[a] in the Two Molecules of **2** and in the Molecule of **6**

	2	2′	6
N(1)–C(2)–N(3)–C(3a)	−3.7	−3.3	−13.0
N(1)–C(6a)–C(3a)–N(3)	−9.4	2.5	−9.7
N(1)–C(6a)–C(3a)–C(4)	−129.6	−117.1	−129.3
N(1)–C(6a)–C(6)–S(5)	93.8	81.8	88.7
C(2)–N(1)–C(6a)–C(3a)	8.1	−4.9	3.0
C(2)–N(1)–C(6a)–C(6)	−108.6	−122.1	−113.6
C(2)–N(3)–C(3a)–C(4)	124.4	117.4	130.8
C(2)–N(3)–C(3a)–C(6a)	8.4	0.3	14.1
N(3)–C(2)–N(1)–C(6a)	−3.2	5.3	5.7
N(3)–C(3a)–C(4)–S(5)	−75.9	−86.1	−73.9
N(3)–C(3a)–C(4)–C(7)	48.0	39.3	51.3
N(3)–C(3a)–C(6a)–C(6)	108.1	121.4	110.5
C(3a)–N(3)–C(2)–O(17)	172.2	178.8	169.2
C(3a)–C(4)–S(5)–C(6)	−42.3	−39.8	−46.0
C(3a)–C(4)–S(5)–O(18)			−155.7
C(3a)–C(4)–C(7)–C(8)	−178.6	175.2	176.3
C(3a)–C(6a)–C(6)–S(5)	−18.4	−30.5	−25.0
C(4)–C(3a)–C(6a)–C(6)	−12.2	1.8	−9.1
C(4)–S(5)–C(6)–C(6a)	35.1	40.8	41.8
C(4)–C(7)–C(8)–C(9)	−55.8	−175.5	−178.8
S(5)–C(4)–C(3a)–C(6a)	37.3	27.0	38.5
S(5)–C(4)–C(7)–C(8)	−58.7	−64.2	−64.9
C(6)–S(5)–C(4)–C(7)	−166.9	−165.9	−171.3
C(6a)–N(1)–C(2)–O(17)	−179.1	−176.8	−176.5
C(6a)–C(3a)–C(4)–C(7)	161.3	152.4	163.7
C(6a)–C(6)–S(5)–O(18)			152.6
C(7)–C(4)–S(5)–O(18)			79.1
C(7)–C(8)–C(9)–C(12)	−164.1	177.5	−173.8
C(8)–C(9)–C(12)–C(11)	−82.9	61.2	−98.3
C(8)–C(9)–C(12)–C(12a)	89.9	−120.3	77.3
C(9)–C(12)–C(11)–N(10)	174.4	176.7	177.7
C(9)–C(12)–C(12a)–C(13)	8.4	1.5	3.4
C(9)–C(12)–C(12a)–C(16a)	−174.0	−177.9	−177.2
N(10)–C(11)–C(12)–C(12a)	0.4	−2.1	1.4
N(10)–C(16a)–C(12a)–C(12)	−0.2	0.6	0.0
N(10)–C(16a)–C(12a)–C(13)	177.8	−178.9	179.5
N(10)–C(16a)–C(16)–C(15)	−178.6	176.6	−179.1
C(11)–N(10)–C(16a)–C(12a)	0.5	−1.9	0.9
C(11)–N(10)–C(16a)–C(16)	−178.0	−179.0	178.4
C(11)–C(12)–C(12a)–C(13)	−177.7	−179.7	179.8
C(11)–C(12)–C(12a)–C(16a)	−0.1	0.9	−0.8
C(12)–C(11)–N(10)–C(16a)	−0.5	2.5	−1.4
C(12)–C(12a)–C(13)–C(14)	178.7	−178.1	178.8
C(12)–C(12a)–C(16a)–C(16)	178.4	178.0	−177.9
C(12a)–C(13)–C(14)–C(15)	1.4	0.6	−0.2
C(12a)–C(16a)–C(16)–C(15)	3.1	−0.1	−1.8
C(13)–C(12a)–C(16a)–C(16)	−3.6	−1.5	1.6
C(13)–C(14)–C(15)–C(16)	−1.8	−2.4	0.1
C(14)–C(13)–C(12a)–C(16a)	1.3	1.2	−0.6
C(14)–C(15)–C(16)–C(16a)	−0.4	2.1	0.9

[a] The sign of the torsion angle A–B–C–D is considered positive if, when looking along the B–C bond, atom A has to be rotated clockwise to eclipse atom D.

Propyl Chain and Indole Substituent. The propyl chain joining the bicyclic framework and the indole system has the fully extended, all-trans conformation (Figures 1 and 2 and Table V) in **2′** and **6**. However, in **2**, the chain is twisted such that the conformation about the C(7)–C(8) bond is nearly gauche. The indole substituent of each molecule is planar (Table IV). The differences in disposition of the indole ring with respect to the trimethylene chain can be seen by inspection of the torsion angles about the C(9)–C(12) bond (Table V). The bond distances and angles in the propyl chains and indole substituents are within normal ranges.

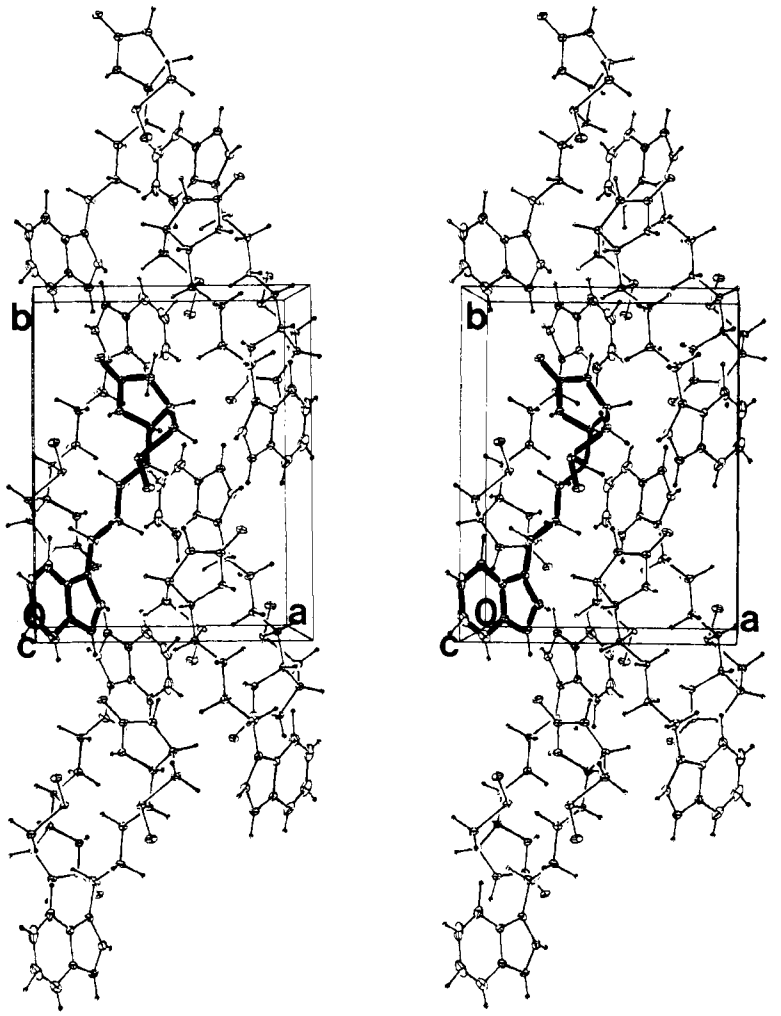


Figure 3. Stereoscopic view of the packing of **6** in the crystal. The reference molecule is shown by dark lines.

Table VI. Some Details of the Hydrogen Bonding in **6**

		N- - -O, Å	H- - -O, Å	N-H- - -O, deg
N(10)- - - O(17)	$x, 1+y, z$	2.801 (5)	1.87 (7)	162 (6)
N(1)- - - O(18)	$1-x, \frac{1}{2}+y, 1\frac{1}{2}-z$	2.936 (5)	2.07 (5)	148 (4)
N(3)- - - O(18)	$\frac{1}{2}-x, 1-y, -\frac{1}{2}+z$	3.582 (6)	2.85 (4)	149 (4)

Packing of the Molecules in the Crystals. Some details of the hydrogen bonding in **6** are given in Table VI. In Biot(SO)-C_3-Ind (**6**), there are chains of molecules that run along the *b* axis and that are held together by N(10)-H- - -O(17) intermolecular hydrogen bonding (Figure 3). In addition, there are hydrogen-bonded interactions in which O(18), i.e., the oxygen atom attached to sulfur, participates. These involve N(1) and a very much weaker interaction with N(3).

Details of the hydrogen bonding in Ind-C_3-Biot (**2**) are given in Table VII. In the crystal of **2**, a water molecule is interposed between N(10′) and O(17) in the two independent molecules and participates in hydrogen bonds (Figure 4). N(10) also appears to interact with the water molecule at $1\frac{1}{2}-x, -y, \frac{1}{2}+z$. The water molecule is also involved in hydrogen bonding to O(17) (the interposition described above) and in a hydrogen bond to the sulfur atom in the primed molecule at $1-x, -\frac{1}{2}+y, 1\frac{1}{2}-z$. There is a pseudocentrosymmetrically related pair of hydrogen bonds between an N-H and an O(17) in each of the two independent molecules. N(1)-H(N1) in the unprimed reference molecule forms a hydrogen bond to O(17′) in the molecule at $1-x, \frac{1}{2}+y, 1\frac{1}{2}-z$ and O(17) in the un-

Table VII. Some Details of the Hydrogen Bonding in **2**[a]

		A- - -B, Å	H- - -B, Å	A-H- - -B, deg
N(1)-H- - - O(17′)	$1-x, \frac{1}{2}+y, 1\frac{1}{2}-z$	2.872 (9)	2.00 (7)	162 (6)
N(3)-H- - - O(17′)	$1-x, -\frac{1}{2}+y, 1\frac{1}{2}-z$	2.962 (9)	1.95 (7)	158 (5)
N(10)-H- - - O(W)	$1\frac{1}{2}-x, -y, \frac{1}{2}+z$	2.977 (9)	2.47 (9)	125 (8)
N(3′)-H- - - O(17)	$1-x, -\frac{1}{2}+y, 1\frac{1}{2}-z$	2.962 (9)	2.18 (7)	166 (8)
N(10′)-H- - - O(W)		2.972 (10)	2.19 (8)	163 (9)
O(W)-HA(W)- - -O(17)		2.832 (8)	1.90 (9)	159 (8)
O(W)-HB(W)- - -S′	$1-x, -\frac{1}{2}+y, 1\frac{1}{2}-z$	3 ˙˙8 (6)	2.46 (9)	151 (7)
O(17)- - -O(W)- - -S′ angle is 100.9 (2)°				

[a] Distances and angles are given in terms of an A-H- - -B type interaction.

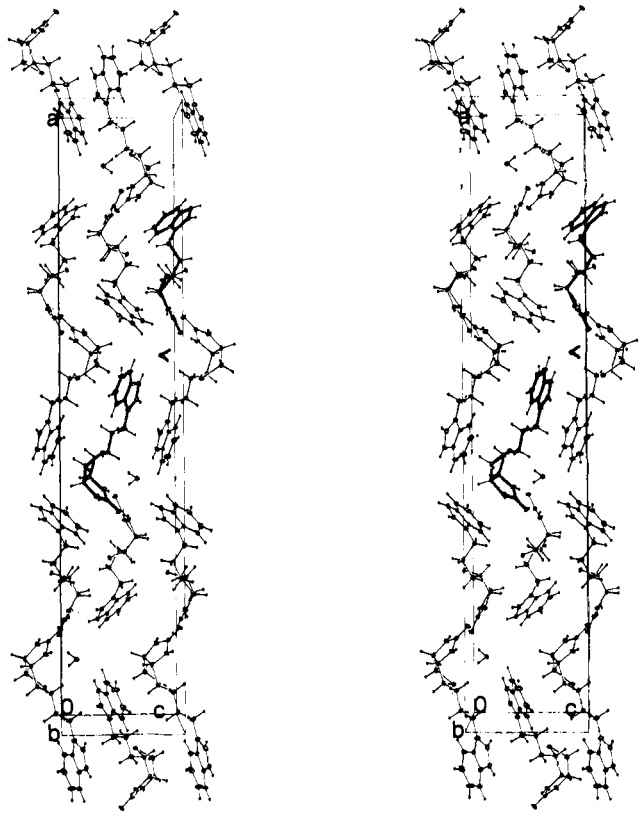


Figure 4. A stereoscopic view of the packing of molecules of **2** in the crystal. The reference molecules are shown by dark lines; that of **2** is to the upper right; that of **2′** is to the center of the cell.

primed reference molecule is involved in a hydrogen bond to the N(3′)-HN(3′) bond. N(3)-HN(3) forms a hydrogen bond to O(17′) in the molecule at $1 - x, -\frac{1}{2} + y, 1\frac{1}{2} - z$. The N(1′) atom in the primed molecule does not appear to participate in hydrogen bonding.

The biotin portion of the primed molecule makes a close approach to an indole group in the unprimed molecule at $-\frac{1}{2} + x, \frac{1}{2} - y, 1 - z$ (Figure 5a) and to an indole group in the primed molecule at $1 - x, \frac{1}{2} + y, \frac{1}{2} - z$ (Figure 5b). The shortest contacts for the first approach are N(1′)- - -C(12a) 3.457 Å, H(1′)- - -C(13) 3.11 Å, H(6a′)- - -N(10) 2.82 Å, and H(6a′)- - -C(16a) 2.78 Å. In the second approach, the shortest contacts are H(4′)- - -C(13′) 2.73 Å, H(4′)- - -C(14′) 2.77 Å, and H(6b′)- - -N(10′) 2.92 Å. The shortest intermolecular contact involving the biotin portion of the unprimed molecule with an indole group is one of 2.84 Å between H(3a) and C(12a) in the molecule at $1\frac{1}{2} - x, -y, -\frac{1}{2} + z$.

Discussion

The X-ray results confirm many of the ideas proposed on the basis of the spectroscopic study.[3] The sulfoxide (**6**) that is the major component of the oxidation of **2,** when bound to avidin and using aqueous NBS, is shown to have the α configuration (i.e., oxygen pointing toward the exo side of the biotin ring system) with a pseudoequatorial orientation. The side chain to the indole ring is also in a pseudoequatorial orientation. The H(C6′)-C(6)-C(6a)-H(C6a) torsion angle is 94.5°, in good agreement with the "close to 90°" anticipated from the low value of the coupling constant.

There appears to be no strong intramolecular interaction between the indole and biotin rings in any of the three molecules studied by X-ray methods. The bond lengths, angles, and general conformations of the two ring systems appear to be similar to those found in other indole and biotin derivatives. The most unusual feature of the molecular geometry is the very definite dissymmetry in the two five-membered rings of the biotin groups in **2** and **6** when compared to the geometry of biotin itself.[8] The effect is much less apparent in the molecule of **2′**.

While there is intermolecular hydrogen bonding in the crystals of both **2** and **6,** it does not appear to be unusually strong. As molecules **2, 2′,** and **6** all show lengthening of the carbonyl bond and attendant shortening of the adjacent C-N bonds, we are not convinced that this general feature of biotin molecules is caused by strong hydrogen bonding. Only in the case of **6** is there direct hydrogen bonding between the indole group and the biotin group.

At the conclusion of the spectroscopic study, the suggestion was made that the biotin molecule almost fills a hydrophobic pocket in the protein molecule.[3] While clearly not all the components of such a pocket are present in the crystal of **2,** it is of considerable interest that the exo side of the biotin ring in the primed molecule of **2** does fit neatly and snugly into a hydrophobic V-shaped crevice provided by two indole rings (Figures 4 and 5). The two bridge hydrogen atoms (H(3a) and H(6a)) point into the axis of the V-shaped wall. The geometrical arrangement of molecules participating in this interaction can be seen on the left of Figure 4 about halfway up the cell involving the reference molecule **2′**. The existence of a hy-

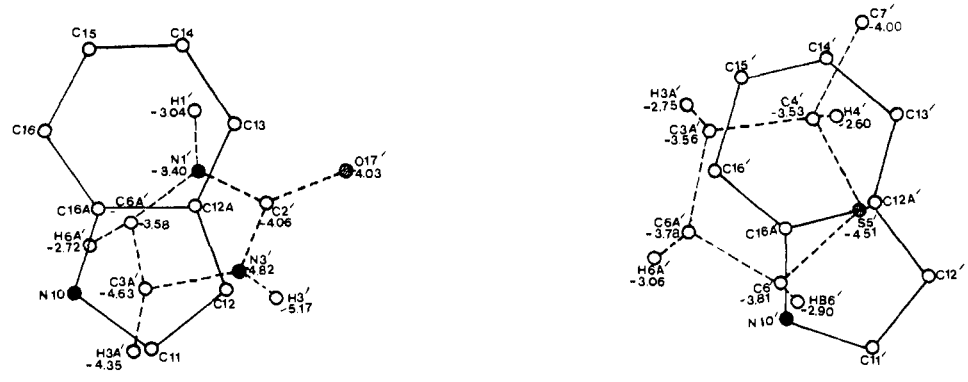


Figure 5. Drawings of the overlap that the biotin ring system makes with indole rings in the molecule of **2.** The atoms of the biotin groups are projected onto the best planes defined by the nine atoms of the indole rings. Distances of atoms from the planes are shown in Å: (a) the overlap of the biotin portion of the primed reference molecule with the indole ring of an unprimed molecule at $-\frac{1}{2} + x, \frac{1}{2} - y, 1 - z$; (b) the overlap of the biotin of the primed molecule with the indole ring of a primed molecule at $1 - x, \frac{1}{2} + y, \frac{1}{2} - z$.

drophobic pocket with a correspondingly snug fit could assist in a strong complex between avidin and biotin.

Acknowledgment. We thank Professor N. J. Leonard for stimulating discussions on this problem. F.-T.L. was supported by a research grant (CHE-76-23543) from the National Science Foundation to Professor Leonard.

Supplementary Material Available: Tables of observed and calculated structure factors and thermal parameters for the crystals of **2** and **6** (42 pages). Ordering information is given on any current masthead page.

References and Notes

(1) This work was supported by NIH GM 19336.
(2) N. M. Green, *Adv. Protein Chem.*, **29**, 84 (1975), and references cited therein.
(3) F.-T. Liu and N. J. Leonard, *J. Am. Chem. Soc.*, preceding paper in this issue.
(4) K. Mutai, B. A. Gruber, and N. J. Leonard, *J. Am. Chem. Soc.*, **97**, 4095 (1975), and references cited therein.
(5) N. J. Leonard, R. S. McCredie, M. W. Logue, and R. L. Cundall, *J. Am. Chem. Soc.*, **95**, 2320 (1973).
(6) (a) N. J. Leonard, K. Golankiewicz, R. S. McCredie, S. M. Johnson, and I. C. Paul, *J. Am. Chem. Soc.*, **91**, 5855 (1969); (b) J. K. Frank and I. C. Paul, *ibid.*, **95**, 2324 (1973); (c) P. L. Johnson, C. A. Maier, and I. C. Paul, *ibid.*, **95**, 5370 (1973); (d) P. L. Johnson, J. K. Frank, and I. C. Paul, *ibid.*, **95**, 5377 (1973); (e) J. K. Frank, N. N. Thayer, and I. C. Paul, *ibid.*, **95**, 5386 (1973).
(7) (a) W. Traub, *Nature (London)*, **178**, 649 (1956); (b) *Science*, **129**, 210 (1959).
(8) G. T. DeTitta, J. W. Edmonds, W. Stallings, and J. Donohue, *J. Am. Chem. Soc.*, **98**, 1920 (1976).
(9) C. Bonnemere, J. A. Hamilton, L. K. Steinrauf, and J. Knappe, *Biochemistry*, **4**, 240 (1965); (b) J. Trotter and J. A. Hamilton, *ibid.*, **5**, 713 (1966).
(10) C.-S. Chen, R. Parthasarathy, and G. T. DeTitta, *J. Am. Chem. Soc.*, **98**, 4983 (1976).
(11) M. D. Glick, H. C. Wormser, and H. N. Abramson, *Acta Crystallogr., Sect. B*, **33**, 1095 (1977).
(12) G. Germain, P. Main, and M. M. Woolfson, *Acta Crystallogr., Sect. A*, **27**, 368 (1971).
(13) Virtually all crystallographic calculations were carried out on the Syntex EXTL structure determination system, provided by Syntex Analytical Instruments, Cupertino, Calif. 95014.
(14) $R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$; $R_w = [\Sigma w||F_o| - |F_c||^2/\Sigma w|F_o|^2]^{1/2}$.
(15) J. A. Ibers and W. C. Hamilton, Ed., "International Tables for X-ray Crystallography", Vol. IV, Kynoch Press, Birmingham, England, 1974, pp 71–147.
(16) See paragraph at end of paper regarding supplementary material.
(17) B. M. Craven, C. Cusatis, G. L. Gartland, and E. A. Vizzini, *J. Mol. Struct.*, **16**, 331 (1973).

Nonoxidative Cyclization of Squalene by *Tetrahymena pyriformis*. Incorporation of a 3β Hydrogen (Deuterium) Atom into Tetrahymanol

D. John Aberhart and Eliahu Caspi*

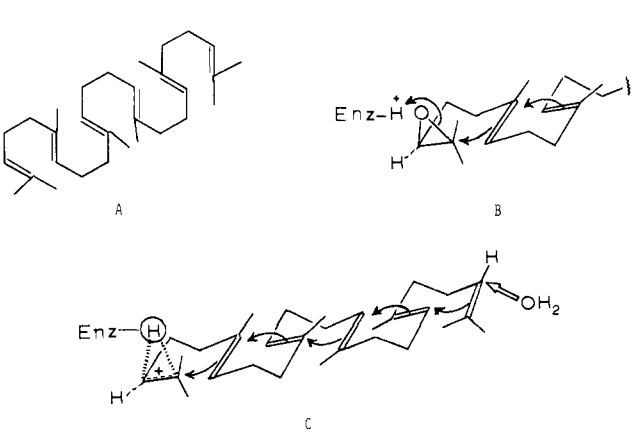
Contribution from The Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts 01545. *Received September 5, 1978*

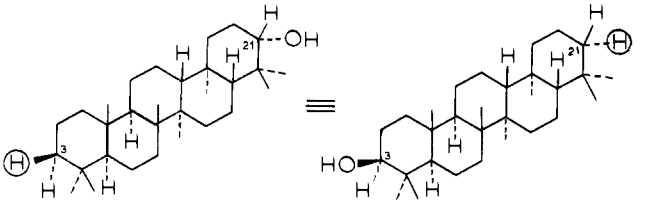
Abstract: 3α- and 3β-[$^2H_1$]-5α-cholestanes and 3α- and 3β-[$^2H_1$]-4,4-dimethyl-5α-cholestanes were synthesized, and their ^2H NMR spectra were determined. The chemical shifts of the 3α- and 3β-deuterio isomers were separated by 0.46 ppm in the cholestane isomers and by 0.23 ppm in the 4,4-dimethylcholestanes. The results were then used to establish that the nonoxidative cyclization of squalene by enzymes of *Tetrahymena pyriformis* in D_2O proceeds with the introduction of a 3β-deuterium atom into tetrahymanol. The ^2H NMR results were supported by comparison of the C-^2H stretching patterns of 3α- and 3β-[$^2H_1$]-4,4-dimethyl-5α-cholestanes with the C-^2H stretching pattern of biosynthetic tetrahymanol. The mechanism of the nonoxidative cyclization of squalene is also discussed.

Introduction

It is now generally accepted that the biosynthesis of C-3 oxygenated triterpenes and sterols is an oxidative process involving the initial formation of 2,3(*S*)-oxidosqualene[1] (B). Enzymatic "cationic" cleavage of the oxirane ring is thought to generate an electron deficiency at C-2 which initiates the cyclization process. Depending on the species, different C-3 oxygenated products are then formed.[1]

In contrast, we have demonstrated that the biosynthesis of the triterpene tetrahymanol (**1**) by *Tetrahymena pyriformis* is not an oxygen-dependent process.[2] Accordingly, the biosynthesis of tetrahymanol (**1**) by an enzyme preparation of *T.*





1a 1b

Structures **1a** and **1b** are equivalent. For convenience of comparison of the ^2H NMR spectra of [^2H]tetrahymanol and the [3-$^2H_1$] model compounds, the numbering system (cf. ref 2) is used in the present paper. This is in contrast to the numbering system used in our paper, ref 3b.

Ⓗ: HYDROGEN (DEUTERIUM) ATOM INTRODUCED IN THE BIOSYNTHESIS FROM SQUALENE

0002-7863/79/1501-1013$01.00/0