Precision Neutron Diffraction Structure Determination of Protein and Nucleic Acid Components. I. The Crystal and Molecular Structure of the Amino Acid L-Alanine[1]

M. S. Lehmann,[2a] T. F. Koetzle,[2b] and W. C. Hamilton*

Contribution from the Chemistry Department, Brookhaven National Laboratory, Upton, New York 11973. Received July 19, 1971

Abstract: A neutron diffraction study of L-alanine, $C_3H_7NO_2$, has been carried out. The structure is orthorhombic, space group $P2_12_12_1$, with four molecules per unit cell; cell parameters $a = 6.032$ (1), $b = 12.343$ (1), and $c = 5.784$ (1) Å. A total of 786 symmetry-independent reflections was used in the least-squares refinement leading to a final conventional R value of 2.2%. The nonhydrogen atoms are located with a precision of 0.001–0.002 Å and the hydrogen atoms are located with a precision of 0.002–0.003 Å. The molecule is in the zwitterion form and the three hydrogen bonds N—H···O are nearly identical in length and angles. The average length of the C–H bond in the methyl group is 1.096 (3) Å and the average N–H bond length is 1.039 (6) Å. Average angles in the methyl group are C–C–H = 110.2 (1)° and H–C–H = 108.7 (2)°. The barriers to rotation for the methyl group and the ammonium group are estimated to be 5.6 and 20 kcal/mol, respectively. The molecules are linked together by hydrogen bonds to form a three-dimensional network. The methyl groups are situated in channels formed by this network.

Hydrogen bonds and van der Waals contacts between hydrogen atoms are two of the important factors determining the secondary and tertiary structures of polypeptide and polynucleotide chains. Precise hydrogen atom positions in these compounds and their monomers are unknown except in a few instances; the only general way of obtaining such information is by neutron diffraction. Present techniques allow the determination of hydrogen and other atom positions and root-mean-square amplitudes of vibration to a precision of better than 0.01 Å. In a series of neutron diffraction studies of the common amino acids, small peptides, nucleosides, and nucleotides we hope to provide not only definitive stereochemical information on hydrogen atom positions but also in most cases better framework geometry than is available from the prior X-ray studies. This information should be useful for all those concerned with the energetics and morphology of biomacromolecules.

In this paper we report the results of the first of these studies, the crystal and molecular structure of L-alanine. This structure has previously been determined by X-ray diffraction.[3,4]

Experimental Section

Crystal Growth and Data Collection. Crystals of L-alanine were easily grown by evaporation of aqueous solutions at room temperature. The crystals were prismatic with main faces [120] and [011]; the volume of the crystal used for the measurement was 7.8 mm³. Neutron diffraction data were collected at room temperature on the four-circle diffractometers located at the Brookhaven National Laboratory high-flux beam reactor using the computer-controlled Multi-Spectrometer control system.[5] Two sets of data were collected. On one of the diffractometers, intensities were measured for reflections with sin $\theta/\lambda < 0.54$ with a neutron wave-

length λ 1.248 Å; on the other diffractometer, intensities were measured for reflections with $0.48 < \sin \theta/\lambda < 0.70$ with a wavelength λ 1.013 Å. For both series of measurements the crystal was mounted with the c axis near the ϕ axis of the instrument; the crystal orientation as well as the unit cell constants were determined by a least-squares refinement of the setting angles observed for 27 strong reflections evenly distributed in reciprocal space. The reflections were measured with the θ–2θ step scan technique using a scan range $\Delta(2\theta) = 2.3°(1 + 3 \tan \theta)$. The step size was varied to give approximately 50 points in each scan. A substantial amount of time was saved during data collection by omitting the weakest reflections. The intensity of each reflection was estimated using a rapid peak-height scan, and reflections with intensity $I < 3\sigma(I)$ were skipped. To check the stability of the crystal and the detectors, two standard reflections were recorded for each 30 reflections; there was no deterioration in crystal quality during the data collection. Two sets of symmetry-related reflections (hkl) and $(-hkl)$ with $k \geq 0$ and $l \geq 0$ were measured, giving 712 and 794 reflections for λ 1.248 and 1.013 Å, respectively.

Results

Crystal Data. The neutron diffraction results confirm that the space group is $P2_12_12_1$. Unit cell constants given by Simpson and Marsh[3] are: $a = 6.032$ (1), $b = 12.343$ (1), $c = 5.784$ (1) Å. (Numbers in parentheses here and throughout this paper are estimated standard deviations in units of the last digit.) Unit cell constants determined in this study are $a = 6.025$ (7), $b = 12.324$ (14), $c = 5.783$ (6) Å, in agreement with the more precise values of Simpson and Marsh which are used in the subsequent discussion. The calculated density assuming four molecules of alanine per unit cell is 1.37 g/cm³. The neutron absorption coefficient was calculated to be 2.63 cm⁻¹ assuming an incoherent neutron cross section for hydrogen of 40 barns.

Data Reduction, Absorption Correction, Refinement. Background corrections were made by use of a method which divides the peak and the background in such a way that $\sigma(I)/I$ is minimized,[6] where I is the integrated intensity and $\sigma(I)$ is the standard deviation based on counting statistics. This is done by dividing the peak into two halves at the maximum counting number, and for each half calculating $\sigma(I)/I$ for every possible point

(1) Research performed under the auspices of the U. S. Atomic Energy Commission.

(2) (a) On leave from Kemisk Institut, Aarhus Universitet, DK 8000 Aarhus C, Denmark; (b) National Institutes of Health Postdoctoral Fellow.

(3) H. J. Simpson, Jr., and R. E. Marsh, *Acta Crystallogr.*, **20**, 550 (1966).

(4) J. D. Dunitz and R. R. Ryan, *ibid.*, **21**, 617 (1966).

(5) D. R. Beaucage, M. A. Kelley, D. Ophir, S. Rankowitz, R. J. Spinrad and R. van Norton, *Nucl. Instrum. Methods*, **40**, 26 (1966).

(6) M. S. Lehmann and F. K. Larsen, to be published.

Table I. Fractional Coordinates and Temperature Parameters[a]

Atom	x	y	z	u_{11}	u_{22}	u_{33}	u_{12}	u_{13}	u_{23}
O^1	0.72829 (21)	0.08451 (10)	0.62814 (22)	419 (6)	331 (5)	244 (5)	126 (5)	−50 (4)	17 (5)
O^2	0.45007 (22)	0.18505 (11)	0.76089 (18)	428 (7)	389 (6)	146 (4)	102 (5)	9 (5)	−26 (4)
N	0.65566 (11)	0.13796 (6)	0.18570 (11)	322 (4)	279 (4)	153 (3)	14 (3)	8 (2)	1 (2)
C	0.56049 (16)	0.14145 (7)	0.60153 (14)	310 (4)	203 (4)	142 (4)	15 (3)	−24 (3)	4 (3)
C^α	0.47674 (14)	0.16143 (7)	0.35533 (14)	281 (4)	211 (4)	142 (4)	26 (3)	−20 (3)	1 (3)
C^β	0.27468 (18)	0.09160 (10)	0.30245 (19)	329 (5)	427 (6)	246 (5)	−70 (4)	−40 (4)	−12 (4)
H^α	0.43571 (38)	0.24726 (16)	0.33760 (37)	537 (11)	300 (8)	324 (9)	109 (8)	−2 (9)	23 (7)
H^1	0.71070 (42)	0.05934 (19)	0.20033 (41)	508 (12)	383 (11)	382 (10)	107 (10)	66 (9)	18 (8)
H^2	0.59477 (36)	0.14894 (19)	0.01777 (34)	494 (11)	471 (11)	208 (8)	−17 (10)	−17 (7)	−6 (8)
H^3	0.78421 (37)	0.19174 (21)	0.21106 (38)	412 (11)	506 (12)	337 (10)	−65 (10)	27 (8)	−14 (9)
$H^{\beta 1}$	0.31887 (52)	0.00671 (22)	0.30574 (57)	743 (19)	400 (12)	657 (17)	−152 (13)	−29 (14)	−66 (12)
$H^{\beta 2}$	0.14588 (44)	0.10501 (29)	0.42954 (48)	454 (12)	880 (21)	504 (14)	−74 (13)	83 (11)	−42 (15)
$H^{\beta 3}$	0.21002 (49)	0.11058 (29)	0.13293 (49)	600 (14)	896 (22)	381 (12)	−158 (15)	−203 (11)	36 (14)

[a] The form of the anisotropic thermal ellipsoid is $\exp(-2\pi^2(u_{11}(ha^*)^2 + u_{22}(kb^*)^2 + u_{33}(lc^*)^2 + 2u_{12}ha^*kb^* + 2u_{13}ha^*lc^* + 2u_{23}kb^*lc^*))$. The u_{ij}'s are multiplied by 10^4.

of division between the peak and the background. The two points, one for each half, which minimize the $\sigma(I)/I$ determine the position of the peak. An absorption correction was carried out by integration in a Gaussian grid. The crystal shape was defined by seven rational planes, and the resulting transmission factors were in the range 0.63–0.69. Data for symmetry-equivalent reflections were averaged to give intensities for 406 and 407 independent reflections collected at λ 1.248 and 1.013 Å, respectively. Squared structure amplitudes were obtained as $F_o^2 = I \sin 2\theta$. The agreement factors ($\Sigma|F_o^2 - \bar{F}_o^2|/\Sigma F_o^2$, where $\bar{F}_o^2$ is the mean of the symmetry equivalent reflections) were 2.3 and 2.6% for the two sets of data. A total of 27 reflections having $F_o^2 < 3\sigma_{count}(F_o^2)$ was omitted from subsequent refinements. The starting parameters for the refinement were the coordinates and temperature factors given by Simpson and Marsh,[3] and the weights used were $w = 1/\sigma^2(F_o^2)$ with

$$\sigma^2(F_o^2) = \sigma^2{}_{count}(F_o^2) + (0.04F_o^2)^2$$

with $\sigma_{count}(F_o^2)$ based on counting statistics. This gave values of $||F_o^2| - |F_c^2||/\sigma^2(F_o^2)$ which were independent of F_o^2 when divided into classes of F_o^2, as expected for a proper weighting scheme. Neutron scattering lengths used were $b_O = 0.575$, $b_N = 0.94$, $b_C = 0.6626$, and $b_H = -0.372$ (10^{-12} cm). The quantity minimized by the refinement was $\Sigma w||F_o^2| - |F_c^2||^2$, where F_c is the calculated structure factor including extinction correction. An isotropic extinction parameter and two scale factors, one for each set of data, were included in the refinement. The extinction correction factor E,[7] which multiplies the calculated structure factor, is

$$E = \left[1 + \frac{(2\bar{T}F_c^2 g\lambda^3)(10^4)}{V^2 \sin 2\theta}\right]^{-1/4}$$

where g is the extinction parameter, $\bar{T}$ is the average beam path length in the crystal for a given reflection with Bragg angle θ, V is the unit cell volume, and F_c is the calculated structure factor on an absolute scale. Refinements were carried out until the shifts in the parameters were less than 0.01 of the estimated standard deviations. The final R value on F^2 ($\Sigma||F_o|^2 - |F_c|^2|/\Sigma|F_o|^2$) was 3.8% and the corresponding weighted R value [$(\Sigma w(|F_o|^2 - |F_c|^2)^2/\Sigma w|F_o|^4)^{1/2}$] was 5.2%. The conventional R factor ($\Sigma||F_o| - |F_c||/\Sigma|F_o|$) was 2.2%

(7) W. H. Zachariasen, *Acta Crystallogr.*, **16**, 1138 (1963).

and the error in an observation of unit weight was 1.48. The extinction parameter g was 0.050 (5), corresponding to values for E in the range 0.98–1.00, except for the two strongest reflections, where E was 0.84. The final values for the structure amplitudes are tabulated as $100|F_o|^2$, $100\sigma(F_o^2)$, and $100|F_c|^2$ for those reflections used in the refinement.[8] The computer programs used have been described briefly by Schlemper, Hamilton, and La Placa.[9]

Discussion

Atomic coordinates and thermal vibration parameters are given in Table I. Table II contains distances and

Table II. Distances (Å) and Angles (deg) within the Molecule

Bond distances		Bond angles	
C^α–C	1.531 (1)	C–C^α–N	110.05 (7)
C^α–N	1.487 (1)	C–C^α–C^β	111.07 (8)
C^α–C^β	1.524 (1)	C–C^α–H^α	108.55 (13)
C^α–H^α	1.093 (2)	N–C^α–C^β	109.74 (8)
C–O^1	1.242 (2)	N–C^α–H^α	106.94 (13)
C–O^2	1.258 (1)	C^β–C^α–H^α	110.40 (15)
N–H^1	1.029 (2)	C^α–C–O^1	118.39 (9)
N–H^2	1.047 (2)	C^α–C–O^2	115.97 (9)
N–H^3	1.031 (3)	O^1–C–O^2	125.65 (10)
C^β–$H^{\beta 1}$	1.081 (3)	C^α–N–H^1	111.33 (15)
C^β–$H^{\beta 2}$	1.082 (3)	C^α–N–H^2	109.40 (13)
C^β–$H^{\beta 3}$	1.081 (3)	C^α–N–H^3	109.05 (14)
		H^1–N–H^2	108.15 (19)
		H^1–N–H^3	110.65 (22)
		H^2–N–H^3	108.20 (18)
		C^α–C^β–$H^{\beta 1}$	110.33 (18)
		C^α–C^β–$H^{\beta 2}$	110.57 (18)
		C^α–C^β–$H^{\beta 3}$	110.37 (18)
		$H^{\beta 1}$–C^β–$H^{\beta 2}$	108.90 (26)
		$H^{\beta 1}$–C^β–$H^{\beta 3}$	108.25 (27)
		$H^{\beta 2}$–C^β–$H^{\beta 3}$	108.36 (28)

angles in the molecule, and Table III gives distances and angles in the hydrogen bonds. Figure 1 is a stereoscopic drawing of the molecule and Figure 2 is a stereoscopic drawing of the content of two unit cells.

(9) E. O. Schlemper, W. C. Hamilton and S. J. La Placa, *J. Chem. Phys.*, **54**, 3990 (1971).
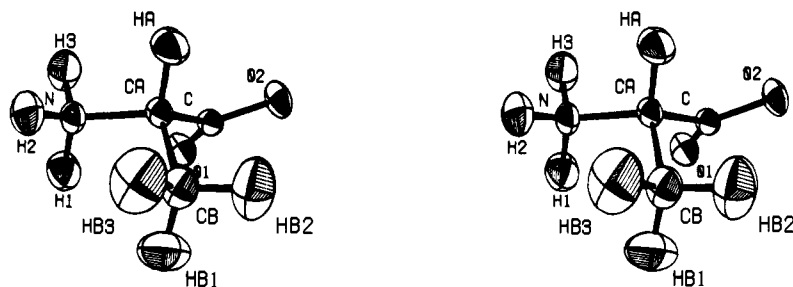
Figure 1. Stereoscopic drawing of the molecule, with 50% probability vibrational ellipsoids; $A = \alpha$, $B = \beta$.
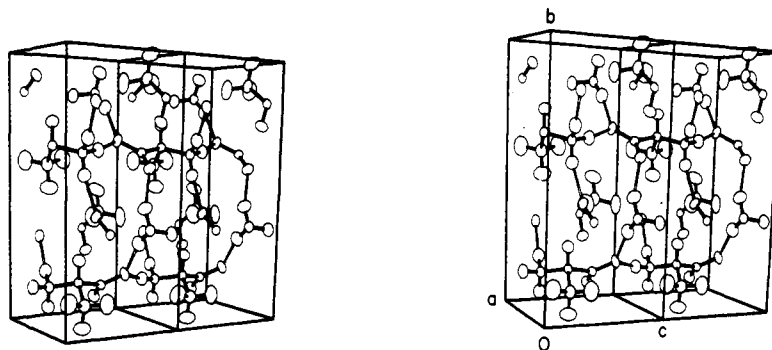

Figure 2. Stereoscopic view of two unit cells; molecular bonds solid, hydrogen bonds open.

The atomic positions found earlier have been confirmed by this study. The main new feature is the precise determination of the hydrogen atom positions. The molecule is in the zwitterion form and the three hydrogen atoms bound to N form three independent hydrogen bonds, which are nearly identical in lengths and angles. Two of the hydrogen bonds involve O^2

Table III. Distances (Å) and Angles (deg) in the Hydrogen Bonds

Bond distances		Bond angles	
$O^1 \cdots H^1$	1.861 (3)	$O^1 \cdots H^1$—N	160.9 (2)
$O^1 \cdots N$	2.853 (1)		
$O^2 \cdots H^2$	1.780 (2)	$O^2 \cdots H^2$—N	168.1 (2)
$O^2 \cdots N$	2.813 (1)		
$O^2 \cdots H^3$	1.828 (3)	$O^2 \cdots H^3$—N	163.7 (2)
$O^2 \cdots N$	2.832 (2)		

and only one involves O^1, resulting in a small but significant lengthening of the bond $C-O^2$ in comparison with $C-O^1$. One hydrogen bond links the alanine molecules together to form a chain along the c axis of the crystal; the two other bonds bind these chains together in a three-dimensional network. The channels formed by this network are occupied by the methyl groups.

The angles in the methyl group differ significantly from the tetrahedral angle. The average angle $C^\alpha-C^\beta-H^{\beta i}$ is 110.4 (1)° and the averange angle $H^{\beta i}-C^\beta-H^{\beta j}$ is 108.5 (2)°. The principal motion of the methyl group is a rotation around the $C^\alpha-C^\beta$ axis (see Figure 1), and the distance of the hydrogen atoms from this axis might be observed too small if no correction for libration were applied. In order to obtain the best geometrical parameters for the methyl group, a correction was therefore carried out using the **T**, **L**, and **S**

tensors[10] in the following way. First the nonhydrogen atoms of the molecular framework were assumed to behave as a rigid body. The analysis of the corresponding atomic vibration tensors in terms of **T**, **L**, and **S** tensors led to an rms difference between observed and calculated thermal parameters u_{ij} of 0.001 Å². The appropriate mean-square amplitudes corresponding to movements of the framework were subtracted from the vibration tensor components of the hydrogen atoms, and a calculation was carried out taking the atoms $H^{\beta 1}$, $H^{\beta 2}$, and $H^{\beta 3}$ to be a rigid body. The site symmetry was assumed to be 3, with the principal axis along the $C^\alpha-C^\beta$ axis. The rms goodness of fit for the residual u_{ij} was in this case 0.005 Å². The only components of these residual **T** and **L** tensors which differ from zero by more than two estimated standard deviations are the components along the symmetry axis, giving a mean-square amplitude of translation of 0.015 (7) Å² and a mean-square amplitude of rotation of 0.0276 (25) rad² with respect to this axis. The screw components are negligibly small. The distances corrected for rigid-body motions together with the angles involving the methyl group are given in Table IV. The average angles after correction are 110.2 (1)° for $C^\alpha-C^\beta-H^{\beta i}$ and 108.7 (2)° for $H^{\beta i}-C^\beta-H^{\beta i}$, so the deviation from tetrahedral geometry seems to be real and not unexpected. Similar results have been found in hexamethylbenzene,[11] where the C–C–H angle is 112.4 (4)°. The mean value of the $C^\beta-H^\beta$ distance of 1.096 (3) Å compares well with the standard value of 1.094 Å.[12]

An equivalent calculation was carried out for the ammonium group. The rms difference between observed and calculated u_{ij} was 0.003 Å². A similar result was obtained for the components of **T** and **L**

(10) V. Schomaker and K. N. Trueblood, *Acta Crystallogr., Sect. B*, **24**, 63 (1968).
(11) J. Edmonds and W. C. Hamilton, unpublished work.
(12) K. Kuchitsu and L. S. Bartell, *J. Chem. Phys.*, **36**, 2470 (1962).

Table IV. Distances (Å) and Angles (deg) Corrected for Rigid-Body Motion

Bond distances		Bond angles	
C^α–C	1.537	C–C^α–N	110.0
C^α–N	1.497	C–C^α–C^β	111.2
C^α–C^β	1.534	C–C^α–H^α	108.4
C^α–H^α	1.091	N–C^α–C^β	109.8
C–O^1	1.249	N–C^α–H^α	107.0
C–O^2	1.266	C^β–C^α–H^α	110.4
N–H^1	1.032	C^α–C–O^1	118.3
N–H^2	1.050	C^α–C–O^2	116.0
N–H^3	1.034	O^1–C–O^2	125.7
C^β–$H^{\beta 1}$	1.096	C^α–N–H^1	111.2
C^β–$H^{\beta 2}$	1.097	C^α–N–H^2	109.3
C^β–$H^{\beta 3}$	1.096	C^α–N–H^3	108.9
		H^1–N–H^2	108.3
		H^1–N–H^3	110.8
		H^2–N–H^3	108.3
		C^α–C^β–$H^{\beta 1}$	110.1
		C^α–C^β–$H^{\beta 2}$	110.3
		C^α–C^β–$H^{\beta 3}$	110.1
		$H^{\beta 1}$–C^β–$H^{\beta 2}$	108.5
		$H^{\beta 1}$–C^β–$H^{\beta 3}$	108.6
		$H^{\beta 2}$–C^β–$H^{\beta 3}$	109.1

with a mean-square translation along the principal axis C^α–N of 0.011 (4) Å^2 and a mean-square rotation around the axis of 0.0106 (13) rad^2. The effective translations for the screw tensor are again small. The corrected distances and angles are given in Table IV.

An estimate of the librational frequencies ν and the barrier heights V_0 can now be obtained for the two groups. A harmonic oscillator approximation to an n-fold cosine-hindered rotor with $n = 3$ was used.[8] The potential function was described as $V(\alpha) = \frac{1}{2}V_0 (1 - \cos n\alpha) \approx \frac{1}{4}V_0 n^2 \alpha^2$, where α is the angle of rota-

tion around the principal axis. The mean-square rotation is then given by $\langle \alpha^2 \rangle = h/(8\pi^2 I\nu) \coth (h\nu/2KT)$, where I is the moment of inertia of the rotor and the frequency ν is given by $\nu = (n/2\pi)(V_0/(2I))^{1/2}$. For the methyl group we found $V_0 = 5.6$ kcal/mol and $\nu = 309$ cm^{-1}, and for the ammonium group we found $V_0 = 20$ kcal/mol and $\nu = 600$ cm^{-1}. The precision of the barrier heights is about 20% for the methyl group and about 35% for the ammonium group. The effect of the hydrogen bonds in raising the ammonium group barrier is dramatic.

Torsional angles are given in Table V. IUPAC–IUB conventions[13] have been used. The difference between

Table V. Torsion Angles (deg)

$\phi^{1 \, a}$	58.3 (2)
ϕ^2	177.8 (1)
ϕ^3	−64.0 (2)
$\chi^{1 \, b}$	57.6 (2)
χ^2	177.3 (2)
χ^3	−62.2 (2)
$\psi^{1 \, c}$	−18.6 (1)
ψ^2	161.5 (1)

^{a} ϕ^i is the torsional angle C–C^α–N–H^i. ^{b} χ^i is the torsional angle N–C^α–C^β–$H^{\beta i}$. ^{c} ψ^i is the torsional angle N–C^α–C–O^i.

the angles ψ^1 and ψ^2 is 180°, showing the group C^α–C–O1–O2 to be planar.

Partial Resolution of Amino Acids by Column Chromatography on a Polystyrene Resin Containing an Optically Active Copper(II) Complex[1]

Richard V. Snyder,[2] Robert J. Angelici,*[3] and Ronald B. Meck

Contribution from the Department of Chemistry, Iowa State University, Ames, Iowa 50010. Received August 6, 1971

Abstract: A labile metal complex, (*N*-carboxymethyl-L-valine)copper(II), which had been previously shown to coordinate L-amino acids more strongly than their D enantiomers, has been chemically bound to a styrene–0.8% divinylbenzene copolymer. The resulting ligand-exchange resin has been used chromatographically to partially resolve several optically active amino acids. In all cases, the D enantiomers eluted first, and the degree of resolution increased with an increase in the bulkiness of the side chain on the α-carbon of the amino acid.

Chromatographic resolutions of optically active amino acids have been attempted on a wide variety of asymmetric sorbents.[4,5] By far the best resolutions

have been achieved by gas chromatography of volatile amino acid derivatives,[6] but this technique is not easily adapted to resolutions on a preparative scale. In principle, such preparative resolutions could be carried out by liquid chromatography, but in practice they have been only partially successful. After the work

(1) Presented at the 161st National Meeting of the American Chemical Society, Los Angeles, Calif., April 1, 1971.
(2) National Science Foundation Trainee, 1968–1969.
(3) Fellow of the Alfred P. Sloan Foundation, 1970–1972.
(4) D. R. Buss and T. Vermeulen, *Ind. Eng. Chem.*, 60 (8), 12 (1968).
(5) S. V. Rogozhin and V. A. Davankov, *Russ. Chem. Rev.*, 37, 565 (1968).

(6) See, for example, J. C. Dabrowlak and D. W. Cooke, *Anal. Chem.*, 43, 791 (1971); J. A. Corbin, J. E. Rhoad, and L. B. Rogers, *ibid.*, 43, 327 (1971), and references therein.