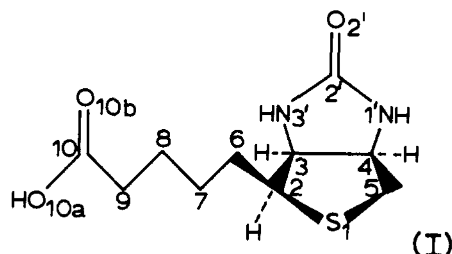
# Molecular Structure of Biotin. Results of Two Independent Crystal Structure Investigations

**George T. DeTitta,* James W. Edmonds,[1a] William Stallings,[1b] and Jerry Donohue***

*Contribution from the Medical Foundation of Buffalo, Buffalo, New York 14203 and University of Pennsylvania, Philadelphia, Pennsylvania 19174.* *Received May 19, 1975*

**Abstract:** The three-dimensional crystal structure of *d*-(+)-biotin, the coenzyme responsible for the fixation and transfer of carbon dioxide in biological systems, has been determined by x-ray diffraction techniques in two laboratories independently. The results of the two determinations are compared and discussed in terms of the mode of nucleophilic activation of the coenzyme. Of particular relevance may be the participation of the ureido carbonyl oxygen as acceptor of a strong hydrogen bond (O $\cdots$ O distance 2.54 Å) which may cause the observed lengthening of the ureido carbonyl bond to 1.25 Å and shortening of the carbonyl carbon–nitrogen bonds to 1.33 and 1.35 Å. These results suggest a partial delocalization of the electronic charge within the ureido group, and are supportive of a proposed activation via polarization mechanism. A close intramolecular nonbonded contact involving the nitrogen atom proximal to the valeryl chain probably precludes carboxylation from that side of the molecule.

Biotin (I) participates as coenzyme in a variety of carboxylase, decarboxylase, and transcarboxylase systems. Its role is to fix $CO_2$ for eventual transfer, and it accomplishes its task in a two-step fashion during which a labile enzyme–biotin–$CO_2$ complex is formed.[2] Early crystallographic




studies of biotin by Traub[3] confirmed the relative stereochemistry at the asymmetric carbons previously suggested by Grob and von Sprecher.[4] In studies of the biotin carboxylation mechanism Bonnemere, Hamilton, Steinrauf, and Knappe[5] prepared the bis-*p*-bromoanilide derivative of the major product in the biotin methyl ester–methyl chloroformate reaction and, by crystallographic determination, confirmed that the primary site of carboxylation in this model reaction was the ureido 1′-N nitrogen. Knappe had already shown that the major product had identical properties with the dimethyl ester of carboxybiotin obtained by diazomethane trapping experiments in biological systems. Subsequently Trotter and Hamilton[6] measured the anomalous dispersion properties of the same compound, 1′-*N*-carboxybiotin bis-*p*-bromoanilide, and established the absolute configuration of biotin.

Continued interest in the mode of action of biotin-dependent enzymes as well as our own interest in careful structural investigation of the biotin vitamers prompted a reinvestigation of the biotin crystal structure, and ultimately led to its determination by two groups working independently of the other. A collaborative report of our results is presented. Where necessary to distinguish between the results of the two groups, work performed by DeTitta and Edmonds, Medical Foundation of Buffalo, is designated I, and by Stallings and Donohue, University of Pennsylvania, II.

## Experimental Section

I. Needle-shaped crystals elongated about the *a* axis were grown by slow cooling of an aqueous solution of biotin (Sigma Co. Lot no. 12C-2000) saturated at 95 °C. A crystal of cross section 0.05 × 0.14 mm and length 0.88 mm was cleaved from a longer crystal and mounted about the *a* axis for data collection. The crystal quality was checked by photographic methods to ensure that it was not damaged in the cleavage process. Cell constants, Table I, were determined from a least-squares refinement of the $2\theta$ values ($2\theta >$ 56°) of 15 reflections centered on a kappa geometry automatic diffractometer. All reflections of the type *hkl* with sin $\theta/\lambda$ less than 0.63 Å$^{-1}$ were measured in the $\theta$–$2\theta$ scan mode using nickel-filtered Cu K$\alpha$ radiation ($\lambda$ = 1.5418 Å) and a fixed scan width of 2°. The data collection was carried out using a diffractometer, the narrowest dimension of whose intensity plateau was 0.85 mm; the takeoff angle was 6°. Therefore, the intensity measurement of only a small number of reflections could be affected by the long dimension of the crystal. These conditions are not particularly critical since, while a maximum of 0.04 (3/88) of the scattering matter lies outside the homogeneous portion of the x-ray beam, the entire crystal is completely within the beam and a large crystal permitted the collection of relatively more data. A total of 1404 data was measured. The data were corrected for Lorentz, polarization, and absorption effects. The variance in the structure factor is estimated to be $\sigma^2(|F|^2) = \sigma^2(I) + 0.06|F|^4$ where $\sigma^2(I) = N$. All reflections were treated as observationally present in subsequent steps.

II. A crystalline sample of *d*-(+)-biotin was obtained from the Nutritional Biochemicals Corp., Cleveland, Ohio; clear, colorless, needle-shaped crystals elongated about the *a* axis were obtained by evaporation to dryness of an aqueous solution of the compound. A crystal of approximate dimensions 0.03 × 0.05 × 0.35 mm was selected for data collection; the quality of the sample was checked by photographic methods. Three-dimensional data were measured on a four-circle automated diffractometer using graphite-monochromatized Cu K$\alpha$ radiation. Lattice constants, Table I, were refined by a least-squares treatment of 12 automatically centered reflections whose $2\theta$ values were 50° or greater. Using a $\theta$–$2\theta$ scan mode and a dispersion-corrected $2\theta$ base width of 1.25°, the data were collected to a maximum sin $\theta/\lambda$ value of 0.58 Å$^{-1}$. A total of 1094 intensities was measured and consequently corrected for Lorentz and polarization factors as well as attenuation of the diffracted beam by aluminum foil. Of the measured intensities, 169 were considered to be unobserved, being less than 3.09 times the respective standard deviations. The variance in the structure factor is estimated to be $\sigma^2(|F|^2) = \sigma^2(I) + (0.02F^2)^2$ where the final term of the sum is included to account for electronic instability of the counting instrument and where, for an unattenuated intensity, $\sigma^2(I) = N + m^2(b_1 + b_2)$; $m$ is a factor which places the background counting times on a common basis with the total counting time.

**Structure Determination and Refinement.** I. The structure was determined via the multisolution tangent formula procedure MULTAN.[7] The 16 nonhydrogen atoms were located in the first *E* map calculated. Their positional and isotropic thermal parameters were refined by block diagonal least squares to a residual *R* of 0.10. At this point a difference synthesis using those reflections with sin $\theta/\lambda$ $\leq$ 0.3 Å$^{-1}$ was calculated and the positions of the hydrogen atoms obtained. Full-matrix least-squares refinement of the positional and anisotropic thermal parameters of the nonhydrogen atoms and

Table I. Crystal Data for *d*-Biotin

| | Group I | Group II | $C_{10}H_{16}O_3N_2S$ |
|---|---|---|---|
| *a* | 5.241 (1) Å | 5.238 (1) Å | Mol wt 244.3 |
| *b* | 10.350 (2) | 10.346 (2) | $\rho_0^{3a}$ = 1.42 g cm$^{-3}$ |
| *c* | 21.042 (2) | 21.038 (3) | $\mu_{CuK\alpha}$ = 24.2 cm$^{-1}$ |
| *V* | 1141.4 Å$^3$ | 1140.1 Å$^3$ | Space group $P2_12_12_1$ |
| $\rho_c$ | 1.421 g cm$^{-1}$ | 1.423 g cm$^{-1}$ | $Z = 4$ |

the positional and isotropic thermal parameters of the hydrogen atoms converged in four cycles with a conventional residual $R$ = 0.053 and weighted residual $R_w$ = 0.111. The goodness of fit parameter was $S$ = 1.05. The function minimized was $\Sigma w(|F_o| - 1/k|F_c|)^2$ where $w = 1/\sigma^2(F_o)$ and $\sigma(F_o) = \sigma(F_o^2)/2F$; $k$ is the scale factor for the observed structure factors. Two of the hydrogen atom isotropic thermal parameters, for $H_3$ and $H_{6b}$, had converged to 0.0 and in fact the thermal parameter for $H_3$ had gone negative and was reset to 0.0 in the final cycle of refinement. A final difference map was essentially flat, and since the two hydrogen atoms had refined to chemically sensible locations, no further refinement was deemed justified. Values of the atomic scattering factors for carbon, oxygen, nitrogen, and sulfur are those of Cromer and Waber.[8] The atomic scattering factor for hydrogen is that of Stewart, Davidson, and Simpson.[9] Anomalous dispersion corrections, $\Delta f'$ = 0.31 and $\Delta f''$ = 0.58, were applied for the sulfur atom throughout the refinement employing the absolute configuration as determined by Trotter and Hamilton.[6] Values for $\Delta f'$ and $\Delta f''$ are those of Cromer.[10] Computer programs for data processing and absorption correction were those of Coppens, Leiserowitz, and Rabinovich[11] as modified by one of us (J.W.E.).

II. The structure was determined using standard Patterson-superposition techniques. The sulfur atom was readily located from the three Harker sections of a sharpened Patterson map; a structure factor calculation using just the sulfur atom yielded a residual of 0.55. Fourfold superposition of the sharpened Patterson on the sulfur coordinates resulted in the location of 11 of the remaining 15 atoms. It is interesting to note that superposition of an unsharpened Patterson map produced results which were generally uninterpretable. The remaining four nonhydrogen atoms were readily located in a Fourier map calculated from the coordinates of the located atoms. Three cycles of isotropic and three cycles of anisotropic least squares reduced the residual to 0.07. Using all data whose values of sin $\theta/\lambda$ were less than 0.35 Å$^{-1}$, a difference Fourier was calculated at this point to locate the hydrogen atoms; of the 16, only two, $H_{8a}$ and $H_{10a}$, were not well located. Two additional least-squares cycles reduced $R$ to 0.056. To refine the hydrogen atom parameters, the heavy atoms were fixed and the hydrogen atoms allowed to refine isotropically on all data whose values of sin $\theta/\lambda$ were less than 0.35 Å$^{-1}$. After four cycles, during which the temperature factors of the two hydrogen atoms not well located appeared to be "blowing up", convergence was obtained. Four more cycles of anisotropic refinement on the heavy atoms produced convergence; $R$ = 0.046 and $R_w$ = 0.057. The function minimized was $\Sigma w(|F_o| - 1/k|F_c|)^2$ where $k$ is the scale factor for the observed structure factors. The data were given weights, $w$, inversely proportional to the standard errors of the observed structure amplitudes, where $\sigma(F_o) = \{|F_o|^2 + \sigma(I)\}^{1/2} - |F_o|$. A final difference Fourier revealed the presence of a weak residual peak in the vicinity of the carboxylic hydrogen suggesting the possibility of disorder in that hydrogen atom; the possibility that the residual peak is spurious must also be weighed very heavily. The same argument cannot be applied in the case of $H_{8a}$. The superposition programs of Hubbard and Jacobson[12a] were used in the fourfold superposition of the Patterson map; these programs utilize a minimum function technique. The program UCLALS4[12b] was used in the refinement of the structure. Atomic scattering factors for all atoms were taken from the "International Tables for X-Ray Crystallography", Vol. II, Kynoch Press, Birmingham, England.

**Comparison of the Structural Results.** In order to analyze the derived results of the two independent structure determinations, the positional parameters, Table II, and all intramolecular distances were compared by half-normal probability plot techniques.[13a] Graphs of $\delta_{real}$ vs. $\delta_{expected}$ for normal distributions of positional parameters and distances are shown in Figure 1. In both cases, the points fall quite close to the line fit by least squares to those points

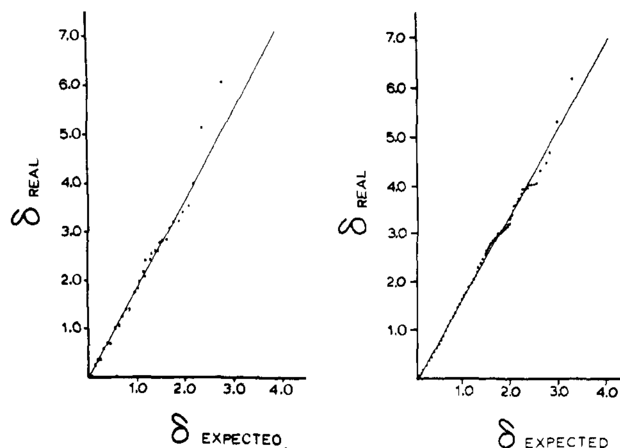



Figure 1. Half-normal probability plots of $\delta_{real}$ vs. $\delta_{expected}$ for positional parameters (left) and intramolecular distances (right) of *d*-biotin. All points with $\delta_{real} \geq 3.0$ are shown while, for clarity, every fifth point with $\delta_{real} < 3.0$ is shown.

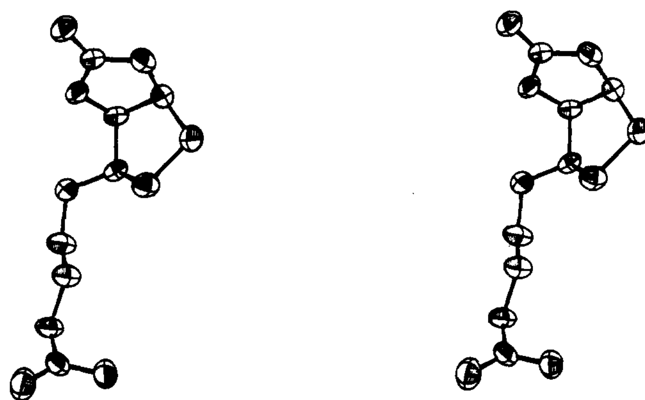


Figure 2. The molecular conformation of *d*-biotin. The positional parameters, Table II, of II were employed to construct the stereoview.

with $\delta_{real} \leq 3.0$, and deviate slightly from the lines only at the upper extremes of the graphs. This is a good indication that the errors in each are distributed normally. The slopes (1.839, 1.776) and intercepts (−0.051, −0.067) of the two graphs are comparable. Following the rationale of Abrahams and Keve[13a] it may be concluded that, on the average, the standard deviations in the positional parameters and bond lengths have been underestimated by a factor of approximately 1.8. That the standard deviations obtained from routine least-squares refinement are generally underestimated, the so-called Hughes effect,[13b] has been recognized (and generally ignored) for some time. By similar reasoning the differences at the extremes of the plots are only marginally significant if a correction factor of 1.8 is applied to the estimated standard deviations. It is interesting to note that the estimated $\sigma_z$ for the sulfur atom, Table II, is of the order $5 \times 10^{-5}$ for both determinations, or approximately $\sigma_z/4$ for the remaining nonhydrogen atoms; in all likelihood $\sigma_z$ for sulfur has been grossly underestimated.

From this analysis it follows that a realistic estimate of the standard deviation in a bond distance between nonhydrogen atoms, $\sigma_{BL}$, would be the average estimated $\sigma$ of 0.006 Å multiplied by the correction factor 1.8, $\sigma_{BL}$ = 0.11 Å. Similarly the corrected average involving the hydrogen atoms, $\sigma'_{BL}$, is of the order 0.19 Å and for bond angles among nonhydrogen atoms $\sigma_{BA}$ = 0.6°.

## Results

Positional and thermal parameters for biotin, as well as the estimated standard deviation in a parameter derived from the least-squares refinement, are collected in Table II. Bond distances and angles are collected in Table III. A stereoview of the molecule is presented in Figure 2.

**Ureido Ring.** The ureido ring, including the carbonyl oxygen, is essentially planar. The maximum deviation from the

Table II. Positional and Thermal Parameters for *d*-(+)-Biotin[a]

| | X | Y | Z | U(1,1) | U(2,2) | U(3,3) | U(1,2) | U(1,3) | U(2,3) |
|---|---|---|---|---|---|---|---|---|---|
| $S_1$ | −0.0575 (2) | 0.3701 (1) | 0.45313 (5) | 0.0394 (6) | 0.0493 (7) | 0.0407 (6) | 0.0072 (5) | −0.0055 (4) | 0.0004 (4) |
| | −0.0565 (3) | 0.3705 (1) | 0.45273 (6) | 0.0398 (8) | 0.0510 (8) | 0.0383 (7) | 0.0078 (8) | −0.0067 (8) | 0.0017 (6) |
| $C_2$ | −0.3653 (9) | 0.4431 (4) | 0.4323 (2) | 0.032 (2) | 0.034 (2) | 0.030 (2) | 0.000 (2) | 0.000 (1) | 0.001 (1) |
| | −0.361 (1) | 0.4423 (5) | 0.4324 (2) | 0.034 (3) | 0.039 (3) | 0.029 (2) | 0.000 (3) | 0.000 (3) | 0.009 (2) |
| $C_3$ | −0.3714 (9) | 0.4368 (4) | 0.3602 (2) | 0.032 (2) | 0.039 (2) | 0.034 (2) | −0.005 (2) | −0.009 (2) | 0.002 (1) |
| | −0.374 (1) | 0.4376 (5) | 0.3592 (2) | 0.039 (3) | 0.043 (3) | 0.024 (2) | 0.000 (3) | −0.002 (3) | 0.002 (2) |
| $C_4$ | −0.225 (1) | 0.3166 (4) | 0.3360 (2) | 0.048 (3) | 0.038 (2) | 0.034 (2) | −0.002 (2) | 0.000 (2) | −0.006 (2) |
| | −0.229 (1) | 0.3162 (5) | 0.3359 (2) | 0.052 (4) | 0.035 (3) | 0.037 (3) | −0.003 (3) | 0.000 (3) | −0.002 (2) |
| $C_5$ | −0.111 (1) | 0.2474 (5) | 0.3927 (2) | 0.054 (3) | 0.040 (2) | 0.043 (2) | 0.009 (2) | 0.009 (2) | −0.003 (2) |
| | −0.109 (1) | 0.2469 (5) | 0.3944 (2) | 0.065 (5) | 0.039 (3) | 0.047 (3) | 0.014 (4) | 0.006 (3) | 0.002 (2) |
| $C_6$ | −0.414 (1) | 0.5776 (4) | 0.4571 (2) | 0.054 (3) | 0.033 (2) | 0.032 (2) | 0.004 (2) | 0.002 (2) | 0.000 (1) |
| | −0.410 (1) | 0.5773 (5) | 0.4571 (2) | 0.057 (4) | 0.037 (3) | 0.035 (2) | 0.005 (3) | −0.003 (3) | 0.002 (3) |
| $C_7$ | −0.449 (1) | 0.5934 (5) | 0.5298 (2) | 0.042 (2) | 0.046 (2) | 0.033 (2) | 0.014 (2) | 0.002 (2) | −0.006 (2) |
| | −0.451 (1) | 0.5921 (5) | 0.5295 (2) | 0.050 (3) | 0.063 (4) | 0.032 (2) | 0.015 (4) | −0.006 (3) | −0.008 (3) |
| $C_8$ | −0.2030 (9) | 0.5805 (5) | 0.5680 (2) | 0.037 (2) | 0.050 (2) | 0.035 (2) | 0.001 (2) | −0.004 (2) | 0.006 (2) |
| | −0.203 (1) | 0.5791 (5) | 0.5683 (2) | 0.047 (4) | 0.056 (4) | 0.034 (3) | 0.004 (3) | 0.000 (3) | −0.004 (3) |
| $C_9$ | −0.245 (1) | 0.6254 (5) | 0.6369 (2) | 0.037 (2) | 0.055 (3) | 0.036 (2) | 0.006 (2) | −0.007 (2) | −0.001 (1) |
| | −0.247 (1) | 0.6272 (5) | 0.6366 (2) | 0.053 (4) | 0.053 (3) | 0.027 (2) | 0.010 (4) | −0.003 (3) | −0.004 (3) |
| $C_{10}$ | −0.013 (1) | 0.6055 (4) | 0.6764 (2) | 0.042 (2) | 0.036 (2) | 0.034 (2) | 0.002 (2) | 0.000 (2) | −0.005 (1) |
| | −0.011 (1) | 0.6046 (5) | 0.6763 (2) | 0.052 (4) | 0.037 (3) | 0.034 (2) | −0.004 (3) | 0.003 (3) | −0.010 (2) |
| $O_{10a}$ | 0.034 (1) | 0.4846 (4) | 0.6883 (2) | 0.071 (3) | 0.043 (2) | 0.055 (2) | 0.006 (2) | −0.030 (2) | −0.008 (1) |
| | 0.0335 (9) | 0.4837 (3) | 0.6885 (2) | 0.078 (3) | 0.036 (2) | 0.060 (2) | 0.003 (3) | −0.028 (3) | −0.002 (2) |
| $O_{10b}$ | 0.121 (1) | 0.6919 (4) | 0.6971 (2) | 0.055 (2) | 0.043 (2) | 0.075 (2) | −0.005 (2) | −0.015 (2) | 0.003 (2) |
| | 0.119 (1) | 0.6905 (4) | 0.6962 (2) | 0.069 (3) | 0.041 (2) | 0.074 (3) | −0.004 (3) | −0.022 (3) | 0.007 (2) |
| $N_{1'}$ | −0.0362 (9) | 0.3713 (3) | 0.2933 (2) | 0.051 (2) | 0.030 (2) | 0.043 (2) | 0.003 (3) | 0.010 (2) | −0.002 (1) |
| | −0.035 (1) | 0.3725 (4) | 0.2940 (2) | 0.057 (3) | 0.041 (2) | 0.039 (2) | 0.000 (3) | 0.011 (3) | −0.007 (2) |
| $C_{2'}$ | −0.052 (1) | 0.5002 (4) | 0.2901 (2) | 0.048 (2) | 0.035 (2) | 0.029 (2) | 0.004 (2) | 0.002 (2) | 0.000 (1) |
| | −0.051 (1) | 0.5002 (5) | 0.2902 (2) | 0.054 (3) | 0.039 (3) | 0.026 (2) | −0.001 (4) | 0.003 (3) | 0.002 (2) |
| $O_{2'}$ | 0.0801 (8) | 0.5732 (3) | 0.2561 (2) | 0.059 (2) | 0.043 (2) | 0.052 (2) | −0.004 (2) | 0.018 (2) | 0.000 (1) |
| | 0.0798 (8) | 0.5731 (3) | 0.2560 (1) | 0.063 (3) | 0.044 (2) | 0.046 (2) | −0.003 (3) | 0.021 (3) | 0.009 (2) |
| $N_{3'}$ | −0.2394 (9) | 0.5428 (4) | 0.3290 (2) | 0.053 (2) | 0.038 (2) | 0.030 (2) | 0.003 (2) | 0.013 (2) | 0.006 (1) |
| | −0.237 (1) | 0.5419 (4) | 0.3296 (2) | 0.052 (3) | 0.034 (2) | 0.035 (2) | 0.003 (3) | 0.014 (3) | 0.007 (2) |

| | X | Y | Z | U(1,1) | | X | Y | Z | U(1,1) |
|---|---|---|---|---|---|---|---|---|---|
| $H_2$ | −0.49 (1) | 0.387 (4) | 0.445 (2) | 0.01 (1) | $H_{7b}$ | −0.49 (3) | 0.674 (9) | 0.539 (5) | 0.06 (4) |
| | −0.48 (1) | 0.391 (5) | 0.454 (2) | 0.02 (2) | | −0.49 (2) | 0.685 (6) | 0.543 (3) | 0.06 (3) |
| $H_3$ | −0.54 (2) | 0.453 (6) | 0.346 (3) | 0.00 (2) | $H_{8a}$ | −0.06 (2) | 0.623 (8) | 0.548 (3) | 0.05 (2) |
| | −0.53 (2) | 0.424 (6) | 0.350 (4) | 0.11 (3) | | −0.03 (2) | 0.623 (7) | 0.550 (3) | 0.08 (3) |
| $H_4$ | −0.33 (2) | 0.264 (8) | 0.316 (3) | 0.07 (2) | $H_{8b}$ | −0.12 (3) | 0.469 (9) | 0.574 (5) | 0.08 (3) |
| | −0.32 (2) | 0.245 (7) | 0.305 (3) | 0.08 (3) | | −0.13 (1) | 0.483 (6) | 0.562 (3) | 0.04 (2) |
| $H_{5a}$ | 0.05 (2) | 0.218 (8) | 0.380 (4) | 0.02 (2) | $H_{9a}$ | −0.30 (1) | 0.720 (6) | 0.634 (2) | 0.03 (1) |
| | 0.03 (2) | 0.218 (7) | 0.383 (3) | 0.08 (3) | | −0.32 (2) | 0.714 (6) | 0.648 (3) | 0.05 (2) |
| $H_{5b}$ | −0.26 (2) | 0.181 (9) | 0.406 (4) | 0.05 (2) | $H_{9b}$ | −0.41 (2) | 0.569 (8) | 0.663 (4) | 0.06 (2) |
| | −0.22 (2) | 0.165 (7) | 0.403 (3) | 0.05 (3) | | −0.42 (2) | 0.553 (9) | 0.654 (5) | 0.16 (4) |
| $H_{6a}$ | −0.59 (1) | 0.604 (6) | 0.432 (3) | 0.04 (1) | $H_{10a}$ | 0.17 (2) | 0.474 (9) | 0.719 (5) | 0.07 (3) |
| | −0.59 (2) | 0.613 (9) | 0.441 (4) | 0.11 (3) | | 0.14 (2) | 0.453 (11) | 0.722 (5) | 0.19 (5) |
| $H_{6b}$ | −0.26 (2) | 0.616 (9) | 0.451 (5) | 0.00 (1) | $H_{1'}$ | 0.09 (2) | 0.322 (9) | 0.267 (4) | 0.04 (2) |
| | −0.27 (2) | 0.645 (7) | 0.433 (3) | 0.06 (3) | | 0.03 (2) | 0.324 (8) | 0.264 (4) | 0.11 (3) |
| $H_{7a}$ | −0.57 (2) | 0.526 (8) | 0.549 (3) | 0.06 (2) | $H_{3'}$ | −0.34 (2) | 0.637 (8) | 0.329 (3) | 0.07 (2) |
| | −0.64 (2) | 0.553 (9) | 0.536 (4) | 0.13 (4) | | −0.29 (2) | 0.624 (6) | 0.323 (3) | 0.04 (2) |

[a] Parameters for I are followed by parameters for II. The standard deviation in the least significant figure of a parameter as estimated from the least-squares procedure (see section on comparison of the structural results), is given in parentheses. The temperature factor expression is $T = \exp\{-2\pi^2[U(1,1)a^{*2}h^2 + U(2,2)b^{*2}k^2 + U(3,3)c^{*2}l^2 + 2U(1,2)a^*b^*hk + 2U(1,3)a^*c^*hl + 2U(2,3)b^*c^*kl]\}$.

plane fit by least squares to the ureido atomic positions is 0.03 Å for $N_{3'}$. The bond distances and angles within the ring are similar to those observed by Wheatley[14] in ethylenethiourea, which is also observed to be essentially planar. The ureido carbonyl bond, 1.25 Å, is significantly longer than the average 1.21 Å found for the comparable bond in 12 barbiturates studied by Craven, Cusatis, Gartland, and Vizzini,[15] and approaches the very long 1.27-Å value found in urea by Caron and Donohue.[16] Similarly the carbonyl carbon–nitrogen bond distances, average 1.34 Å, are shorter than the average 1.37 Å found for the comparable bonds in the barbiturates, and approach the 1.33-Å urea C–N bond value. Since the standard deviations in the barbiturate C=O bond and C–N bond lengths, as calculated from the 12 and 24 measurements, respectively, are quite small, $\sigma_{C=O} = 0.006$ Å and $\sigma_{C-N} = 0.012$ Å, the observed ureido carbonyl bond lengthening and carbonyl C–N bond shortening are considered to be real effects of particular significance. It is noteworthy that the ureido carbonyl oxygen accepts a strong (O · · · O separation 2.54 Å) hydrogen bond.

**Tetrahydrothiophene Ring.** The tetrahydrothiophene ring is envelope shaped with the sulfur atom 0.87 Å out of the plane fit by least squares to the positions of the four carbon atoms; the maximum deviation from this plane is 0.02 Å for both $C_3$ and $C_4$. The sulfur atom extends from the tetrahydrothiophene ring so as to minimize the S · · · $C_2$, transannular separation, 3.68 Å. The dihedral angle of 42.4° defines the intersection of the four-atom plane $C_2$–$C_3$–$C_4$–$C_5$ and the three-atom $C_5$–S–$C_2$ plane. The C–S bond lengths, 1.82 and 1.81 Å, are within the range of C–S bond lengths, 1.80 (3) Å, observed by Girling and Jeffrey[17] for the thioether bonds of the methyl thioribopyranosides. The C–S–C bond angle has contracted from the C–Se–C value of 93° observed by Hope and McCullough[18] in the tetrahydroselenophene iodine complex to 89.4° in biotin.

**Symmetry of the 2′-Keto-3,4-imidazolidotetrahydrothio-**

**Table III.** Interatomic Bond Distances and Angles

| | I | II | Av[a] | | I | II | Av[b] |
|---|---|---|---|---|---|---|---|
| $S_1-C_2$ | 1.835 (5) Å | 1.811 (5) Å | 1.823 Å | $C_2-H_2$ | 0.91 (5) Å | 0.94 (9) Å | 0.93 Å |
| $S_1-C_5$ | 1.819 (5) | 1.794 (5) | 1.807 | $C_3-H_3$ | 0.95 (10) | 0.85 (10) | 0.90 |
| $C_2-C_3$ | 1.519 (6) | 1.542 (6) | 1.531 | $C_4-H_4$ | 0.88 (9) | 1.09 (8) | 0.99 |
| $C_2-C_6$ | 1.508 (6) | 1.512 (7) | 1.510 | $C_5-H_{5a}$ | 0.94 (10) | 0.82 (10) | 0.88 |
| $C_3-C_4$ | 1.548 (6) | 1.548 (7) | 1.548 | $C_5-H_{5b}$ | 1.08 (10) | 1.04 (8) | 1.06 |
| $C_3-N_{3'}$ | 1.454 (6) | 1.438 (7) | 1.446 | $N_{1'}-H_{1'}$ | 1.00 (10) | 0.88 (9) | 0.94 |
| $C_4-C_5$ | 1.514 (6) | 1.557 (7) | 1.536 | $N_{3'}-H_{3'}$ | 1.11 (9) | 0.90 (7) | 1.00 |
| $C_4-N_{1'}$ | 1.452 (6) | 1.466 (7) | 1.459 | $C_6-H_{6a}$ | 1.10 (6) | 1.07 (10) | 1.09 |
| $N_{1'}-C_{2'}$ | 1.338 (5) | 1.326 (7) | 1.332 | $C_6-H_{6b}$ | 0.91 (10) | 1.13 (9) | 1.02 |
| $C_{2'}-O_{2'}$ | 1.250 (6) | 1.247 (6) | 1.249 | $C_7-H_{7a}$ | 1.03 (9) | 1.08 (10) | 1.06 |
| $C_{2'}-N_{3'}$ | 1.352 (7) | 1.350 (7) | 1.351 | $C_7-H_{7b}$ | 0.88 (10) | 1.02 (6) | 0.95 |
| $C_6-C_7$ | 1.549 (6) | 1.546 (6) | 1.548 | $C_8-H_{8a}$ | 0.97 (9) | 1.08 (10) | 1.03 |
| $C_7-C_8$ | 1.525 (7) | 1.540 (7) | 1.533 | $C_8-H_{8b}$ | 1.24 (10) | 1.07 (6) | 1.15 |
| $C_8-C_9$ | 1.538 (6) | 1.538 (6) | 1.538 | $C_9-H_{9a}$ | 1.02 (6) | 1.01 (7) | 1.02 |
| $C_9-C_{10}$ | 1.487 (7) | 1.510 (7) | 1.499 | $C_9-H_{9b}$ | 1.18 (10) | 1.24 (10) | 1.21 |
| $C_{10}-O_{10a}$ | 1.300 (6) | 1.298 (6) | 1.299 | $O_{10a}-H_{10a}$ | 0.97 (11) | 0.95 (11) | 0.96 |
| $C_{10}-O_{10b}$ | 1.218 (6) | 1.195 (7) | 1.207 | | | | |

| | I | II | Av[c] | | I | II | Av |
|---|---|---|---|---|---|---|---|
| $S_1-C_2-C_3$ | 103.9 (3)° | 105.2 (3)° | 104.6° | $C_5-C_4-N_{1'}$ | 113.8 (3)° | 112.2 (4)° | 113.0° |
| $S_1-C_2-C_6$ | 116.5 (3) | 116.6 (3) | 116.6 | $C_4-N_{1'}-C_{2'}$ | 112.2 (3) | 112.8 (3) | 112.5 |
| $C_2-S_1-C_5$ | 89.1 (2) | 89.8 (2) | 89.4 | $N_{1'}-C_{2'}-O_{2'}$ | 126.7 (3) | 127.0 (3) | 126.9 |
| $S_1-C_5-C_4$ | 106.3 (3) | 105.9 (3) | 106.2 | $N_{1'}-C_{2'}-N_{3'}$ | 109.8 (3) | 109.1 (3) | 109.5 |
| $C_2-C_3-C_4$ | 110.7 (3) | 108.7 (4) | 109.7 | $O_{2'}-C_{2'}-N_{3'}$ | 123.5 (3) | 123.9 (3) | 123.7 |
| $C_2-C_3-N_{3'}$ | 114.1 (3) | 112.8 (4) | 113.5 | $C_6-C_7-C_8$ | 114.3 (3) | 113.4 (4) | 113.9 |
| $C_3-C_2-C_6$ | 112.4 (3) | 111.4 (4) | 111.9 | $C_7-C_8-C_9$ | 110.4 (3) | 109.9 (4) | 110.2 |
| $C_2-C_6-C_7$ | 117.3 (4) | 117.0 (4) | 117.2 | $C_8-C_9-C_{10}$ | 111.6 (3) | 110.1 (4) | 110.9 |
| $C_3-C_4-C_5$ | 108.5 (3) | 108.8 (4) | 108.6 | $C_9-C_{10}-O_{10a}$ | 113.3 (3) | 113.9 (4) | 113.6 |
| $C_3-C_4-N_{1'}$ | 103.2 (3) | 102.0 (3) | 102.6 | $C_9-C_{10}-O_{10b}$ | 124.7 (3) | 123.0 (4) | 123.9 |
| $C_4-C_3-N_{3'}$ | 102.8 (3) | 103.1 (3) | 103.0 | $O_{10a}-C_{10}-O_{10b}$ | 121.9 (3) | 123.0 (4) | 122.5 |
| $C_3-N_{3'}-C_{2'}$ | 111.9 (3) | 112.7 (3) | 112.3 | | | | |

Estimates for the standard deviations (see comparison of structural results section) are *a*, $\sigma_{BL} \simeq 0.011$ Å; *b*, $\sigma'_{BL} \simeq 0.19$ Å; *c*, $\sigma_{BA} \simeq 0.6°$.

**Table IV.** Torsion Angles[a] about the Nonhydrogen Bonds

| | I | II | Av | | I | II | Av |
|---|---|---|---|---|---|---|---|
| $C_5-S_1-C_2-C_3$ | −41.4° | −42.5° | −42.0° | $C_4-C_3-N_{3'}-C_{2'}$ | 2.3° | 4.3° | 3.3° |
| $C_5-S_1-C_2-C_6$ | −165.6 | −166.5 | −166.0 | $C_3-C_4-C_5-S_1$ | −28.0 | −27.4 | −27.7 |
| $C_2-S_1-C_5-C_4$ | 40.9 | 40.6 | 40.8 | $N_{1'}-C_4-C_5-S_1$ | 86.2 | 84.7 | 85.5 |
| $S_1-C_2-C_3-C_4$ | 31.7 | 32.1 | 31.9 | $C_3-C_4-N_{1'}-C_{2'}$ | 0.4 | −0.6 | −0.1 |
| $S_1-C_2-C_3-N_{3'}$ | −83.7 | −81.6 | −82.7 | $C_5-C_4-N_{1'}-C_{2'}$ | −116.9 | −116.8 | −116.9 |
| $C_6-C_2-C_3-C_4$ | 158.5 | 159.3 | 158.9 | $C_4-N_{1'}-C_{2'}-O_{2'}$ | −177.8 | −176.2 | −177.0 |
| $C_6-C_2-C_3-N_{3'}$ | 43.1 | 45.6 | 44.4 | $C_4-N_{1'}-C_{2'}-N_{3'}$ | 1.0 | 3.3 | 2.2 |
| $S_1-C_2-C_6-C_7$ | −68.3 | −70.0 | −69.1 | $N_{1'}-C_{2'}-N_{3'}-C_3$ | −2.2 | −4.9 | −3.6 |
| $C_3-C_2-C_6-C_7$ | 172.0 | 169.3 | 171.7 | $O_{2'}-C_{2'}-N_{3'}-C_3$ | 176.8 | 174.6 | 175.7 |
| $C_2-C_3-C_4-C_5$ | −2.8 | −3.2 | −3.0 | $C_2-C_6-C_7-C_8$ | 73.5 | 74.3 | 73.9 |
| $C_2-C_3-C_4-N_{1'}$ | −123.8 | −122.0 | −122.9 | $C_6-C_7-C_8-C_9$ | 168.1 | 166.2 | 167.2 |
| $N_{3'}-C_3-C_4-C_5$ | 119.5 | 116.6 | 118.1 | $C_7-C_8-C_9-C_{10}$ | 176.1 | 175.4 | 175.8 |
| $N_{3'}-C_3-C_4-N_{1'}$ | −1.5 | −2.1 | −1.8 | $C_8-C_9-C_{10}-O_{10a}$ | −70.7 | −70.0 | −70.4 |
| $C_2-C_3-N_{3'}-C_{2'}$ | 122.2 | 121.3 | 121.8 | $C_8-C_9-C_{10}-O_{10b}$ | 110.5 | 112.0 | 111.3 |

[a] Torsion angle defined after the Klyne–Prelog convention.

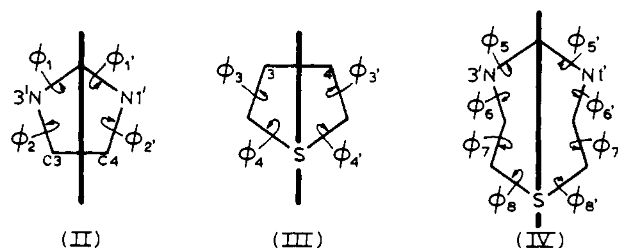
**phene Moiety.** The bicyclic ring system possesses nearly perfect mirror symmetry about the plane passing through the sulfur and carbonyl carbon and normal to the $C_3-C_4$ bond. A measure of the mirror asymmetry, or departure from exact symmetry, of a ring system has been described by Duax and Rohrer.[19] The asymmetry parameters for the ureido ring (II), tetrahydrothiophene ring (III), and the bicyclic ring (IV) are respectively given by

$$\Phi_U = \sqrt{[(\phi_1 + \phi_{1'})^2 + (\phi_2 + \phi_{2'})^2]/2} = 2.5°$$

$$\Phi_T = \sqrt{[(\phi_3 + \phi_{3'})^2 + (\phi_4 + \phi_{4'})^2]/2} = 3.1°$$

$$\Phi_B = \sqrt{[(\phi_5 + \phi_{5'})^2 + (\phi_6 + \phi_{6'})^2 + (\phi_7 + \phi_{7'})^2 + (\phi_8 + \phi_{8'})^2]/4} = 3.0°$$

where the individual $\phi$'s are the internal ring torsion angles. Torsion angles are collected in Table IV. A value of $\Phi = 0°$




is necessary for exact mirror symmetry while values of $\Phi$ in the range 15-35° describe a highly asymmetric ring with respect to the mirror plane of interest.

The dihedral angle of 122° describes the intersection of the six atoms of the ureido portion and the four carbon atoms of the tetrahydrothiophene ring, atoms $C_3$ and $C_4$ being common to both. The sulfur atom assumes the endo configuration with respect to the bicyclic ring system. Model building shows quite clearly that if the sulfur atom

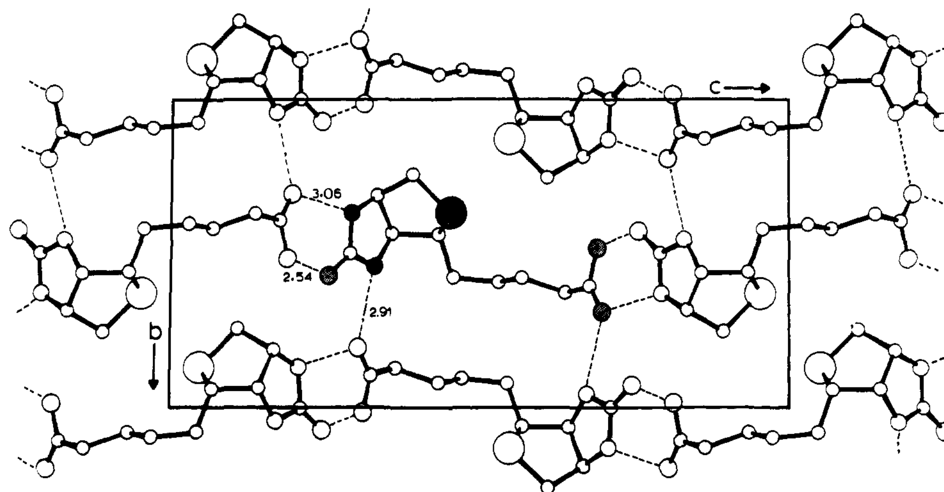



**Figure 3.** The crystal packing and hydrogen bonding of *d*-biotin viewed along the *a* axis. Bold lines are covalent bonds; dashed lines are hydrogen bonds. The positional parameters, Table II, of 1, employed to compose the drawing, refer to the molecule with the oxygen atoms (light hatch), sulfur atom (dark hatch), and nitrogen atoms (black) indicated. Hydrogen bonding donor–acceptor relationships for *d*-biotin:

| Donor | Acceptor | Coordinates of acceptor |
|---|---|---|
| $N_1'$ | $O_{10b}$ | $\frac{1}{2} - x, 1 - y, -\frac{1}{2} + z$ |
| $N_3'$ | $O_{10b}$ | $-\frac{1}{2} + x, \frac{3}{2} - y, 1 - z$ |
| $O_{10a}$ | $O_2'$ | $\frac{1}{2} - x, 1 - y, \frac{1}{2} + z$ |

The position of the acceptor is related to its position in Table II by the equations given. The position of the donor is given in Table II.

were in the exo configuration, the valeryl $C_6$ methylene group would swing in even closer to the $N_{3'}$ nitrogen (see next section) if bicyclic mirror plane symmetry is to be preserved. Calculations assuming that mirror plane symmetry is preserved and the sulfur atom is in the exo configuration indicate a $C_6 \cdots N_{3'}$ nonbonded contact of approximately 2.75 Å.

**Valeryl Chain.** The valeryl chain is severely twisted from the maximally extended all-trans conformation, Figure 2 and Table IV. Bond distances and angles are within normal ranges except for the $C_2$–$C_6$–$C_7$ bond angle which has opened up to 117.2°. This may have been expected to be due to the close $S \cdots C_7$ nonbonded contact, 3.48 Å, which is considerably shorter than the sum, 3.85 Å, of the sulfur-methylene van der Waals radii.[20] However, this simple rationale does not explain why $C_7$ is not twisted about the $C_2$–$C_6$ bond away from the sulfur atom, assuming that rotation about a C–C single bond would be more energetically favorable than opening a C–C–C bond angle; nor does it explain why the $S \cdots C_8$ nonbonded contact, 3.34 Å, is even shorter than the $S \cdots C_7$ contact. There is another particularly close approach, between the valeryl $C_6$ and the ureido $N_{3'}$ (2.86 Å), which would be expected to hinder close approach from solution to the side of the ureido ring proximal to the valeryl chain. The carboxylic acid group is planar, with the carboxylic carbon atom deviating 0.01 Å from the plane defined by the $O_{10a}$, $O_{10b}$, and $C_9$ positions.

**Hydrogen Bonding.** All potential hydrogen bond donors are utilized in the packing arrangement, Figure 3. The ureido $N_{1'}$ and $O_{2'}$ and a neighboring carboxylic acid group form a $\{\cdots O_{2'}\text{–}C_{2'}\text{–}N_{1'}H \cdots O_{10b}\text{=}C_{10}\text{–}O_{10a}H \cdots\}$ hydrogen bonded ring. The acid function carbonyl oxygen $O_{10b}$ accepts a second hydrogen bond from another neighboring $N_{3'}H$ group. The hydrogen bonding is three-dimensional. The sulfur atom, known to be protonated in strong acid,[21] does not participate in the hydrogen bonding. There are no sulfur–sulfur nonbonded approaches which are smaller than the sum of van der Waals radii.[20] Donor–acceptor relationships and distances are indicated in Figure 3. The $O_{2'} \cdots O_{10a}$ hydrogen bond distance, 2.54 Å, is indicative of a strong hydrogen bond.

**Discussion**

The details of the biotin–$HCO_3^-$ reaction mechanism are not well understood. Employing 2-imidazolidone (ethyleneurea) as a biotin model, Caplow[22] showed that the ureido group is a poor nucleophile in bimolecular reactions and concluded that enzymatic activation of biotin for nucleophilic attack on $HCO_3^-$ is probably necessary. Hegarty, Bruice, and Benkovic[23] suggested that the enol form of biotin would be a highly active nucleophile. In a series of later papers Bruice and Hegarty[24] explored the possibility of ureido oxygen vs. nitrogen attack and concluded, from studies on compounds which bear little resemblance to biotin, that oxygen attack upon bicarbonate was the more likely. At the time only inconclusive evidence favoring ureido 1′-N attack was available: trapping experiments in which enzyme-bound carboxylated biotin was treated with diazomethane yielded 1′-*N*-methoxycarbonyl biocytin upon proteolytic digest.[2] Bruice et al. argued that, under these conditions, *N*-carboxybiotin derivatives would be isolated whether the initial product of the biotin–$CO_2$ reaction was *N*-carboxy- or *O*-carboxybiotin, since *O*-carboxybiotin would undergo a rapid O → N acyl shift. However, recent evidence in favor of ureido 1′-N attack has been obtained by Guchhait, Moss, Lane, and co-workers[25] who have purified the enzymatic components of the acetyl coenzyme A carboxylase system of *Escherichia coli*. Employing the purified enzymes, they demonstrated that chemically synthesized 1′-*N*-carboxybiotin derivatives can serve as substrate in both the reverse biotin carboxylase and carboxyltransferase half-reactions catalyzed by the system. In light of these results the inviting theory of ureido oxygen attack as proposed by Bruice and Hegarty[24] seems untenable.

It has been suggested by Griesser, Sigel, Wright, and McCormick[26] that nucleophilic activation of biotin would result from a polarization of the ureido carbonyl bond via hydrogen bond formation, with an attendant increase in double bond character of the carbonyl carbon–nitrogen bonds and a concomitant increase in NH acidity. Under these conditions the NH groups would be more easily deprotonated, yielding a highly active biotin anion. Traub[3b]

had suggested that an intramolecular hydrogen bond donated by the carboxylic acid function to the ureido oxygen might displace the keto-enol equilibrium toward the enol and affect the reactivity of the nitrogen atoms.[23] It is now known that such an intramolecular hydrogen bond is not possible in the normal biochemical situation since biotin is covalently bound through an amide linkage to a lysyl $\epsilon$-amino group in all of the biotin-dependent enzymes studied to date.[2] However, the same effects could be obtained by the formation of a hydrogen bond between the biotin ureido oxygen and an amino acid donor in the active site. Whether the effect of a strong hydrogen bond donated to the ureido oxygen is to promote the formation of a biotin anion as suggested by Griesser et al.[26] or the tautomerization of biotin to the enol form as suggested by Traub[3b] and Hegarty et al.,[23] in either case the net result is the activation of the nitrogen atom for bicarbonate attack. Infrared studies by Griesser et al.[26] confirm the formation of a hydrogen bond donated by phenol to the ureido oxygen of 2-imidazolidone in chloroform solution. Additional results supportive of such a model have been obtained in the present study. The lengthening of the carbonyl bond, whose oxygen atom accepts a strong hydrogen bond, the shortening of the carbonyl carbon-nitrogen bonds, and the overall planarity of the ureido ring are effects consistent with a delocalization of electronic charge within the ring in response to a hydrogen bond polarization of the carbonyl bond.

Theoretical studies of the electronic structure of biotin[27a] and various forms of carboxybiotin[27b] by Maggiora explored these questions of biotin tautomerism and activation by hydrogen bond polarization. In both cases the results were unfavorable. However, it is now apparent that the model employed for the structure of biotin in these studies is inadequate. Most importantly, it does not reflect the observed effects in the ureido group, which is the most important moiety of the molecule in terms of biological function. It is unclear whether the results of the CNDO/2 calculations of population overlap and atomic net charge will be greatly altered by the new model presented herein, but it does seem clear that the activation via polarization mechanism cannot be dismissed until further calculations are performed.

Comparison of the structure of biotin with that of 1′-*N*-carboxybiotin, as observed in its bis-*p*-bromoanilide derivative by Bonnemere et al.,[5] gives a picture of the effects of carboxylation upon the ureido ring. Although the precision of the 1′-*N*-carboxybiotin bis-*p*-bromoanilide structure determination is uncertain,[28] a qualitative comparison is possible. Upon carboxylation the ureido carbonyl bond shortens and the carbonyl carbon-nitrogen bonds lengthen to values more closely approximating those found in the barbiturates. The ureido remains essentially planar. This suggests that the planarity of the ureido ring in biotin may be more a consequence of its incorporation into the bicyclic 2′-keto-3,4-imidazolidotetrahydrothiophene ring system than the effect of a delocalization of charge within the ureido ring. (Additional evidence that the tetrahydrothiophene ring may be necessary for the planarity of the ureido ring comes from the crystal structural study of *dl*-dethiobiotin by Chen, Parthasarathy, and DeTitta.[29] In the dethio analogue of biotin the ureido ring is puckered.) It also suggests that the delocalization of charge which may be critical for nucleophilic activation of biotin is no longer necessary once biotin has been carboxylated. It should be emphasized that, because of the uncertain precision of the 1′-*N*-carboxybiotin bis-*p*-bromoanilide determination, conclusions of this nature are tentative.

It has been also suggested that a polarization of the ureido carbonyl bond might be a consequence of a transannular sulfur-carbonyl carbon interaction. Calculations by Bowen, Rauscher, and Ingraham[30] indicate little or no overlap potential for transannular S $\cdots$ $C_{2'}$ separations of 3.5 Å or greater, which eliminates this through-space mechanism (S $\cdots$ $C_{2'}$ distance 3.68 Å) as a possible activation mechanism.

In summary, a highly precise model of the structure of biotin has been obtained. Results supportive of a proposed activation via carbonyl bond polarization mechanism are observed. Of primary importance are effects suggestive of ureido ring delocalization, chiefly the significant lengthening of the carbonyl bond and shortening of the carbonyl carbon-nitrogen bonds, in response to participation of the carbonyl oxygen in a strong hydrogen bond. Until the structure of the active site of a biotin-dependent enzyme system has been elucidated, corroborative evidence of such a hydrogen bond formation is awaited.

**Acknowledgments.** I. We wish to thank Dr. Jane F. Griffin for her helpful discussions of the structural results and Dr. Douglas Rohrer for his supervision of the data collection. This project was supported in part by USPH Grant HL-15378 from the National Heart and Lung Institute. II. We wish to thank Mr. Henry Katz for his technical assistance. This work was supported in part by the National Science Foundation and the National Institutes of Health.

**Supplementary Material Available:** A listing of observed and calculated structure factor amplitudes for both determinations (8 pages). Ordering information is given on any current masthead page.

## References and Notes

(1) Present address: (a) Dow Chemical U.S.A., Midland, Mich. 48640; (b) The Institute for Cancer Research, The Fox Chase Cancer Center, 7701 Burholme Ave., Philadelphia, Pa. 19111.
(2) J. Moss and M. D. Lane, *Adv. Enzymol. Relat. Areas Mol. Biol.*, **35**, 321 (1971). This is an excellent review of the biotin literature.
(3) (a) W. Traub, *Nature (London)*, **178**, 649 (1956). The crystal structure of biotin had been determined in projection. Positional parameters, bond distances, and angles were not reported. The results of this determination have not, to our knowledge, subsequently appeared. (b) W. Traub, *Science*, **129**, 210 (1959).
(4) C. A. Grob and H. von Sprecher, *Helv. Chim. Acta*, **35**, 885 (1952).
(5) C. Bonnemere, J. A. Hamilton, L. K. Steinrauf, and J. Knappe, *Biochemistry*, **4**, 240 (1965).
(6) J. Trotter and J. A. Hamilton, *Biochemistry*, **5**, 713 (1966).
(7) G. Germain, P. Main, and M. M. Woolfson, *Acta Crystallogr., Sect. A*, **27**, 368 (1971).
(8) D. T. Cromer and J. T. Waber, "International Tables for Crystallography", Vol. IV, Kynoch Press, Birmingham, England, 1974, p 71.
(9) R. F. Stewart, E. R. Davidson, and W. T. Simpson, *J. Chem. Phys.*, **42**, 3175 (1965).
(10) D. T. Cromer, *Acta Crystallogr.*, **18**, 17 (1965). More recent values for $\Delta f'$ and $\Delta f''$ have been compiled by Cromer; see ref 8, p 148. Sulfur values, Cu K$\alpha$ radiation, are $\Delta f' = 0.319$ and $\Delta f'' = 0.517$.
(11) P. Coppens, L. Leiserowitz, and D. Rabinovich, *Acta Crystallogr.*, **18**, 1035 (1965).
(12) (a) C. R. Hubbard and R. A. Jacobson, "A Fortran IV Crystallographic System of Programs for Generalized Superpositions", Ames Laboratory, Iowa State University of Science and Technology; (b) P. Gantzel, R. A. Sparks, K. N. Trueblood, and R. E. Long, "UCLALS4: A Fortran IV Version of A.C.A. Program No. 232", University of California, Los Angeles, 1965.
(13) (a) S. C. Abrahams and E. T. Keve, *Acta Crystallogr., Sect. A*, **27**, 157 (1971); (b) E. W. Hughes, "Structural Chemistry and Molecular Biology", A. Rich and N. Davidson, Ed., W. H. Freeman, San Francisco, Calif., 1968.
(14) P. J. Wheatley, *Acta Crystallogr.*, **6**, 369 (1953).
(15) B. M. Craven, C. Cusatis, G. L. Gartland, and E. A. Vizzini, *J. Mol. Struct.*, **16**, 331 (1973).
(16) A. Caron and J. Donohue, *Acta Crystallogr., Sect. B*, **25**, 404 (1969).
(17) R. L. Girling and G. A. Jeffrey, *Acta Crystallogr., Sect. B*, **30**, 327 (1974).
(18) H. Hope and J. D. McCullough, *Acta Crystallogr.*, **17**, 712 (1964).
(19) (a) W. L. Duax and D. C. Rohrer, American Crystallographic Association Summer Meeting, University of Connecticut, Storrs, Conn., June 17–22, 1973, p 166. "Atlas of Steroid Structure", Vol. 1, W. L. Duax and D. A. Norton, Ed., Plenum Press, New York, N.Y., 1975.
(20) L. Pauling, "The Nature of the Chemical Bond", 3d ed, Cornell University Press, Ithaca, N.Y., 1960.
(21) (a) G. A. Olah and A. M. White, *J. Am. Chem. Soc.*, **90**, 6087 (1968); (b) R. Griesser and D. B. McCormick, *Arch. Biochem. Biophys.*, **160**, 667 (1974).

(22) M. Caplow, *J. Am. Chem. Soc.*, **87**, 5774 (1965).
(23) A. F. Hegarty, T. C. Bruice, and S. J. Benkovic, *Chem. Commun.*, 1173 (1969).
(24) (a) T. C. Bruice and A. F. Hegarty, *Proc. Natl. Acad. Sci. U.S.A.*, **65**, 805 (1970); (b) A. F. Hegarty and T. C. Bruice, *J. Am. Chem. Soc.*, **92**, 6561 (1970); (c) *ibid.*, **92**, 6568 (1970); (d) *ibid.*, **92**, 6575 (1970); (e) A. F. Hegarty, R. F. Pratt, T. Giudici, and T. C. Bruice, *ibid.*, **93**, 1428 (1971).
(25) (a) R. B. Guchhait, S. E. Polakis, P. Dimroth, E. Stoll, J. Moss, and M. D. Lane, *J. Biol. Chem.*, **249**, 6633 (1974); (b) R. B. Guchhait, S. E. Polakis, D. Hollis, C. Fenselaw, and M. D. Lane, *ibid.*, **249**, 6646 (1974); (c) S. E. Polakis, R. B. Guchhait, E. E. Zwergel, and M. D. Lane, *ibid.*, **249**, 6657 (1974).
(26) R. Griesser, H. Sigel, L. D. Wright, and D. B. McCormick, *Biochemistry*, **12**, 1917 (1973).
(27) (a) G. M. Maggiora, *J. Theor. Biol.*, **41**, 523 (1973); (b) *ibid.*, **43**, 51 (1974).
(28) The stated aim of the authors, ref 5, was the unequivocal identification of the major product of the reaction of biotin methyl ester with methyl chloroformate as 1′-*N*-methoxycarbonyl biotin methyl ester. Because the only interest was the elucidation of the structure, no special precautions were taken in the data collection. Some significant discrepancies between bond lengths as calculated from the positional parameters, Table I, and as listed in Table II, ref 5, exist.
(29) C. S. Chen, R. Parthasarathy, and G. T. DeTitta, *J. Am. Chem. Soc.*, in press.
(30) C. E. Bowen, E. Rauscher, and L. L. Ingraham, *Arch. Biochem. Biophys.*, **125**, 865 (1968).

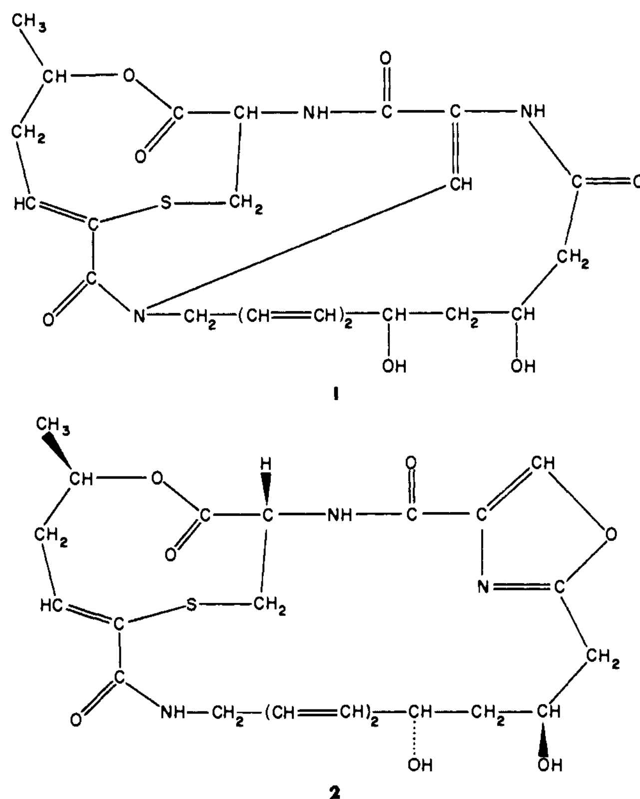
# Structure of the Antibiotic Griseoviridin[1]

**George I. Birnbaum*[2a] and Sydney R. Hall[2b]**

*Contribution from the Division of Biological Sciences, National Research Council of Canada, Ottawa, Canada K1A 0R6, and Department of Energy, Mines and Resources, Ottawa, Canada K1A 0G1.* *Received July 21, 1975*

**Abstract:** The three-dimensional structure of griseoviridin methanolate, $C_{22}H_{27}O_7N_3S{\cdot}CH_3OH$, was determined by x-ray crystallography. The substance crystallizes in the monoclinic space group $P2_1$ and the unit cell dimensions are $a$ = 10.559 (1), $b$ = 13.066 (3), and $c$ = 9.433 (1) Å, $\beta$ = 99.00 (1)°. Intensity data were collected with a diffractometer and the structure was solved by a combination of Patterson superposition and tangent refinement techniques. Contrary to a previous proposal the structure contains an oxazole ring. The relative contributions of three canonical structures are assessed on the basis of bond lengths in that ring. It is estimated that the uncharged species contributes 80% to the structure while delocalizations of oxygen electrons to the ring nitrogen and to a conjugated amide oxygen account for an additional 10% each. The molecular structure is quite rigid and most atoms lie in one of four major planes. There are several short intramolecular contacts which give rise to some unusual bond angles and torsion angles.

Griseoviridin is a broad-spectrum antibiotic with inhibitory activity in vitro toward various pathogenic bacteria and fungi.[3] It has also been shown to inhibit bacterial protein synthesis in cell-free systems and poly-U-directed poly(phenylalanine) synthesis in yeast and in human tonsil cell-free systems.[4] It was isolated from *Streptomyces griseus*[5] and its chemical structure has been the subject of intensive investigations by Ames, Bowman, and their colleagues[6-8] as well as by de Mayo, Stoessl, and their collaborators.[9-12] On the basis of degradative and spectroscopic studies the latter group proposed **1** for the structure of griseoviridin.[11] An x-ray analysis was undertaken in order to establish unambiguously the structure and complete stereochemistry. We report here the results of this work which revealed that the correct structure is **2**. It should be pointed out that the Canadian workers considered the skeleton of **2** a plausible one but rejected it in favor of **1**.[11]



## Experimental Section

Colorless crystals of griseoviridin methanolate, $C_{22}H_{27}O_7N_3S{\cdot}CH_3OH$, were sent to us by Drs. A. Stoessl and P. de Mayo. Precession photographs showed them to be monoclinic and the space group $P2_1$ was indicated by systematic absences of $0k0$ reflections for $k$ odd. A crystal with dimensions 0.29 × 0.29 × 0.14 mm was mounted on a Picker four-circle diffractometer equipped with a Mo target and a graphite monochromator. The cell parameters were obtained from a least-squares fit of the diffractometer angles $2\theta$, $\phi$, and $\chi$ for 42 independent reflections, using Mo $K\alpha_1$ radiation ($\lambda$ 0.70926 Å) and assuming a triclinic cell. The crystal data are as follows: $a$ = 10.559 (1), $b$ = 13.066 (3), and $c$ = 9.433 (1) Å; $\beta$ = 99.00 (1)°; $V$ = 1285.4 Å$^3$; $d_x$ = 1.316 g cm$^{-3}$; $Z$ = 2; $F(000)$ = 540; $\mu$ (Mo $K\alpha$) = 1.8 cm$^{-1}$.

The moving-crystal-moving-counter technique ($\theta$-$2\theta$ scan) was used to collect the intensity data. The $2\theta$ scan rate was 2° min$^{-1}$ and the scan width was varied according to the angular dispersion (2.5-3.0°). Background counts were measured for 30 s on each side of the scan. Three reference reflections were monitored every 50 measurements; their intensities decreased by 3.3% during the course of data collection. A net count of 18 was determined as threshold intensity below which reflections were considered unob-