

intensity of  $\beta$  and  $\gamma$  bands. Sodium silicate and lauryl sulfate micelles compete for dye ions with the organic detergent having the greater tendency to interact with pinacyanol chloride.

The results indicate that salts affect the absorption spectra of the dye directly as well as by influencing aggregation.

PHILADELPHIA 6, PA.

RECEIVED SEPTEMBER 1, 1949

[CONTRIBUTION No. 1353 FROM THE GATES AND CRELLIN LABORATORIES OF CHEMISTRY, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA]

## Interatomic Distances and Bond Angles in the Polypeptide Chain of Proteins<sup>1,2</sup>

BY ROBERT B. COREY AND JERRY DONOHUE

In many problems of protein structure, it has long been necessary to assume probable distances between the atoms which make up a polypeptide chain. The magnitude assigned to these interatomic bonds and to the angles which are associated with them, have undergone progressive revisions corresponding to advancing knowledge of the structures of molecules in general. Recent X-ray determinations of the structures of amino acids, peptides, and related compounds now provide direct experimental data regarding the dimensions of the peptide linkage and other atomic groupings characteristic of protein molecules.

Some time ago, data from the structures of diketopiperazine<sup>3</sup> and glycine<sup>4</sup> were used to derive probable dimensions of the polypeptide chain,<sup>5</sup> and more recently the data obtained from an X-ray examination of DL-alanine<sup>6</sup> were incorporated in a similar discussion.<sup>7</sup> No data from a linear peptide were then available, so that bond angles were still in doubt by as much as 5°, and some of the distances required the support of additional data before they could be accepted with complete confidence. In particular, no explanation could be advanced for the short  $\alpha$ C-N distance (about 1.43 Å.) found in alanine (about 1.40 Å. in the less precisely determined glycine structure). Further X-ray studies have recently provided satisfactory answers to these questions. They include determinations of the crystal structures of L-threonine<sup>8</sup> and N-acetylglycine<sup>9</sup> and a revision of the parameters of DL-alanine,<sup>10</sup> in all of which the refinement of atomic positions was carried out by means of three-dimensional Fourier syntheses involving the simultaneous use of all

intensity data. Also a complete structure determination for crystals of  $\beta$ -glycylglycine<sup>11</sup> now supplies the first direct structural data for a linear peptide.

Practical methods for the rapid calculation of three-dimensional Fourier summations have been devised only within the last few years. The power of this technique is well illustrated by the case of DL-alanine, for which two-dimensional Fourier projections were inadequate to satisfactorily resolve any atom of the structure whereas a three-dimensional Fourier plot calculated from the data originally published completely resolved all C, N and O atoms.<sup>10</sup>

Interatomic distances and bond angles found in these recent structure determinations are compared in Table I. Figure 1 represents the probable dimensions of the fully-extended polypeptide

TABLE I  
A COMPARISON OF CERTAIN DIMENSIONS FOUND IN SOME CRYSTALS OF AMINO ACIDS AND PEPTIDES

	L-Threonine <sup>8</sup>	DL-Alanine <sup>10</sup>	N-Acetylglycine <sup>9</sup>	$\beta$ -Glycylglycine <sup>11</sup>
A. Interatomic Distances, Å.				
Carboxyl C-O	1.24	1.21	1.19	1.21
	1.25	1.27	1.31	1.27
Carboxyl C- $\alpha$ C	1.52	1.54	1.51	1.53
$\alpha$ C-N	1.49	1.50	1.45	1.48
$\alpha$ C- $\beta$ C	1.54	1.51	....	....
$\beta$ C-C	1.50	....	....	....
N-Carbonyl C'	....	....	1.32	1.29
Carbonyl C'-O'	....	....	1.24	1.23
Carbonyl C'- $\alpha$ C'	....	....	1.50	1.53
$\alpha$ C'-N'	....	....	....	1.51
B. Bond Angles				
Carboxyl O-C-O	127°	125°	124°	124.5°
Carboxyl O-C- $\alpha$ C	117	121	124	123
	116	113	112	112
Carboxyl C- $\alpha$ C-N	110	108	110	110.5
Carboxyl C- $\alpha$ C- $\beta$ C	113	111	....	....
N- $\alpha$ C- $\beta$ C	108	110	....	....
$\alpha$ C-N-Carbonyl C'	....	....	120	122
N-Carbonyl C'-O'	....	....	121	125
N-Carbonyl C'- $\alpha$ C'	....	....	118	121
$\alpha$ C'-Carbonyl C'-O'	....	....	121	121
Carbonyl C'- $\alpha$ C'-N'	....	....	....	110

(1) Aided by a grant from the National Foundation for Infantile Paralysis.

(2) Presented at the Symposium on the Chemistry of Proteins, held at the One-hundred and Sixteenth meeting of the American Chemical Society in Atlantic City, New Jersey, September 1949, under the auspices of the Division of Biological Chemistry of the American Chemical Society.

(3) R. B. Corey, *THIS JOURNAL*, **60**, 1598 (1938).

(4) G. Albrecht and R. B. Corey, *ibid.*, **61**, 1087 (1939).

(5) R. B. Corey, *Chem. Revs.*, **26**, 227 (1940).

(6) H. A. Levy and R. B. Corey, *THIS JOURNAL*, **63**, 2095 (1941).

(7) R. B. Corey, *Adv. in Protein Chem.*, **4**, 385 (1948).

(8) D. P. Shoemaker, J. Donohue, V. Schomaker and R. B. Corey, *THIS JOURNAL*, **72**, 2328 (1950).

(9) G. B. Carpenter and J. Donohue, *ibid.*, **72**, 2315 (1950).

(10) J. Donohue, *ibid.*, **72**, 949 (1950).

(11) E. W. Hughes and W. J. Moore, *ibid.*, **71**, 2618 (1949).

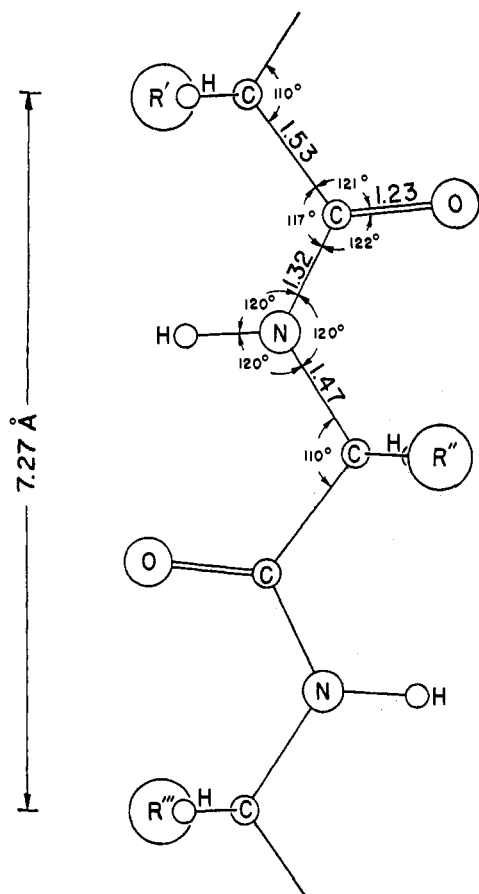


Fig. 1.—Dimensions of the fully-extended polypeptide chain as derived from recent crystal structure data.

chain of proteins as derived from the data listed in the table. With the exception of the revised  $\alpha$ C-N distance (1.47 Å.), these distances and

bond angles differ little from those previously assigned,<sup>5,7</sup> but the probability that they are essentially correct is greatly increased. The bonds around the nitrogen atom form angles which are close to  $120^\circ$ . The N-C bond which is involved in resonance is about 1.32 Å. in length. In the choice of bond angles around the carbonyl carbon atom, values intermediate between those found in N-acetylglycine and  $\beta$ -glycylglycine were taken, with slight weighting in favor of the N-acetylglycine structure which was based upon complete intensity data and on three-dimensional Fourier refinement of atomic parameters. The C-C distance 1.53 Å. is intended to reflect the fact that in amino acids and some other recently determined structures the C-C bond has been reported as slightly, but probably significantly, less than the classical 1.54 Å. Although the dimensions of the polypeptide chain in proteins probably depart little from those here indicated, more direct evidence based on precise determinations of atomic positions in crystals of higher linear peptides is greatly desired.

**Acknowledgment.**—All of the references cited in this paper refer to work which was carried out in the Gates and Crellin Laboratories. We are glad to acknowledge the support provided by the Rockefeller Foundation and by the National Foundation for Infantile Paralysis for this basic attack on the general problem of the structure of proteins.

### Summary

Data from recent crystal structure studies of amino acids and peptides are used to derive values for the bond lengths and bond angles of the polypeptide chain of proteins.

PASADENA, CALIFORNIA RECEIVED DECEMBER 5, 1949

[CONTRIBUTION FROM THE KEDZIE CHEMICAL LABORATORY, MICHIGAN STATE COLLEGE]

## The Exchange Reaction between Anhydrous Deuterium Chloride and Phenol<sup>1</sup>

BY HAROLD HART

Deuterium-hydrogen exchange has been used as a diagnostic tool in the study of the lability of hydrogen atoms in organic compounds since heavy water first became available. In the case of aromatic compounds, this exchange has been shown to follow the common orientation rules of electrophilic substitution.<sup>2</sup>

It is well known that the aromatic ring of phenols is particularly susceptible to such substitution. Thus, it is possible to alkylate phenol in certain cases without the use of any of the

common Friedel-Crafts catalysts.<sup>3,4</sup> In spite of this fact, exchange studies with phenol and anisole with heavy water in the presence of aqueous hydrochloric acid or aqueous alkali have shown the reaction to be slow, even at  $100^\circ$ .<sup>5,6,7,8</sup>

(3) H. Hart and J. H. Simons, *THIS JOURNAL*, **71**, 345 (1949).

(4) G. M. Bennett and F. M. Reynolds, *J. Chem. Soc.*, 131 (1935); J. Van Alphen, *Rec. trav. chim.*, **46**, 287 (1927).

(5) C. K. Ingold, C. G. Raisin and C. L. Wilson, *J. Chem. Soc.*, 1637 (1936).

(6) P. A. Small and J. H. Wolfenden, *ibid.*, 1811 (1936).

(7) T. Titani, *et al.*, *Bull. Chem. Soc. Japan*, **10**, 554 (1935); **11**, 554 (1936); **13**, 681 (1938); **14**, 353 (1939); see also *C. A.*, **30**, 2082 (1936); **31**, 97 (1937); **33**, 2399 (1939); **34**, 1902 (1940).

(8) A. I. Brodskii, *Izvest. Akad. Nauk S. S. R., Otdel. Khim. Nauk*, 3 (1949); *C. A.*, **43**, 5011 (1949).

(1) This paper was presented at the 116th meeting of the American Chemical Society, Atlantic City, September, 1949.

(2) A. P. Best and C. L. Wilson, *J. Chem. Soc.*, 28 (1938).