I

# EQUILIBRIUM SEDIMENTATION OF MACROMOLECULES IN DENSITY GRADIENTS WITH APPLICATION TO THE STUDY OF DEOXYRIBONUCLEIC ACID

II

# THE CRYSTAL STRUCTURE OF N,N'-DIMETHYL MALONAMIDE

Thesis by

Matthew Meselson

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1957

ACKNOWLEDGMENTS

I am grateful to Dr. Raphael Pasternak, my research advisor until he left the Institute in 1956, for his good friendship and lively criticisms.

Professor Linus Pauling, my thesis advisor, has given greatly appreciated advice and aid, especially at critical times.

Dr. Franklin Stahl has been a close companion as well as a valued research partner. Part of this work has been done in collaboration with him and with Dr. Jerome Vinograd, to whom I am grateful for his encouragement and friendship.

The National Science Foundation and the California Institute of Technology have provided me with appreciated financial support.

In many ways my parents have made it possible for me to pursue this work.

ABSTRACT

I. A method is presented for the study of the molecular weight and composition of macromolecules by observing in the ultracentrifuge their equilibrium distribution in a density gradient. The macromolecular material is centrifuged to near equilibrium in a solution of a low molecular weight solute. Under these conditions, sedimentation of the low molecular weight solute results in a continuous increase of density along the direction of the centrifugal field. The macromolecules collect in a band where their effective density is the same as that of the solution. Analysis of the shape and position of such bands yields information regarding molecular weight and composition. By this method, a deoxyribonucleic acid sample from bacteriophage is shown to be monodisperse with molecular weight $1.4 \times 10^7$. Bacteriophage DNA containing 5-bromouracil instead of thymine has been prepared and found to differ considerably in effective density from normal DNA.

II. The crystal structure of N,N'-dimethyl malonamide has been determined by two-dimensional methods. A type of weighted density projection has been devised which is especially suited to the refinement of projections complicated by atomic overlap. The crystal is monoclinic with space group C2/c; $\underline{a} = 8.66$, $\underline{b} = 4.56$, $\underline{c} = 18.51$ Å, $\beta = 94.2^\circ$; $n = 4$. A network of N-H...O hydrogen bonds ties molecules into extended sheets parallel to the ab plane. Although the configuration and dimensions found for the amide group conform quite closely to previous findings for related compounds, there is some indication of an especially large double bond character for the $N\text{-}C_1$ bond.

Table of Contents

Table of Contents (Cont'd)

# I. EQUILIBRIUM SEDIMENTATION OF MACROMOLECULES IN DENSITY GRADIENTS*

## A. Descriptive Part

A solution of a low molecular weight solute is centrifuged until equilibrium is closely approached. The opposing tendencies of sedimentation and diffusion have then produced a stable concentration gradient of the low molecular weight solute. The concentration gradient and compression of the liquid result in a continuously increasing density along the direction of centrifugal force. Let us consider the distribution of a small amount of a single macromolecular species in this density gradient. The initial concentration of the low molecular weight solute, the centrifugal field strength, and the length of the liquid column may be chosen so that the range of density at equilibrium encompasses the effective density of the macromolecular material. The centrifugal field tends to drive the macromolecules into the region where the sum of the forces acting on each molecule is zero. (The effective density of the macromolecular material is here defined as the density of the solution in this region.) This concentrating tendency is opposed by Brownian motion, with the result that at equilibrium the macromolecules are distributed with respect to concentration in a band of width inversely related to their molecular weight.

Gaussian bands. It is shown in a later section that in a constant density gradient under certain attainable conditions, the equilib-

---

* The material contained on pages 1-20 follows the treatment of M. Meselson, F. W. Stahl, and J. R. Vinograd, to be published.

rium concentration distribution of a single macromolecular species is Gaussian. The standard deviation of this Gaussian band is inversely proportional to the square root of the macromolecular weight. The Gaussian distribution is centered about the cylindrical surface corresponding to the effective density of the macromolecular material. Fig. 1 shows a photometer record of the equilibrium distribution of bacteriophage DNA in a density gradient of cesium chloride in water.

If the macromolecular material is composed of species with various molecular weights and effective densities, the observed equilibrium distribution is the sum of the separate Gaussian distributions with standard deviations and means corresponding to the molecular weights and effective densities respectively of the various molecular species. When heterogeneities in molecular weight and effective density are both present it is possible but a priori unlikely that the separate distributions would conspire to produce a single observed band of essentially Gaussian form*.

---

* This possibility is subject to experimental test. By means of a partition cell the material on either side of the mean may be isolated and rebanded. The new band will be skewed if there was density heterogeneity in the original band. Alternatively, one may compare the concentration distributions observed in cells of different shape. For a single species, the concentration distribution is independent of cell shape. However, for material heterogeneous with respect to either molecular weight or effective density or both, the observed concentration distribution is dependent on variations in the area of the cell along the radius of rotation. This dependence is such that for material of homogeneous density the symmetry of the band is not disturbed (although its shape may be). For a material with density heterogeneity, however, the band will show various departures from symmetry depending on cell shape and the particular density distribution present.

Figure 1. The equilibrium distribution of DNA from bacteriophage $T^4$. An aliquot of osmotically shocked $T^4$ r containing 3 $\mu$g of DNA was centrifuged for 80 hours at 58,000 x g in 7.7 molal cesium chloride solution. The density gradient is essentially constant over the region of the band and equal to 0.046 g $cm^{-4}$. The DNA concentration at the maximum is 20 $\mu$g/ml.

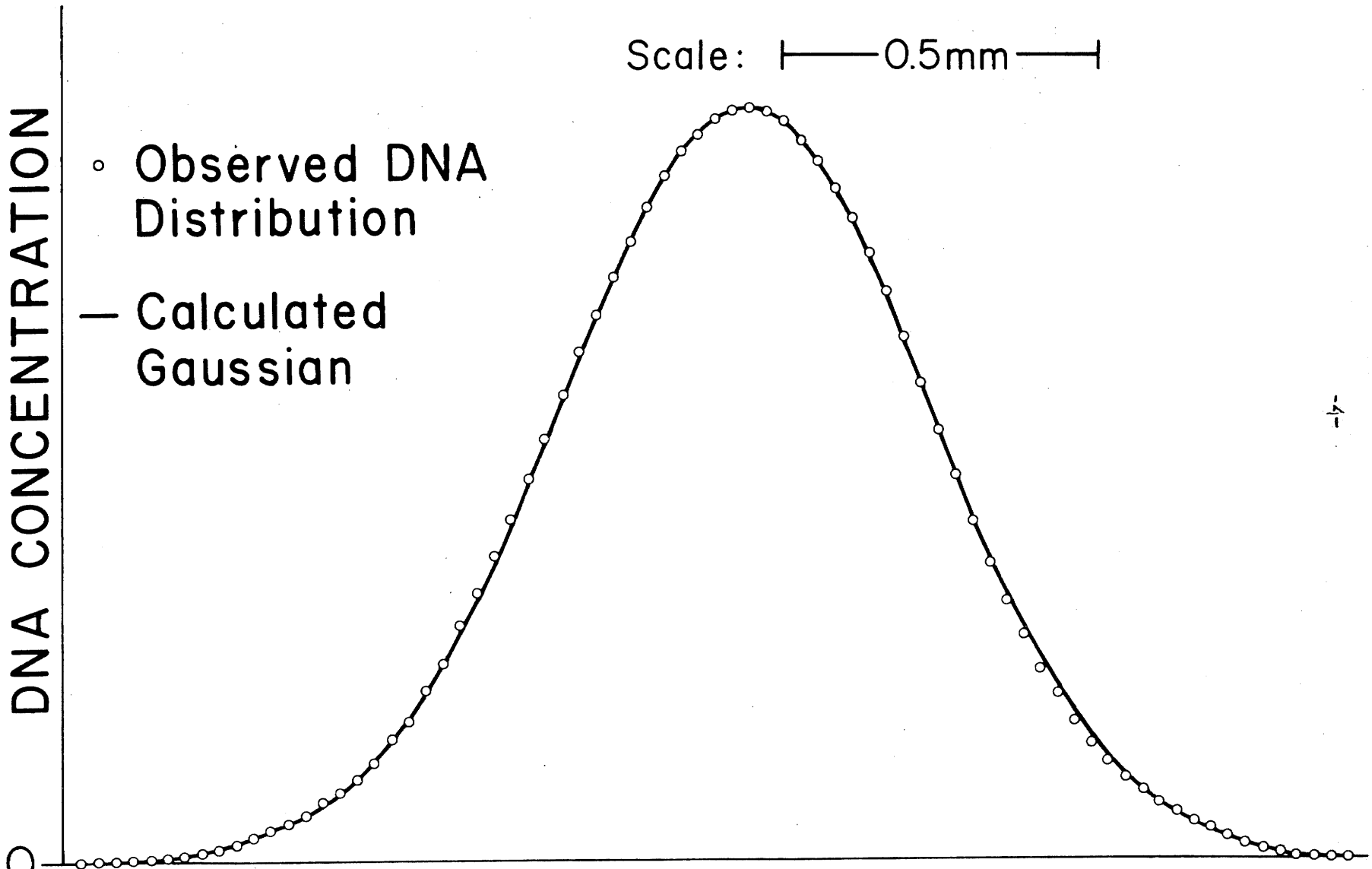

Scale: 0.5mm
Observed DNA Distribution
Calculated Gaussian
DNA CONCENTRATION
0
DISTANCE FROM ROTOR CENTER

Fig. 1

Bimodal or polymodal distribution of banded material. If the effective densities of the macromolecular species are sufficiently distinct a distribution with more than one mode will be observed. In extreme cases this may lead to the formation of discrete bands. An example is the separation of normal DNA from DNA which contains, instead of thymine, the analogue 5-bromouracil. This unusual DNA is considerably more dense than normal DNA, and prepared mixtures of the two give rise to well resolved bands in a cesium chloride gradient (fig. 2). With DNA in cesium chloride density differences of less than 0.001 $g/cm^3$ may be detected.

Skewed unimodal bands. A skewed band indicates the presence of materials heterogeneous with respect to effective density. Such bands are shown in figs. 2 and 3 for bacteriophage DNA containing 5-bromouracil and calf thymus DNA, respectively. The skewness of the former band is the result of compositional heterogeneity; i.e., some molecules contain more 5-bromouracil (in place of thymine) than others. Effective density heterogeneity may in general be either compositional or structural in origin.

Symmetric unimodal non-Gaussian bands. For the Gaussian $y = \exp\left(-\frac{x^2}{2\sigma^2}\right)$ a plot of ln y against $x^2$ yields a straight line with slope $\frac{-1}{2\sigma^2}$. A plot of the log of the concentration in a band against the square of the distance from the maximum provides a convenient test for heterogeneity. Downward concavity anywhere in this plot signifies heterogeneity in effective density. Such a case may be rare in view of the a priori unlikelihood that the effective densities would be distributed in just such a way as to give rise to a symmetrical band (see footnote,

Figure 2. The concentration distribution of a mixture of normal and 5-bromouracil-containing DNA from bacteriophage T4. A mixture of osmotically shocked normal and 5-bromouracil-containing T4 was centrifuged for 18 hours at 150,000 x g in 8.9 molal cesium chloride at pH 8.4. The density gradient is 0.12 g $cm^{-4}$. The position of the normal DNA indicates an effective density of 1.70; the maximum effective density of the substituted DNA is 1.80.

Figure 3. The concentration distribution of calf thymus DNA banded in cesium chloride. 2.7 $\mu$g of calf thymus DNA was centrifuged for 15 hours at 150,000 x g in 7.7 molal cesium chloride at pH 8.4. The density gradient is 0.12 g $cm^{-4}$.

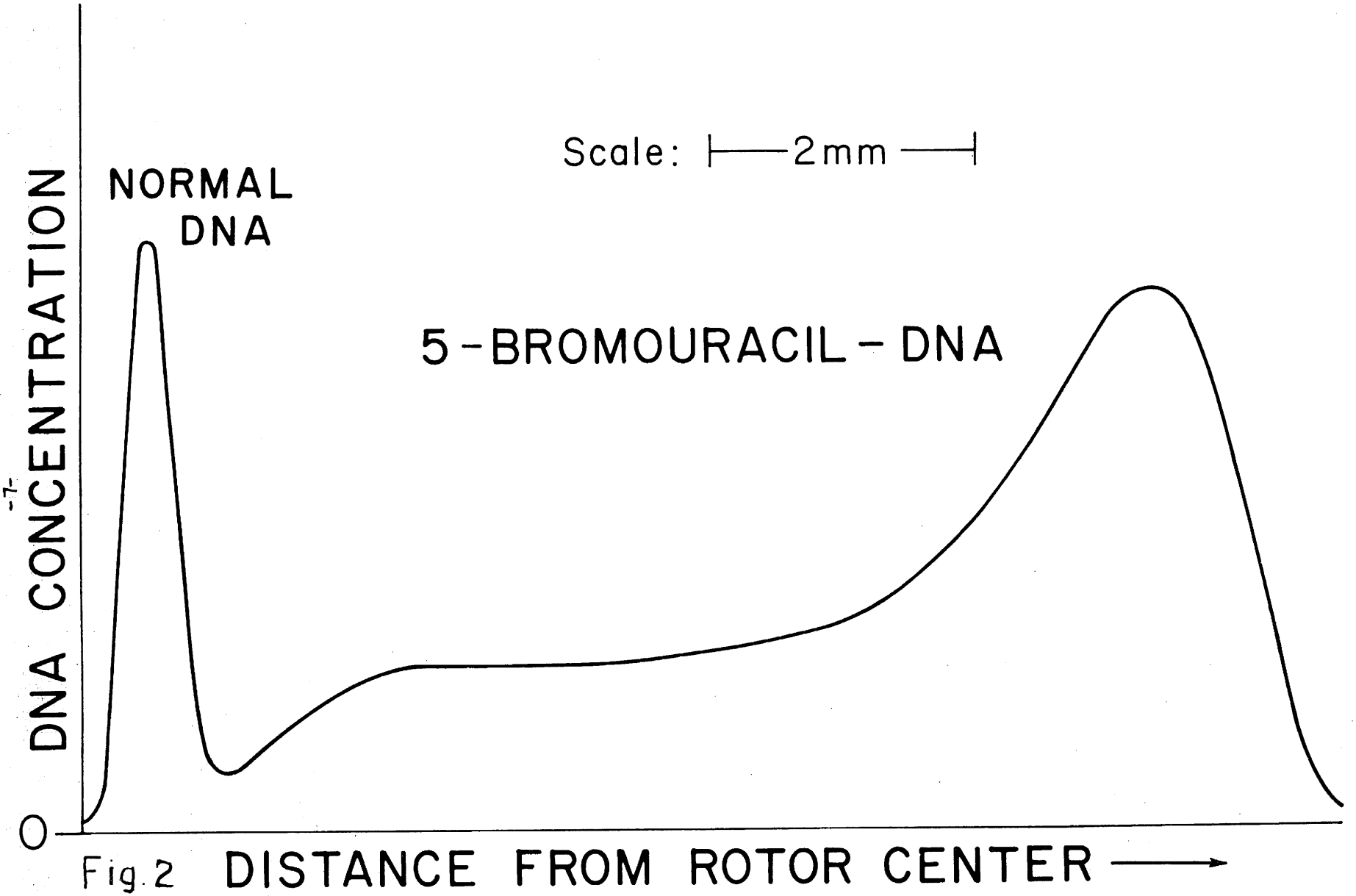

Scale: 2mm
NORMAL DNA
5-BROMOURACIL-DNA
DNA CONCENTRATION
O
DISTANCE FROM ROTOR CENTER

Fig. 2

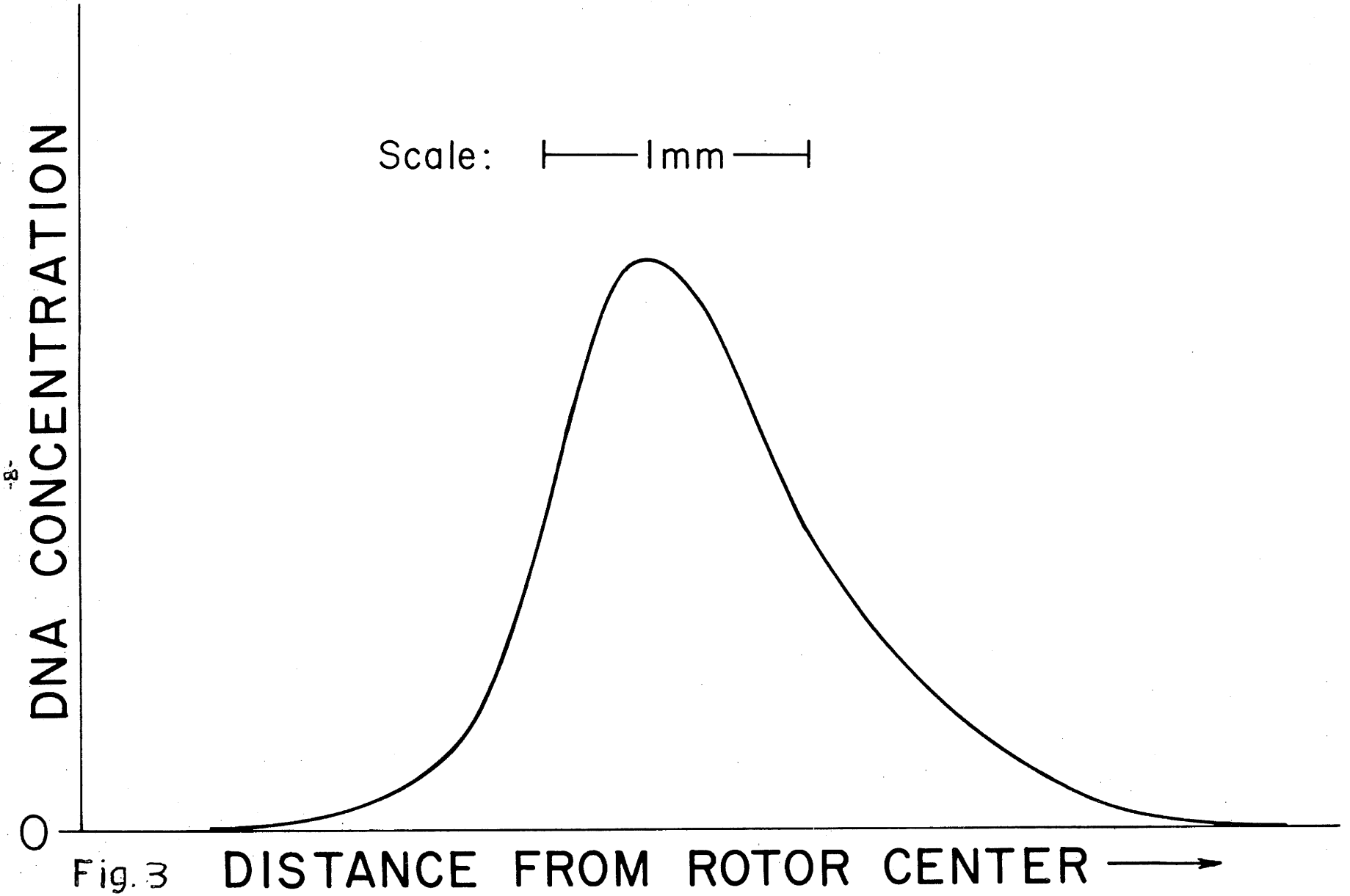



Fig. 3

page 2). The absence of downward concavity coupled with the observed symmetry of the concentration distribution is strong presumptive evidence for density homogeneity. Under this presumption upward concavity is evidence for heterogeneity in molecular weight. The slope at any point is inversely proportional to the weight mean molecular weight of the material at the corresponding position in the band (see the later discussion).

Measurements of effective density. The mean effective density of macromolecular material distributed in any manner within the density range of the solution may be found from the mean of the mass distribution evaluated from the observed concentration distribution. If effective density is influenced by composition, the distribution provides a basis for the analysis of the composition of the material. This application of the method is illustrated by the results with phage DNA containing 5-bromouracil (fig. 2). The effective density of this DNA is found to be related to the degree of substitution of thymine by 5-bromouracil as determined chromatographically (see page 38 ).

Molecular weight determinations. In the absence of density heterogeneity both the number and weight mean molecular weight may be calculated from the observed shape of the equilibrium band, as will be shown later. Molecular weights calculated from concentration distributions should be considered minimal values if there is a possibility of density heterogeneity.

Details of the experimental procedures and results with several materials are described in a later section.

B. Quantitative Relations

The equilibrium distribution. The total potential of any component at equilibrium in a closed system at constant temperature must be uniform throughout the system. In a centrifugal field this requirement results in the rigorous condition

$$1) \quad M_i(1-\bar{v}_i\rho)r\omega^2 dr - \sum_K \frac{\partial \mu_i}{\partial c_K} dc_K = 0$$

where $M_i$, $\bar{v}_i$, $\mu_i$, $c_i$ are molecular weight, partial specific volume, partial molal (Gibbs) free energy, and concentration of the ith component. The summation extends over a complete set of independently variable components. The angular velocity is given by $\omega$, the radial coordinate is r and the density of the solution at r is $\rho$. Consider a system containing water, a low molecular weight electrolyte XY and a macromolecular electrolyte $PX_n$. The discussion will apply equally well to positive or negative polymeric ions and, with $n = 0$, to neutral polymer molecules. In the case of neutral polymers, but not otherwise, the low molecular weight solute need not be an electrolyte.

Three components are necessary and sufficient* to describe the

---

* Although three components are sufficient for a thermodynamic description of the system in the absence of the centrifugal field, more may be required in its presence. A system subject to a centrifugal field may be regarded as being composed of a continuous sequence of phases of infinitismal depth in the direction of the field. A number of components sufficient to describe the chemical composition of each phase must be employed. In the present case, more than three components would be required if any product arises with composition not describable in terms of the three chosen components and which sediments differently than any of the components involved in its formation. This complication is explicitly excluded from the present discussion.

composition of the system. They are chosen here as water, XY, and the neutral unsolvated molecule $PX_n$. Other choices are of course permissible but this one is especially convenient.

For definiteness in making certain approximations and for comparison with experiment we shall often refer to the system water, cesium chloride, cesium deoxyribonucleate. The total amount of polymer will be made so small as to have a negligible effect on the potentials of the salt and water. Therefore we may first calculate the concentration distribution of XY from equation 1 ignoring the polymer, and then, again using equation 1, find the distribution of the polymer in the salt gradient. This gradient is

$$2)\quad \frac{dC_{xy}}{dr} = \frac{\partial C_{xy}}{\partial a_{xy}} \cdot \frac{da_{xy}}{dr} = \frac{\partial C_{xy}}{\partial a_{xy}} \cdot \frac{a_{xy} M_{xy} (1-\bar{v}_{xy}\rho) r\omega^2}{RT}$$

where $C_{XY}$ is the concentration of XY. The sum of equation 1 has been replaced by its equivalent, RT d ln $a_{XY}$ where $a_{XY}$ is the activity. Values of $a_{XY}$, $\rho$, and $\bar{v}_{XY}$ as functions of concentration and pressure are experimentally determinable and may be found in the literature for some systems. The salt concentration as a function of r may be found by numerical integration of equation 2. Alternatively, it may be measured by optical methods in the centrifuge itself (see pages 26-30).

In many cases it is possible to select XY so that the density and concentration gradients are essentially constant over regions suf-

ficiently long to encompass a polymer band. For example, it is found both by calculation from equation 2 and by direct observation of the refractive index gradient in the centrifuge, that the cesium chloride concentration and density gradients are essentially constant over the region of a DNA band. For computational simplicity we shall consider only these linear systems so that upon choosing $r_o$ in the region of a band, we may write $C_{XY}$ and $\rho$ in the band as

3) $$C_{xy}(r) = C_{xy}(r_o) + \left(\frac{dC_{xy}}{dr}\right)_{r_o} (r - r_o)$$

4) $$\rho(r) = \rho(r_o) + \left(\frac{d\rho}{dr}\right)_{r_o} (r - r_o)$$

Having found the distribution of XY, we may now employ equation 1 to determine the distribution of $PX_n$. Making use of equation 4, we write the first term of equation 1, which represents the work per mole done against the centrifugal field moving $PX_n$ from r to $r + dr$, as

5) $$M_{PX_n}\left[1 - \bar{V}_{PX_n}\rho(r_o) - \bar{V}_{PX_n}\left(\frac{d\rho}{dr}\right)_{r_o}(r - r_o)\right] r\omega^2 dr$$

It should be emphasized that $\bar{v}_{PX_n}$ is the partial specific volume of $PX_n$ in a solution of XY at a concentration $C_{XY}(r)$. In order to evaluate the remaining term in equation 1 we consider it to be composed entirely of the osmotic work RT d ln $C_{PX_n}$ + zRT d ln $\gamma_X C_X$ where z is the effective number of counter-ions which must be moved along with the charged polymer molecule $PX_{n-z}$ in order to maintain electrical neutrality and $\gamma_X$ is the activity coefficient of the ion X. We thereby neglect any other change in the free energy of the polymeric component as it is moved through the solution. This should be a valid approximation if the fractional change in the concentration of XY across a polymer band is small. It is especially plausible for the case of DNA in cesium chloride, for which the cesium chloride and water concentrations change by less than one part in 100 over the region of a band. Also, over a small concentration range, the term d ln $\gamma_X$ will be negligible in comparison to d ln $C_X$. Incorporating these approximations in the limit of low polymer concentration, we have

$$M_{PX_n}\left(1-\bar{V}_{PX_n}\rho(r_o)-\bar{V}_{PX_n}\left(\frac{d\rho}{dr}\right)_{r_o}(r-r_o)\right) r\omega^2 dr$$

$$6)\quad -RT\,d\ln C_{PX_n}^{(r)} - zRT\,d\ln\left(C_{xy}(r_o)+\left(\frac{dC_{xy}}{dr}\right)_{r_o}(r-r_o)\right)$$

$$= 0$$

Assuming $\bar{v}_{PX_n}$ and z to be independent of r over the region of a band, this may be integrated with respect to $(r-r_o)$ from $r = r_o$ to $r = r$, yielding

$$7)\quad M_{PX_n}\left[1-\bar{V}_{PX_n}\rho(r_o)\right]\omega^2(r-r_o)\left(\frac{r-r_o}{2}+r_o\right) - M_{PX_n}\bar{V}_{PX_n}\left(\frac{d\rho}{dr}\right)_{r_o}\omega^2(r-r_o)^2\left(\frac{r-r_o}{3}+\frac{r_o}{2}\right) - 2RT\ln\left[1+\left(\frac{dC_{XY}}{dr}\right)_{r_o}\frac{(r-r_o)}{C_{XY}(r_o)}\right] - RT\ln\frac{C_{PX_n}(r)}{C_{PX_n}(r_o)} = 0$$

In many cases, including that of DNA in cesium chloride solution, the band width may be made quite small compared with the distance of the band from the center of rotation so that $|r-r_o| \ll r_o$. Further, because of the small magnitude of the gradient of $C_{XY}(r)$,

$\ln\left[1+\frac{dC_{XY}}{dr}\frac{(r-r_o)}{C_{XY}(r_o)}\right]$ may be expanded as $\frac{dC_{XY}}{dr}\frac{(r-r_o)}{C_{XY}(r_o)}$. Introducing these approximations in equation 7 and completing the square in the variable $(r-r_o)$, we obtain

$$8)\quad C_{PX_n}(r) = C_{PX_n}(r_o)\exp\left[\frac{\alpha^2}{2\sigma^2}\right]\exp\frac{-1}{2\sigma^2}\left[(r-r_o)+\alpha\right]^2$$

where

$$9) \quad \sigma^2 = \frac{RT}{M_{PX_n} \bar{V}_{PX_n} \left(\frac{d\rho}{dr}\right)_{r_o} r_o \omega^2}$$

and

$$10) \quad \alpha = \frac{z RT \left(\frac{dC_{xy}}{dr}\right)_{r_o}}{M_{PX_n} \bar{V}_{PX_n} \omega^2 r_o C_{xy(r_o)}} - \frac{\left(1 - \bar{V}_{PX_n} \rho_{(r_o)}\right)}{\bar{V}_{PX_n} \left(\frac{d\rho}{dr}\right)_{r_o}}$$

This is a Gaussian distribution with standard deviation $\sigma$ . Equation 8 is simplified by choosing $r_o$ as the mean in which case $\alpha = 0$. Therefore the density of the medium at the band center is given by the expression

$$11) \quad \rho(r_o) = \frac{1}{\bar{V}_{PX_n}} \left(1 - \frac{z RT \left(\frac{dC_{xy}}{dr}\right)_{r_o}}{M_{PX_n} \omega^2 r_o C_{xy(r_o)}}\right)$$

i.e., the pull of the counter-ions displaces the origin of the Gaussian band from the region of density $\frac{1}{\bar{V}_{PX_n}}$ to $r_o$, where the density is $\rho(r_o)$.

The final result, then, for the distribution of a single polymeric species at equilibrium in a constant density gradient is

12) $$C_{PX_n}(r) = C_{PX_n}(r_o) \exp -\left[\frac{(r-r_o)^2}{2\sigma^2}\right]$$

From the observed value of $\sigma$ the molecular weight is calculated as

13) $$M_{PX_n} = \frac{RT}{\bar{v}_{PX_n}\left(\frac{d\rho}{dr}\right)_{r_o} r_o \omega^2 \sigma^2_{observed}}$$

Provided that one uses the true value of $\bar{v}_{PX_n}$ rather than the reciprocal of the effective density $\rho(r_o)$ the molecular weight obtained from equation 13 refers to the dry neutral molecule $PX_n$ whether or not the species actually present is solvated or charged. However it is so much more convenient to measure the effective density $\rho(r_o)$ than to determine $\bar{v}_{PX_n}$ by the usual pycnometric method that one might profitably ask under what conditions it is permissible to equate the two quantities. This may be done when

14) $$1 - \frac{2RT\left(\frac{dC_{xy}}{dr}\right)_{r_o}}{M_{PX_n} r_o \omega^2 C_{xy}(r_o)} \cong 1$$

For DNA in cesium chloride solution, even if each primary phosphate carries an effective negative charge (which is surely not the case), the approximation of equation 14 involves an error of less than 10 per cent.

Having shown that a single polymeric species in a constant density gradient will be distributed at equilibrium in a band of Gaussian shape, we now turn to the more general situation of a polymer heterogeneous with respect to molecular weight although homogeneous with respect to effective density. In the limit of low polymer concentration, interactions between polymer molecules may be presumed absent so that the observed band will be the sum of many Gaussians with coincident origins with each Gaussian possessing a standard deviation related to the molecular weight of the corresponding species by equation 9.

The weight average molecular weight of the material at r is defined by the expression

$$15) \quad M_w(r) = \frac{\sum_i C_i(r) M_i}{C(r)}$$

where $C_i(r)$ and $m_i$ are the weight concentration at r and the molecular weight respectively of the ith species. According to equation 12 we have

$$16) \quad \frac{dC_i}{dr} = \frac{-(r-r_0) C_i(r)}{\sigma_i^2} = \frac{-(r-r_0) C_i(r) m_i}{\beta}$$

where $$\beta = m_i \sigma_i^2 = \frac{RT}{\bar{v}_i \left(\frac{d\rho}{dr}\right)_{r_o} r_o \omega^2}$$

It follows that

17) $$M_w(r) = \frac{-\beta \left(\frac{dC}{dr}\right)_r}{(r-r_o)\, C(r)}$$

The number average molecular weight of the material at r is defined by the expression

18) $$M_N(r) = \frac{C(r)}{\sum_i \frac{C_i(r)}{m_i}}$$

According to equation 16 we write

19) $$d\left(\frac{C_i(r)}{m_i}\right) = -\frac{(r-r_o)\, C_i(r)}{\beta}\, dr$$

It follows that

$$20) \quad \frac{C_i(r)}{m_i} = \frac{-1}{\beta} \int_{-\infty}^{r} (r'-r_o)\, C_i(r')\, dr'$$

And the number mean molecular weight of the material at r is given by

$$21) \quad M_N(r) = \frac{-\beta\, C(r)}{\int_{-\infty}^{r} (r'-r_o)\, C(r')\, dr'}$$

The corresponding molecular weight means for the material comprising a band are found by integration of equations 17 and 21 respectively and are given by

$$22) \quad M_W = \frac{-\beta \int \frac{1}{r-r_o} \cdot \frac{dC}{dr} \cdot dr}{\int C(r)\, dr}$$

and

$$23) \quad M_N = \frac{-\beta \int C(r)\, dr}{\int (r-r_o)^2\, C(r)\, dr}$$

The integrations are to extend over the entire band and are written so as to apply to cells having straight walls, whether radial or not. In the completely general case of heterogeneity of both density and molecular weight the molecular weights calculated from the above equations are minimum values. For convenience in numerically estimating the weight mean molecular weight, equations 17 and 22 may be written in the forms

$$24) \quad M_w(r) = -2\beta \frac{d \ln C}{d(r-r_0)^2}$$

$$25) \quad M_w = \frac{-2\beta \int \frac{d \ln C}{d(r-r_0)^2} \cdot C(r) \cdot dr}{\int C(r)\, dr}$$

respectively. See page 5.

The approach to equilibrium. Finally, we turn to the problem of estimating the time necessary for the distribution of the polymer to approach closely to its equilibrium value*. For this purpose, it is suf-

---

* This treatment follows that of G. M. Nazarian and M. Meselson, to be published.

ficient to replace the actual centrifuge system by one of infinite extent bounded by walls parallel to the x axis and with a uniform gravitational field acting in the positive x direction:

The concentration distribution of the polymer C(x,t) must satisfy the continuity equation

$$26) \qquad \frac{\partial}{\partial x}\left(-D\frac{\partial C}{\partial x}+CV\right)=-\frac{\partial C}{\partial t}$$

where D is the diffusion constant and V is the sedimentation velocity acquired under the action of the external force. This velocity is influenced by the buoyancy effect of the surrounding medium and is expressed by Svedberg's equation

$$27) \qquad V=\frac{M(1-\bar{v}\rho)D_f}{RT}$$

where M is the molecular weight of the polymer, $\bar{v}$ its partial specific volume, $\rho$ the density of the medium, and f the constant external force acting on unit mass.

For the medium with constant density gradient, we take

28) $$\rho = \frac{1}{\bar{v}} + \frac{d\rho}{dx} x$$

so that the origin ($x = 0$) is at the position where the density of the medium and the effective density of the polymer are identical. Then the sedimentation velocity is a function of x and is given by

29) $$V = - \frac{D x}{\sigma^2}$$

where $$\sigma^2 = \frac{RT}{M \bar{v} \frac{d\rho}{dx} f}$$

30)

Introducing this expression for V into equation 26 yields the time-dependent differential equation for the problem:

31) $$\frac{\partial}{\partial x} \left( \frac{\partial c}{\partial x} + \frac{x c}{\sigma^2} \right) = \frac{1}{D} \frac{\partial c}{\partial t}$$

The equilibrium distribution is easily obtained as a special case by setting $\frac{\partial C}{\partial t} = 0$ and imposing the condition that there be no flux at infinity. Thus

$$\frac{\partial C}{\partial x} + \frac{xC}{\sigma^2} = 0$$

or

32) $$C(x) = C(0) \exp - \left[ \frac{x^2}{2\sigma^2} \right]$$

which is identical with the Gaussian distribution obtained from thermodynamic considerations. After separation of variables and appropriate substitution, equation 31 may be cast into the form of Hermite's differential equation. Its solutions are restricted to the Hermite polynomials by the boundary condition

33) $$\int_{-\infty}^{\infty} C\,dx$$ bounded

The general solution of equation 6 may be written

$$34) \quad C(x,t) = \exp-\left[\frac{x^2}{2\sigma^2}\right] \sum_{n=0}^{\infty} A_n H_n\left(\frac{x}{\sigma\sqrt{2}}\right) \exp-\left[\frac{nDt}{\sigma^2}\right]$$

where $H_n$ is the nth Hermite polynomial. The coefficients $A_n$ are determined from the initial distribution $C(x, t = 0)$ with the aid of the orthogonality relations for Hermite polynomials. We shall consider the case in which $C(x, t = 0)$ is some constant $C_o > 0$ over the region $\frac{-L}{2} < x < \frac{L}{2}$ and zero elsewhere. Then equation 34 becomes

$$C(x,t) = \frac{L C_o}{\sigma\sqrt{2\pi}} \exp-\left[\frac{x^2}{2\sigma^2}\right] \left\{ 1 + \left(\frac{L^2}{12\sigma^2} - 1\right)\left(\frac{x^2}{2\sigma^2} - \frac{1}{2}\right) \exp-\left[\frac{2Dt}{\sigma^2}\right] \right.$$

$$35) \quad \left. + \left(\frac{L^4}{120\sigma^4} - \frac{L^2}{12\sigma^2} + \frac{1}{8}\right)\left(\frac{x^4}{\sigma^4} - \frac{6x^2}{\sigma^2} + 3\right) \exp-\left[\frac{4Dt}{\sigma^2}\right] + \ldots \right\}$$

This solution applies strictly only to a system of infinite extent along x. However, if $L \gg \sigma$ in the infinite system, the flux at $\pm \frac{L}{2}$ is very nearly zero, which is mathematically equivalent to the condition imposed by walls at $|x| = \frac{L}{2}$.

For $t^*$, the time required to reach a distribution within one per cent of equilibrium we demand that

$$\frac{C(2\sigma, t^*) - C(2\sigma, \infty)}{C(2\sigma, \infty)} = 0.01$$

hence we take

$$36) \quad \left(\frac{L^2}{12\sigma^2} - 1\right)\left(2 - \frac{1}{2}\right) \exp - \left[\frac{2Dt^*}{\sigma^2}\right] = 0.01$$

keeping only the most slowly decaying term.

Solving equation 36 for $t^*$ yields

$$37) \quad t^* = \frac{\sigma^2}{D}\left(\ln \frac{L}{\sigma} + 1.26\right)$$

On introducing this value of time into the more rapidly decaying terms of equation 35 one indeed finds that they have become negligible at this late stage in the approach to equilibrium. This estimate is based on the assumption that the density gradient is fully established at zero time. The time actually required for the equilibrium of XY may be estimated theoretically (1). The sum of these two times may be taken as a maximum

estimate of the time required in practice. In the case of DNA in cesium chloride, the cesium chloride equilibrium is in fact much more quickly approached than that for the polymer and the estimate of equation 37 need not be modified.

C. Studies of DNA in CsCl Density Gradients

The CsCl gradient. Cesium chloride of high spectroscopic purity was treated overnight with "Norite A" decolorizing charcoal and subsequently recrystallized. The charcoal treatment accomplished the removal of a partially sedimentable impurity with an adsorption maximum at 2700 A which, if present, complicates the ultraviolet photometry of DNA bands in CsCl gradients.

Purified CsCl solution buffered at pH 8.4 with 0.1 per cent tris-hydroxymethyl aminomethane was used throughout this work. The density was determined with a Westphal balance to 0.002 $g/cm^3$.

The density gradient for a given set of conditions was calculated according to the equation

$$38) \quad \frac{d\rho}{dr} = \frac{\partial \rho}{\partial a} \frac{a M (1-\bar{v}\rho)\omega^2 r}{RT}$$

Values of the density $\rho$ and the activity a of CsCl solutions as a function of concentration are tabulated in the literature. Values of the density determined by different workers (2,3) agree very well. The activities

have been determined isopiestically (4). No estimate of accuracy is given for the activity coefficients although they are reported to three places. From these published data the values of $\bar{v}$, a and $\frac{d\rho}{da}$ for CsCl of density 1.70 g/cm$^3$ at 25°C are found to be

$$\bar{v} = 0.276 \text{ cm}^3/\text{g}; \quad a = 13.9 \text{ (molal scale)} \quad \frac{d\rho}{da} = 0.0159 \text{ g/cm}^3$$

These values refer to atmospheric pressure. However, the pressure coefficients of the activity and of the density are low enough so that the effect of hydrostatic pressure in the centrifuge may be neglected. The value of the density gradient at the position r in the cell where the density is 1.70 (the effective density of DNA) is obtained from equation 38 as

$$39) \quad \left(\frac{d\rho}{dr}\right)_{\rho=1.70} = 3.5 \times 10^{-10} \omega^2 r$$

where $\omega$ is in radians per second and r is in centimeters. At a representative angular velocity, 27,690 rpm, $\left(\frac{d\rho}{dr}\right)_{\rho=1.70}$ is 0.047 g/cm$^4$, 7 cm from the center of the rotor. Because of the small change of $\frac{\partial\rho}{\partial a}$, a, $\bar{v}$, and $\rho$ with r, variations in $\frac{d\rho}{dr}$ along the cell are primarily due to the variation of centrifugal potential with distance from the center of rotation. Thus equation 39 may be used to estimate the variation of the density gradient across a DNA band. In experiments intended to measure

molecular weight, the field strength is chosen so as to give a DNA band with a standard deviation less than 0.1 cm. According to equation 39 the variation of density gradient across this distance is about one per cent for $r = 7$ cm and may be neglected. Over the small density ranges encountered, the density gradient is directly proportional to the refractive index gradient. In the Spinco model E ultracentrifuge employed throughout this work, the refractive index gradient may be observed directly by means of a modified schlieren optical system (fig. 4). Measurements of the refractive index gradient made with the schlieren system provide a direct means of estimating the density gradient. For this purpose we write

$$\frac{d\rho}{dr} = \frac{\partial\rho}{\partial n} \cdot \frac{dn}{dr}$$

For $\rho = 1.70$, $\frac{d\rho}{dn}$ is 11.7 g cm$^{-3}$. The observed value of $\frac{dn}{dr}$ at that density in a centrifugal field of 58,000 g is 0.0040 cm$^{-1}$. This value has been corrected for a refractive index gradient primarily due to deformation of the cell in the centrifugal field. This correction is obtained from the center of the schlieren diagram immediately upon reaching the full field strength before any significant sedimentation of the CsCl has

Figure 4 is a set of superimposed schlieren diagrams showing successive stages in the development of a cesium chloride equilibrium gradient. The numbers above the curves are the times in hours during which the centrifuge operated at 27,690 rpm. The slight upward slope of the refractive index gradient curve indicates a one per cent per mm departure from constancy as expected from equation 39. The diagrams of fig. 4 are chosen from an experiment in which 3 $\mu$g of DNA was present in the cell; the banded DNA produces a diphasic pip in the final schlieren diagram.

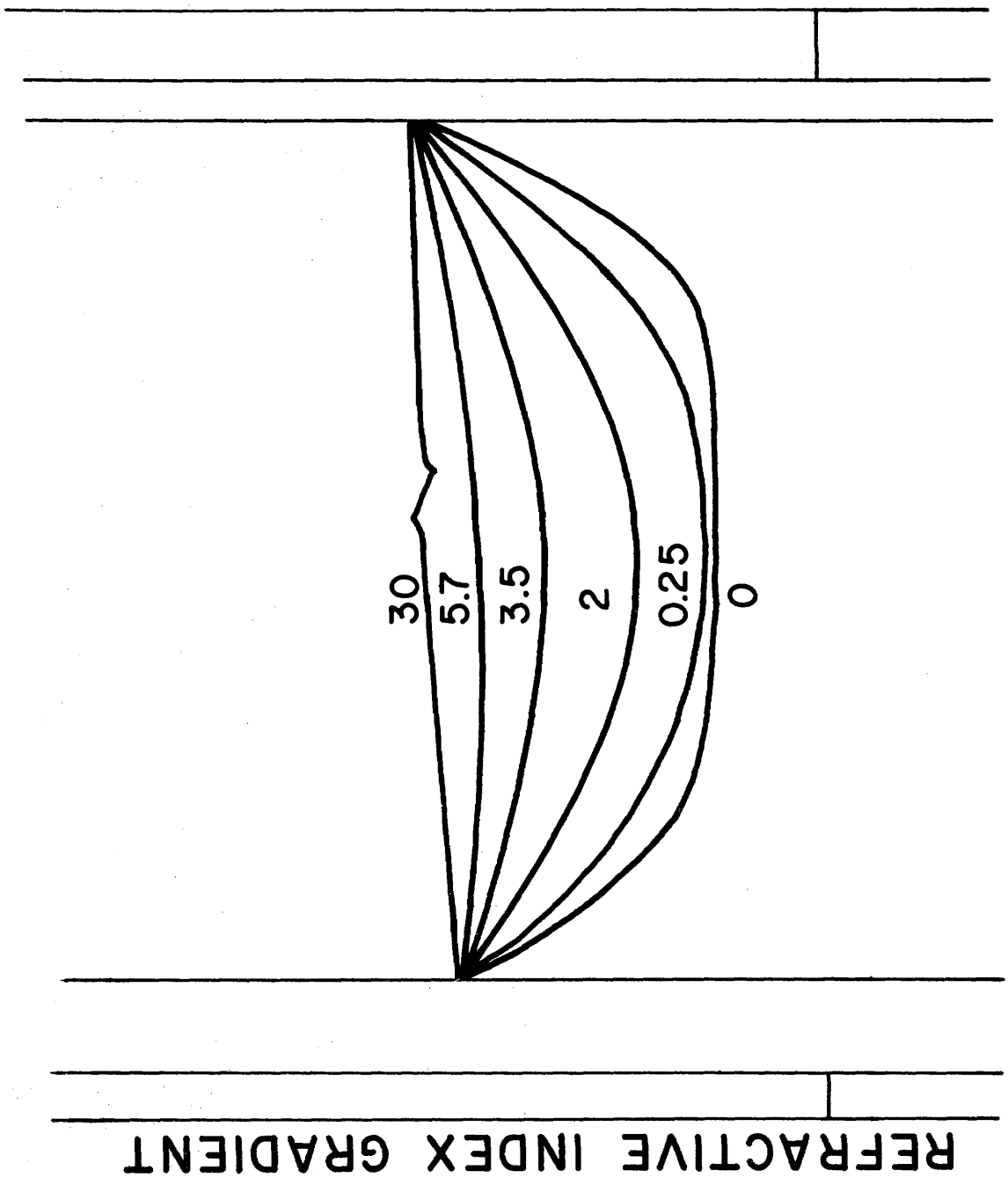

REFRACTIVE INDEX GRADIENT
30
5.7
3.5
2
0.25
0
DISTANCE FROM ROTOR CENTER

Fig. 4

taken place. The density gradient determined in this way is 0.047 g/cm $^4$ which is to be compared with the value 0.0465 calculated from equation 39.

Photometry of DNA. The concentration distribution of DNA in the ultracentrifuge cell was determined by ultraviolet photometry during the operation of the centrifuge. Illumination was provided by a low pressure mercury lamp (General Electric AH-4) equipped with a chlorine-bromine filter (5). An image of the centrifuge cell, linearly enlarged 1.8 fold, was focused upon a blue sensitive photographic plate (Eastman "Commercial"). With this combination of source, filter, and plate, only the mercury radiation at 2480, 2540, and 2650 A contributes significantly to the recorded image.

It was experimentally verified that DNA in 7.7 molal CsCl obeys Beer's Law at each of these three wave lengths and over the entire concentration range encountered in this work.

Plates were photometered with a modified Sinclair Smith microphotometer which records transmission. The linearity of the photometer was established by showing the ratio of transmission of two chromatically neutral filters to be constant over a wide range of incident illumination. Within the linear exposure range of the photographic emulsion, optical density in the centrifuge cell should be directly proportional to optical density produced on the film. That this was the case over a wide range of exposure levels was established by showing that a given DNA concentration distribution gives rise to photometrically measured optical density distributions differing only by an additive constant.

The photometer produced a 14 x linearly enlarged image of the plate on a 0.25 mm slit. Taking into account the factor by which the image of the cell appears enlarged on the film, this slit width corresponds to 0.01 mm in the cell itself.

For the work with bacteriophage DNA, the indicating galvanometer of the photometer was read directly at intervals of 0.045 mm on the film corresponding to every 0.025 mm in the cell. This spacing yielded about 80 readings across a band. Readings were also taken at doubled intervals out to one band width on either side of the band edges in order to establish an accurate base line. All readings were taken twice, scanning first in one direction and then in the other. These duplicate readings were usually identical. Averaged values were used when the forward and backward readings differed.

For the work on salmon sperm DNA, direct reading of the photometer galvanometer was abandoned in favor of continuous recordings by a Brown recording potentiometer. A comparison of directly read transmissions with recorded ones for a given film showed them to be identical. Effects on the plates due to halation and scattered light were judged negligible. The optical resolution of the system was investigated with the use of an ultraviolet photograph of a fine wire in the resting centrifuge cell. When this photograph was projected onto the slit stage of the photometer, the image of the wire was seen to be about one third the width of the slit and appeared quite sharply defined. Thus the system comprising ultraviolet optics, film, and photometer optics was shown able to resolve detail within the slit width employed. The curvature

of the image of a cylindrical surface of constant DNA concentration in the centrifuge cell was too small to reduce appreciably the attainable resolution of the concentration distribution.

For the purely qualitative work with human, calf thymus, and 5-bromouracil substituted bacteriophage DNA, a Spinco Analytrol recording densitometer was employed. This instrument produces a continuous recording of optical density on the film. Although its resolving power is much less than that of the Sinclair Smith instrument, comparisons showed it to be quite adequate for qualitative work.

Test for equilibrium. Evidence for the attainment of equilibrium in the case of bacteriophage DNA was provided by the essential identity of the final concentration distributions whether the DNA was initially distributed uniformly in the cell or in an extremely tight band produced by running at high speed. Although no quantitative study has been made of the approach to equilibrium, between 50 and 100 hours are required in practice before an initially even distribution of DNA approximates its equilibrium distribution as judged by the above experimental criterion. Times of this magnitude are predicted by equation 37 if the diffusion coefficient of DNA is taken as $10^{-8}$ $sec^{-1}$. Although no adequate experimental determination of D has yet appeared, this value is in accord with a reported maximum value of $2 \times 10^{-8}$ $sec^{-1}$ (6).

Ideality of DNA behavior. The maximum concentration of DNA in the bands for which figures are presented is 20 $\mu$g/ml. Light scattering studies of calf thymus DNA in 0.2 molar NaCl show no concentration

dependent departure from ideality below 200 μg/ml. (Higher concentrations were not investigated (7).) With the small DNA concentrations in strong CsCl solutions used in the present work it was expected that errors due to DNA concentration effects would be altogether negligible. As an experimental test of this expectation, equilibrium concentration distributions of bacteriophage DNA at two different concentration levels were compared. The maximum concentration was 20 μg/ml in one experiment and 10 μg/ml in the other. The two distributions were found to be the same and it was concluded that concentration dependent departures from ideal behavior were essentially absent.

Preparation of DNA samples.

(a) Normal bacteriophage DNA. Bacteriophage $T^4$ r240 was propagated on Escherichia Coli strain B in a glucose salts medium. After preliminary purification by differential centrifugation, the virus was treated with deoxyribonuclease and then with ribonuclease in order to remove contaminating nucleic acids. Further purification by differential centrifugation was carried out until the pellet of sedimented virus appeared completely transparent and colorless. A solution of purified virus with unit optical density at 2600 A was found to possess an infective titre of $1.8 \times 10^{11}$ per ml. This indicates a ratio of infectivity to nucleic acid content 20 per cent higher than the most highly purified preparations described by other workers (8).

The viral DNA was released from its protein coat by osmotic shock. For this purpose a suspension of virus of infective titre $1$ to $5 \times 10^{13}$ per ml was mixed with an equal volume of saturated CsCl

solution. After the mixture remained standing one hour in the cold, its osmotic pressure was rapidly lowered by dumping onto it 15 volumes of water. Thereupon the solution became viscous and elastic and lost most of its initial turbidity indicating that the bacteriophage had released their DNA into solution. This procedure was successfully applied to as few as $10^{11}$ bacteriophage in an initial volume of 0.002 ml. The viscous DNA solution was stored in the cold with a trace of chloroform to prevent the growth of micro-organisms. Small aliquots were added directly to CsCl solutions of density appropriate for the banding of DNA. The viral protein, being much less dense than DNA, is separated from the DNA during the centrifugation in CsCl thereby avoiding the need of a previous separation.

(b) 5-Bromouracil bacteriophage DNA. T4 r240 bacteriophage containing DNA substituted with 5-bromouracil was prepared by the method of Litman and Pardee (9). The virus was purified and osmotically shocked as outlined above. The substitution of thymine by 5-bromouracil was about 60 per cent complete as determined chromatographically (10).

(c) Calf thymus*, human leukocyte** and salmon sperm DNA***. These DNA samples were prepared in other laboratories by detergent methods

---

These samples were kindly supplied by

* Dr. C. Jardetsky

** Dr. S. Perry

*** Prof. P. Doty

followed by alcoholic precipitation. Before being used in banding experiments, the solid DNA was allowed at least five days to pass completely into solution.

Studies of various DNA samples. Only the DNA from normal bacteriophage was found to form the Gaussian band expected for a single molecular species (fig. 1). Aside from the remote possibility that the concentration distributions for several molecular species could have combined to give the observed Gaussian distribution, it may be concluded that this viral DNA is homogeneous. The mean of its distribution indicates an effective density of 1.70 $g/cm^3$ and the standard deviation corresponds to a molecular weight for the Cs.DNA salt of $18 \times 10^6$. Assuming the base composition of T4 reported by Wyatt and Cohen (11) and the glucose content reported by Sinsheimer (12), this corresponds to a molecular weight of $14 \times 10^6$ for sodium deoxyribonucleate. If this material is representative of the DNA contained in intact bacteriophage its molecular weight corresponds to about twelve molecules of DNA per virus particle. This estimate rests upon a value of $4.0 \pm 0.4 \times 10^5$ phosphorus atoms per phage (13).

Bands of human leukocyte, salmon sperm (fig. 5) and calf thymus (fig. 3) DNA were all found to be skewed in the direction of higher effective density. Inasmuch as these DNA samples had been prepared by a method other than that used to obtain the virus DNA, it cannot be decided at present whether the observed density heterogeneity is native or had been introduced during isolation.

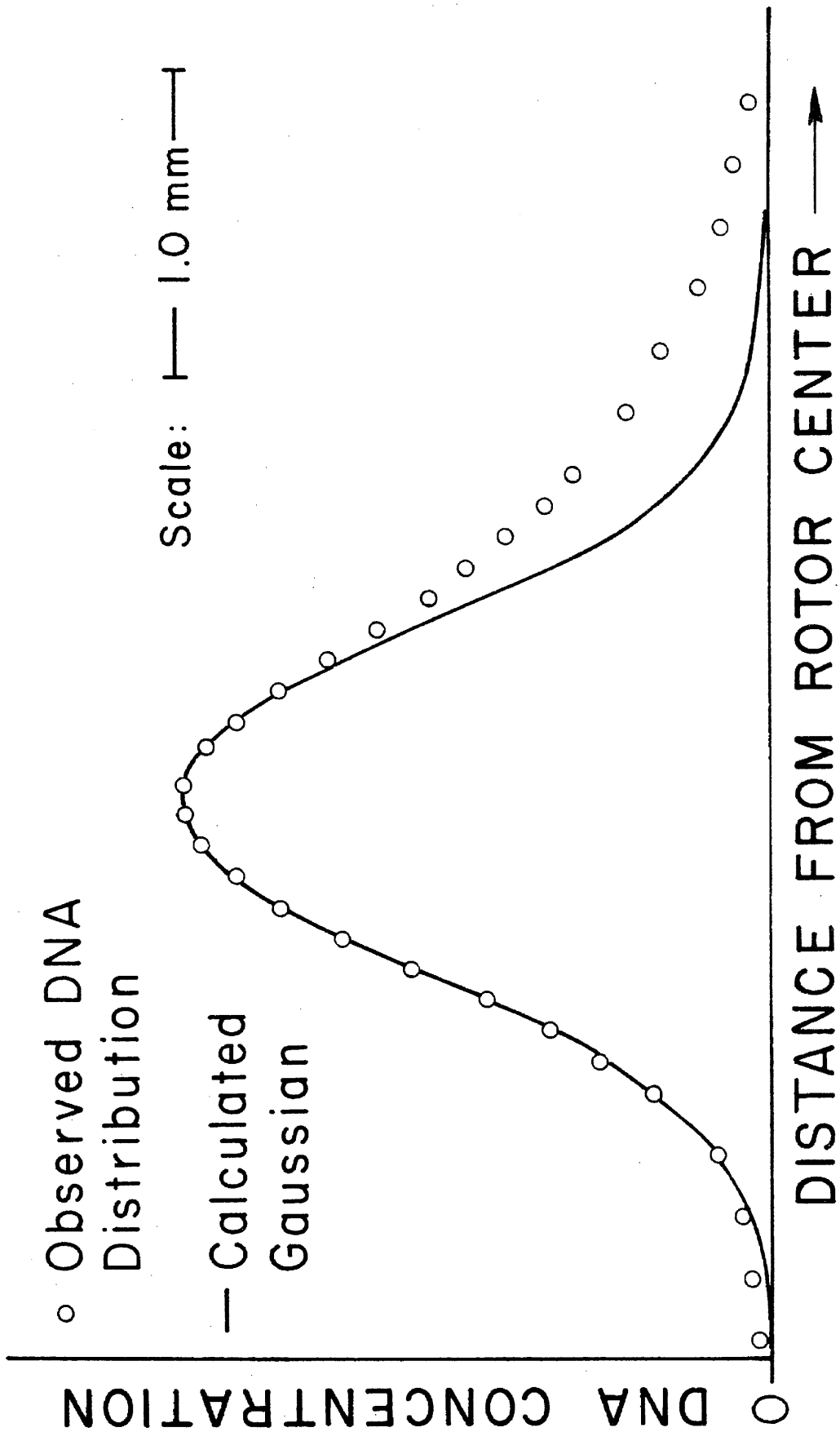

DNA CONCENTRATION
DISTANCE FROM ROTOR CENTER
Scale: 1.0 mm
o Observed DNA Distribution
— Calculated Gaussian

Fig.5

The sample of salmon sperm DNA was supplied by Prof. Paul Doty, who assigned to it a molecular weight from light scattering measurements of 8 ± 1 million. The molecular weight of this material was studied by density gradient centrifugation in order to provide a comparison with an established method of molecular weight determination. Fig. 5 shows the concentration distribution of the Doty salmon sperm DNA in a CsCl gradient. The pronounced skewness in the direction of higher effective density precludes any but a minimal estimate of the molecular weight. Because of the limited available machine time, attainment of equilibrium was not verified by the experimental method employed for the viral DNA. Instead, the centrifugation was continued for 106 hours, which is 30 per cent longer than the time found adequate for equilibration of phage DNA at the same field strength and temperature. Nevertheless, in view of the lower molecular weight and correspondingly lower sedimentation velocity of the salmon DNA, this time may not have been sufficient. Non-attainment of equilibrium would result in a low estimate for the molecular weight. The molecular weight estimate of four million for this material was obtained from the standard deviation of a calculated Gaussian fitted to the low density side of the observed concentration distribution as shown in the figure. Accepting the light scattering value with its associated limit of error, the discrepancy may be due to the factors cited above. Additionally, some degradation may have occurred in this material during the time subsequent to its study by light scattering.

Fig. 2 shows the concentration distribution in a CsCl gradient of a mixture of 5-bromouracil phage DNA and normal phage DNA added as a

density marker. Experiments with each material separately show that the distribution of the mixed material is the arithmetic sum of the two individual distributions. The increase in density to be expected upon complete substitution of 5-bromouracil for thymine in phage DNA may be estimated under the assumption that normal and 5-bromouracil DNA possess the same effective volume in CsCl solution. The fractional density change is then just the ratio of the molecular weights of the two varieties of DNA. For normal phage DNA the mean molecular weight per nucleotide is 360. Thymine comprises 32 per cent of the nucleotides. The molecular weight ratio of substituted to non-substituted DNA is thus 1.06 and a density increase of 0.10 $g/cm^3$ is to be expected upon complete 5-bromouracil substitution. The increase actually found from the positions of the normal and 5-bromouracil DNA bands in the experiment of fig. 2 is 0.10 $g/cm^3$, in striking agreement with the predicted value. The measurement of the density difference between normal and 5-bromouracil substituted DNA assumes that the maximum density observed corresponds to complete substitution. This is plausible in view of the rather sharp cut-off at the high density edge of the 5-bromouracil DNA band. Furthermore, the mean of the distribution of the substituted DNA is located 2/3 of the distance from the normal peak to the position of presumably complete substitution. If the density is proportional to the fraction of substitution, the location of the mean of the 5-bromouracil DNA distribution indicates the material to be 2/3 substituted. The chromatographically established degree of substitution is 0.6. In view of the difficulty involved in the chromatography of the small amount of 5-bromouracil DNA available, this

fair agreement may be regarded as confirmatory.

Additional applications. The usefulness of the method for biological studies is illustrated by the study of intact viruses in cesium chloride gradients. Both bacteriophage and tobacco mosaic virus (14) have been banded without loss of infectivity. Preliminary work with bovine albumin and human hemoglobin suggests the applicability of this method to smaller macromolecules.

References

1) R. A. Pasternak, G. M. Nazarian, J. R. Vinograd, Nature 179, 92 (1957).

2) "International Critical Tables" Vol. 3, McGraw-Hill Co., New York, 1928, p. 94.

3) P. A. Lyons and J. F. Riley, J.A.C.S. 76, 5216 (1954).

4) R. A. Robinson and R. H. Stokes, "Electrolyte Solutions", Butterworths, London, 1955, p. 489.

5) T. Svedberg and K. O. Pedersen, "The Ultracentrifuge", Oxford, 1940, p. 248.

6) S. H. Goodgall, C. S. Rupert, and R. M. Herriott, Trans. Far. Soc. 53, 257 (1957).

7) P. Doty and B. H. Bunce, J.A.C.S. 74, 5029 (1952).

8) A. D. Hershey, Virology 1, 108 (1955).

9) R. M. Litman and A. B. Pardee, Nature 178, 529 (1956).

10) S. Zamenhof and G. Griboff, Nature 174, 306 (1954).

11) G. R. Wyatt and S. S. Cohen, Nature 170, 1072 (1952).

12) R. L. Sinsheimer, Proc. Nat. Acad. Sci. 42, 502 (1956).

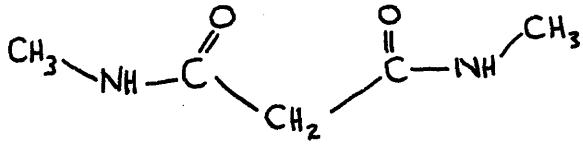
13) A. D. Hershey, J. Gen. Physiol. 36, 777 (1953).

14) A. Siegel, private communication (1957).

## II. THE CRYSTAL STRUCTURE OF N,N'-DIMETHYL MALONAMIDE

### A. Introduction

The determination of the structure of N,N'-dimethyl malonamide



provides additional data of significance to the structure of proteins. This compound provides a case of N-substitution intermediate between that in peptide amides and terminal amides, and the presence of the methyl substituent facilitates X-ray investigation of the configuration about the nitrogen atoms and the planarity of the amide groups.

### B. Preparation and Crystallization

N,N'-Dimethyl malonamide was prepared by the addition of a stoichiometric quantity of diethyl malonate to a 25 per cent aqueous solution of methylamine. Evaporation to near dryness, extraction of the residue with hot benzene, and subsequent cooling gave a crystalline product melting at 128-129°C. Recrystallization of the crude product from benzene gave very thin plates which melted at 134-135°C, in agreement with the published value (1). More nearly equidimensional crystals were obtained from benzene solutions containing a few per cent of ethanol and these were used throughout the present study. Laue photographs of crystals grown from the two solutions were identical. The crystals were bounded by (110) and (001) faces and showed perfect cleavage parallel to the latter.

C. Unit Cell and Space Group

The unit cell dimensions obtained from rotation photographs taken in a Straumanis-type camera with Cu $K_\alpha$ radiation ($\lambda = 1.542$ Å) are, with their estimated limits of error:

$\underline{a} = 8.656 \pm 0.017$, $\underline{b} = 4.56 \pm 0.015$, $\underline{c} = 18.508 \pm 0.08$ Å, $\beta = 94.22 \pm 0.12^\circ$

The density determined by flotation is 1.18 g/cm$^3$ and the number of molecules per unit cell is 4 (calculated 4.01). Precession and Weissenberg photographs reveal systematic absences $\underline{hkl}$ for $\underline{h} + \underline{k}$ odd and $\underline{h0l}$ for $\underline{l}$ odd. These absences are characteristic of the non-centrosymmetric space group Cc ($n=4$) as well as the centrosymmetric space group C2/c ($n=8$). No piezoelectric effect could be detected although a pronounced effect was produced by the weakly piezoelectric substance hexamethylene tetramine (2). The morphology of the crystals and the result of the piezoelectric experiment do not rule out the presence of a center of symmetry.

D. Trial Structure

The search for a trial structure was begun on the assumption that the space group of the crystal is C2/c rather than the acentric Cc. This requires that the four molecules must be located at special positions in the unit cell. The only positions consistent with the possible molecular symmetry of N,N'-dimethyl malonamide are those in which the central carbon atom is located on a two-fold axis of the space group. If the amide groups are planar and in the trans configuration, and if the

bond lengths and angles agree with previously determined values for related compounds, only three parameters are required to determine the positions of all heavy atoms. These parameters may be chosen as (1) the angle between the c axis and the projection of the $C_o$-$C_1$ bond upon the ac plane, (2) the angle between the planes of the two amide groups, and (3) the elevation of the molecule above the ac plane (fig. 1). Considerations based on the unit cell dimensions and the requirement of good packing and hydrogen bonding, serve to fix approximately the first two of these parameters. Their assignment specifies the x and z coordinates of each atom and results in an arrangement of the molecules into hydrogen-bonded layers parallel to the cleavage plane (001) (fig. 2). Because of the operation of the c-glide, the specification of the elevation of the molecule above the ac plane determines the mode of stacking of successive hydrogen bonded layers. This was established by considerations of methyl-group packing supplemented by a Patterson projection onto (100). As a check on the trial structure, some low-order structure factors were calculated and these displayed encouraging agreement with observed values.

E. Intensity Data

Two well-formed crystals of uniform dimensions were mounted for rotation about the a and b axes respectively. Neither crystal exceeded 0.2 mm in any dimension. Multiple film Weissenberg (h0l), (0kl), and (1kl) equi-inclination photographs were taken with filtered Cu $K_\alpha$ radiation. The photographs displayed a total of 229 independent reflections. In addition, a system of very weak diffuse reflections could be seen on all highly exposed photographs (fig. 3). The ordered structure has been refined without

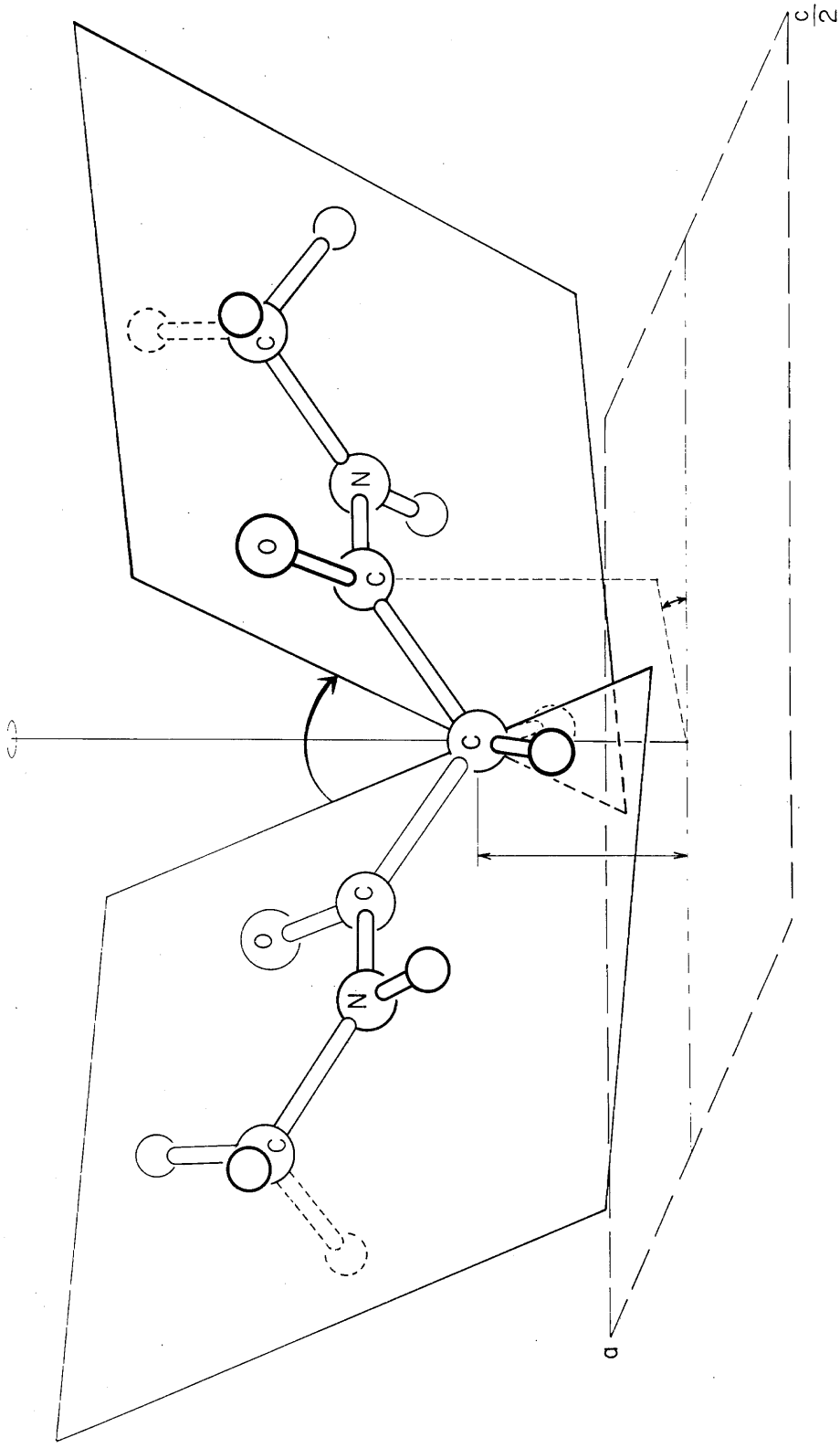



Fig. 1 THE MOLECULAR CONFIGURATION OF N,N'-DIMETHYL MALONAMIDE

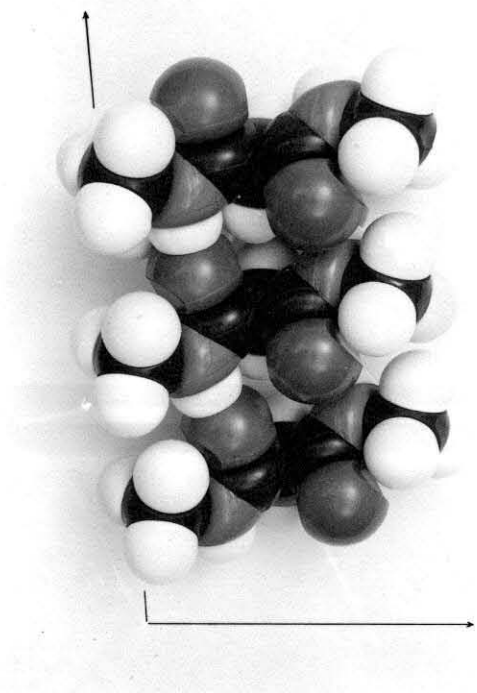


FIG. 2 MOLECULES IN A HYDROGEN-BONDED LAYER PARALLEL TO THE ab PLANE AS VIEWED DOWN THE b AXIS

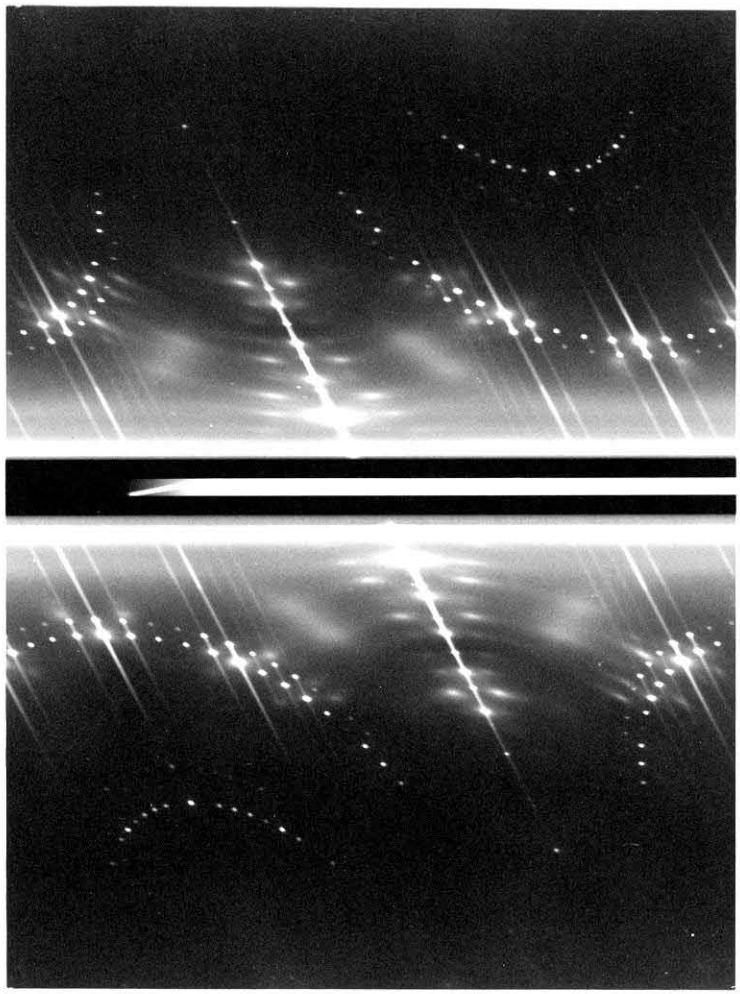


FIG. 3 AN (okl) WEISSENBERG PHOTOGRAPH SHOWING DIFFUSE MAXIMA

regard to them. This neglect is justified by the satisfactory appearance of the final difference maps.

All intensities were visually estimated on two separate occasions. The quantity $\frac{2\sum|F_1^2-F_2^2|}{\sum(F_1^2+F_2^2)}$ was 0.09. The two observations of each intensity were averaged and corrected for Lorentz and polarization factors. Because of the small dimensions of the crystals, no absorption correction was applied. Values of the observed structure factors are given in Appendix I.

F. Refinement of Parameters

The refinement of the projection onto (010) was begun with the calculation of a Fourier projection based on about half of the observed ($h0\bar{l}$) reflections. All heavy atoms were shown clearly resolved in positions consistent with the trial structure. Atomic centers located by parabolic interpolation and a scale and isotropic temperature factor ($B = 6$) obtained from a plot of $\log \frac{F^2_{obs.}}{F^2_{calc.}}$ against $\sin^2\theta$ were employed in the computation of a second set of ($h0\bar{l}$) structure factors. These served to specify the signs of nearly all the remaining observed structure factors. A further cycle of Fourier refinement yielded calculated structure factors for which the discrepancy factor R was 0.30. Three cycles of diagonal least-squares refinement of positional parameters were then carried out. At this stage the positional parameters had converged even though the discrepancy factor was still 0.23. Upon closer examination of the observed and calculated structure factors it was found that, for most reflections, $F(h0\bar{l})$ calculated too high and $F(h0l)$ too low, suggesting an

over-all anisotropic temperature factor. In order to obtain a more certain indication of the anisotropy, an $\underline{h}0\underline{l}$ difference map was prepared. Shifts of less than 0.02 A in atomic positions, a small scale factor correction, and individual anisotropic thermal corrections were obtained from the difference map by the method given in Appendix II. The axes of anisotropy were nearly parallel for all atoms, and the indicated displacement was largest for O and smallest for $C_2$. The individual atomic corrections were averaged to give the over-all temperature factor $\exp-(0.0147\underline{h}^2 + 0.0055\underline{l}^2 + 0.0044\underline{hl})$. This corresponds in projection to a direction of maximum vibration approximately normal to the cleavage plane. Hydrogen atoms $H_1$ and $H_2$ were assigned to positions consistent with reasonable bond lengths and angles in accord with the indications of the difference map. Upon introduction of the positional and thermal corrections and the hydrogen contributions, the discrepancy factor dropped sharply to 0.17. A final $\underline{h}0\underline{l}$ difference projection (fig. 4) was prepared which indicated barely significant positional shifts. Positive peaks attributable to the methyl hydrogens $H_3$, $H_4$, and $H_5$ were apparent but there were no other regions of electron density greater than 0.5 electron $\text{Å}^{-2}$. The introduction of the indicated shifts and of the contribution of these three additional hydrogen atoms (see below) further reduced the discrepancy factor to 0.15 (neglecting unobserved structure factors). Refinement of the projection was now judged satisfactory.

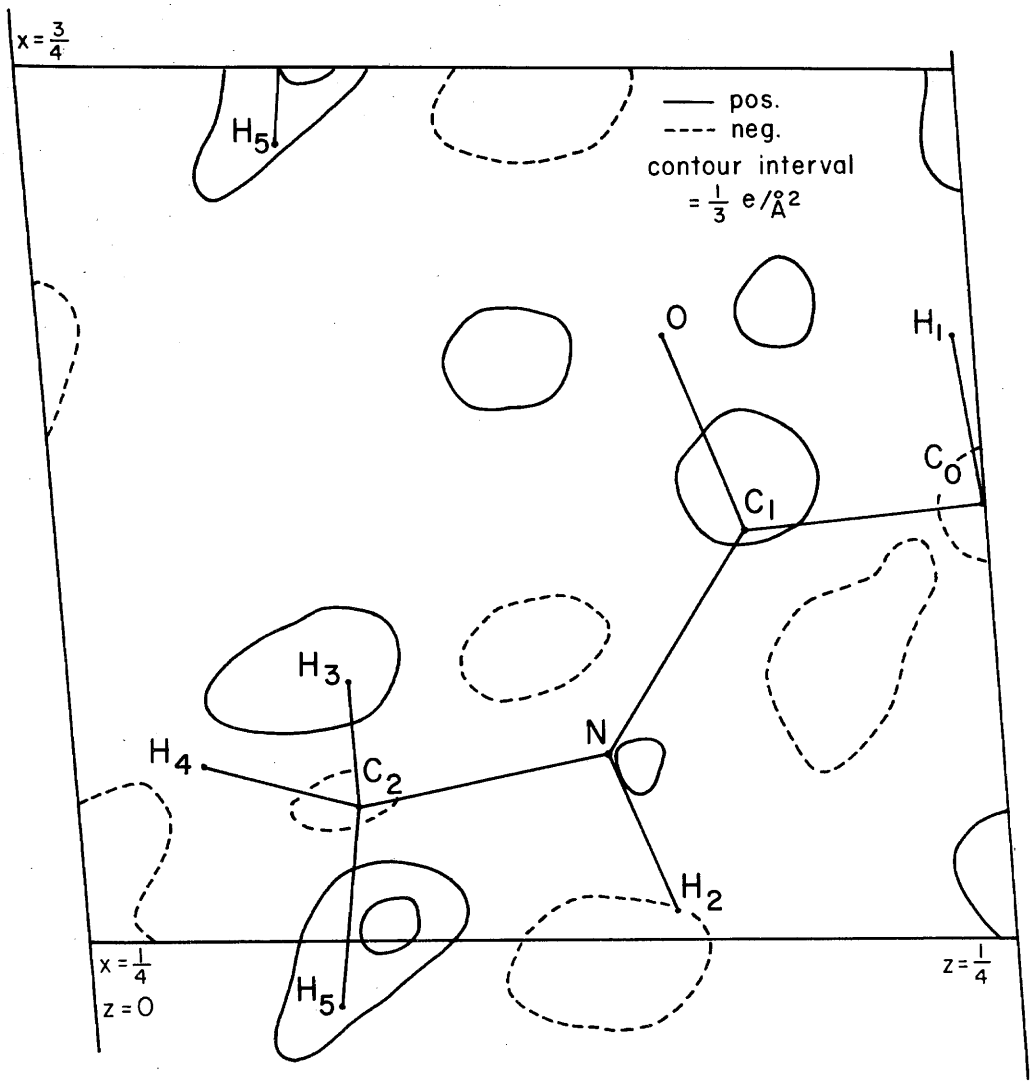



Fig. 4 THE FINAL *h*0*l* DIFFERENCE FOURIER PROJECTION

It was anticipated that satisfactory refinement of the 0kl projection would be prevented by overlap between $C_1$, N, and O and by the relatively small number of observed 0kl reflections. It was therefore decided to base refinement on the 1kl data which, in addition to being more numerous, present an opportunity to use a novel form of weighted density projection especially suited for the reduction of overlap.

Weighted density projections. A weighted density projection is one in which the electron density $\rho(xyz)$ is weighted by some function $g(xyz)$ before projection. In normal projections the weighting function is simply a constant; a section may be considered to be a weighted projection in which the weighting function is zero everywhere except in the plane of the section. The normal projection suffers from the complications of atomic overlap but possesses the advantage of being easily calculated from zero-layer data alone; the section is entirely free of overlap difficulties but requires considerable computational effort and the use of a full set of three-dimensional data. Aside from these two well-known projections, the weighted density projections of greatest practical value are those which, although allowing interfering atoms to be projected with reduced weight, nevertheless may be simply calculated from substantially less than the full three-dimensional set of data. Hereafter simple projections and sections will not be described as weighted density projections. Although weighted density projections have been used to determine atomic coordinates along the axis of projection (3), we shall not discuss this application but shall instead consider weighted projections

for the selection and refinement of coordinates in the plane of projection* (4). We shall assume, as was the case in the refinement of the N,N'-dimethyl malonamide structure, that the atomic positions along the axis of projection have previously been determined, at least approximately.

The most general real weighting function having the periodicity of the lattice and with constant value in planes parallel to the plane of projection, taken as yz, is

$$1) \qquad g(x) = \sum_{n=0}^{\infty} A_n \cos 2\pi n(x-\delta_n)$$

where $A_n$ and $\delta_n$ are constants. The corresponding weighted projection of the electron density is given by

$$\rho_g(y,z) = \int_0^1 \rho(x,y,z)\, g(x)\, dx$$

$$2) \qquad = \int_0^1 \mathcal{R}\mathcal{R} \sum_{h=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} F(hkl) \exp\left[2\pi i(hx+ky+lz)\right] \cdot \sum_{n=0}^{\infty} A_n \cos 2\pi(x-\delta_n)\, dx$$

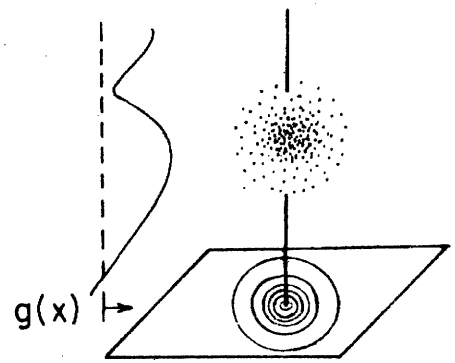
$$= \sum_{n=0}^{\infty} \tfrac{1}{2} A_n \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} \Big[ \left(F(nkl) + F(\bar{n}kl)\right) \cos 2\pi n\delta_n \cos 2\pi(ky+lz) + \left(F(nkl) - F(\bar{n}kl)\right) \sin 2\pi n\delta_n \sin 2\pi(ky+lz) \Big]$$

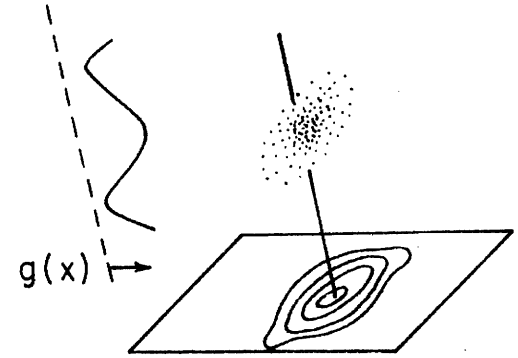
---

* But see the note on "modular" projections in Appendix III.

In general, an isolated atom will be shown in projection as an extremum in the correct position only if the atom is cylindrically symmetric about an axis parallel to the projection axis which must itself be normal to the projection plane (5). However, these severe requirements are relaxed in the special case that $g(\underline{x}_m + \underline{x}) = g(\underline{x}_m - x)$ where $\underline{x}_m$ is the coordinate of a particular atom. Then an isolated centrosymmetric atom will be projected without peak shift (fig. 5). The peak shape will of course depend on both $\rho(\underline{xyz})$ and $g(\underline{x})$ and in some cases may be sharpened in comparison with a simple projection (6). If the above conditions for unshifted peak projection of isolated atoms are fulfilled, then equation 2 may be used for refinement of $\underline{y}$ and $\underline{z}$ parameters of significantly weighted atoms in a manner essentially the same as that employed with a simple Fourier projection. A convenient way to compensate for series termination errors and overlap not remedied by the use of the weighting function is to make use of the weighted difference projection which is obtained from equation 2 by replacing F by $(F_{obs.} - F_{calc.})$. Corrections to the positional, thermal, and scale parameters may be obtained from a weighted difference projection by the method given in Appendix II.

Refinement of parameters in the yz plane. During the refinement of the projection along the a axis, both $F_o$ and $F_o - F_c$ syntheses were prepared from the (1kl) data. According to equation 2, the use of (1kl) Fourier coefficients restricts the value of n in equation 1 to unity so that the weighting function becomes $\cos 2\pi(\underline{x} - \underline{x}_m)$. This proves to be very convenient for, by varying $\underline{x}_m$, the cosine weighting function can be positioned in the cell so as to minimize troublesome overlap in the region




Atom must be cylindrically symetric about projection axis

Projection axis must be normal to projection plane

Arbitrary weighting function

Atom must be centrosymetric

Projection axis may be inclined to projection plane

Weighting function must be symetric about atomic center



Fig. 5 CONDITIONS FOR UNSHIFTED PEAK PROJECTION

of any given atom. We shall call Cos $2\pi(\underline{x}-\underline{x}_m)$ a sliding weighting function to emphasize this possibility. According to equation 2 the weighted projection of electron density is given in this case by the expression

$$\rho_1(yz) = \frac{1}{2}\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\Big\{\big[F(1kl)+F(\bar{1}kl)\big]\cos 2\pi(ky+lz)\cos 2\pi\delta$$

3)

$$+\big[F(1kl)-F(\bar{1}kl)\big]\sin 2\pi(ky+lz)\sin 2\pi\delta$$

By choosing $\delta$ equal to $\underline{x}_m$, the $\underline{x}$ parameter of a given centrosymmetrical atom, that atom is projected with maximum weight and without peak shift. In the present case with $\underline{a}$ = 8.66 Å, Cos $2\pi(\underline{x}-\underline{x}_m)$ does not vary rapidly enough in the region of an atom at $\underline{x}_m$ to require that $\delta$ be precisely equal to the atomic coordinate. In fact, slight displacements may intentionally be introduced in order to give minimal weight to interfering atoms. In practice, the two projections (figs. 7, 8)

4a) $$\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\big[F(1kl)+F(\bar{1}kl)\big]\cos 2\pi(ky+lz)\cos 2\pi\delta$$

and

4b) $$\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\big[F(1kl)-F(\bar{1}kl)\big]\sin 2\pi(ky+lz)\sin 2\pi\delta$$

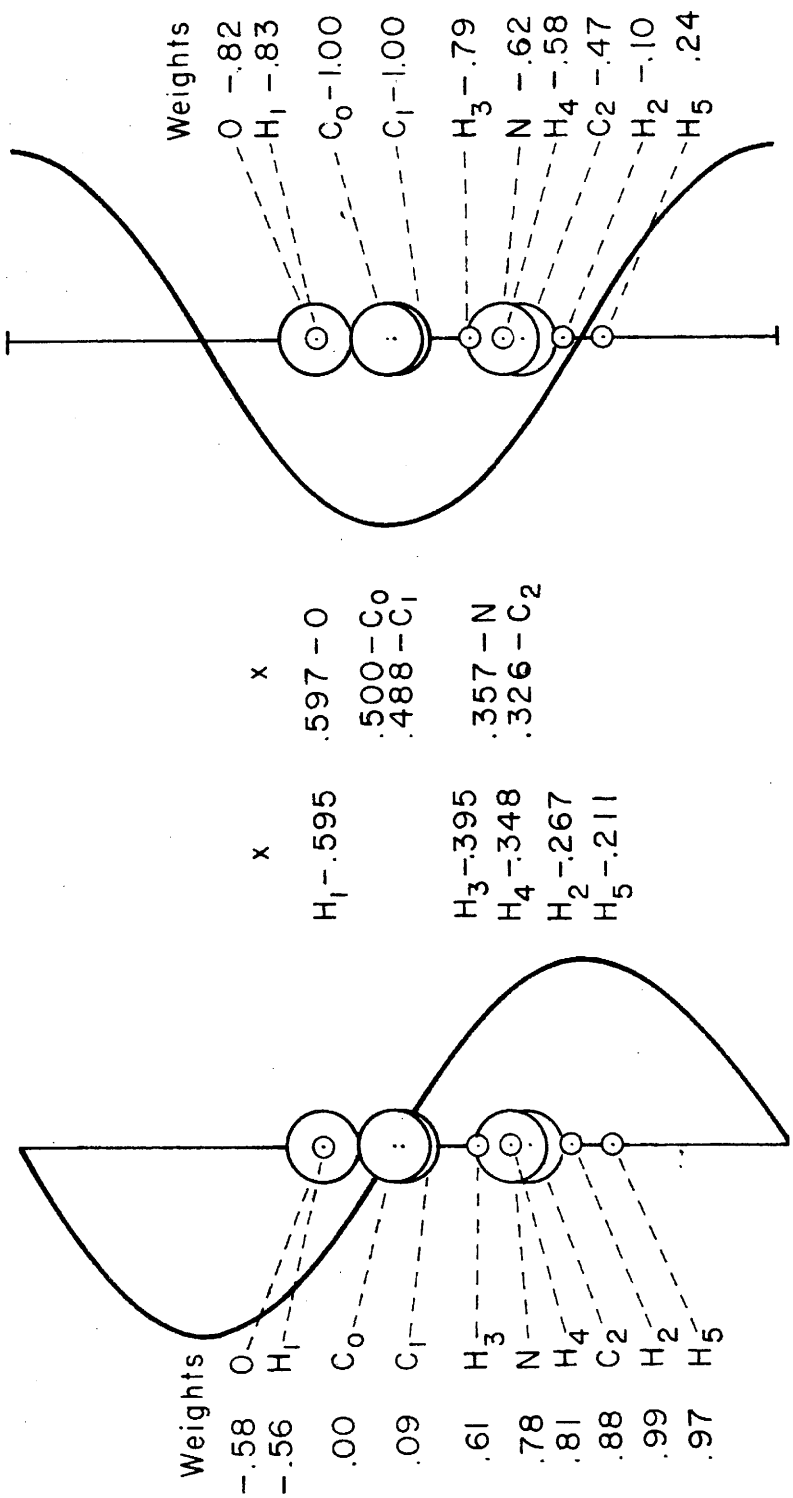



Fig. 6 Variation of the Sine and Cosine Weighting Functions along the a Axis.

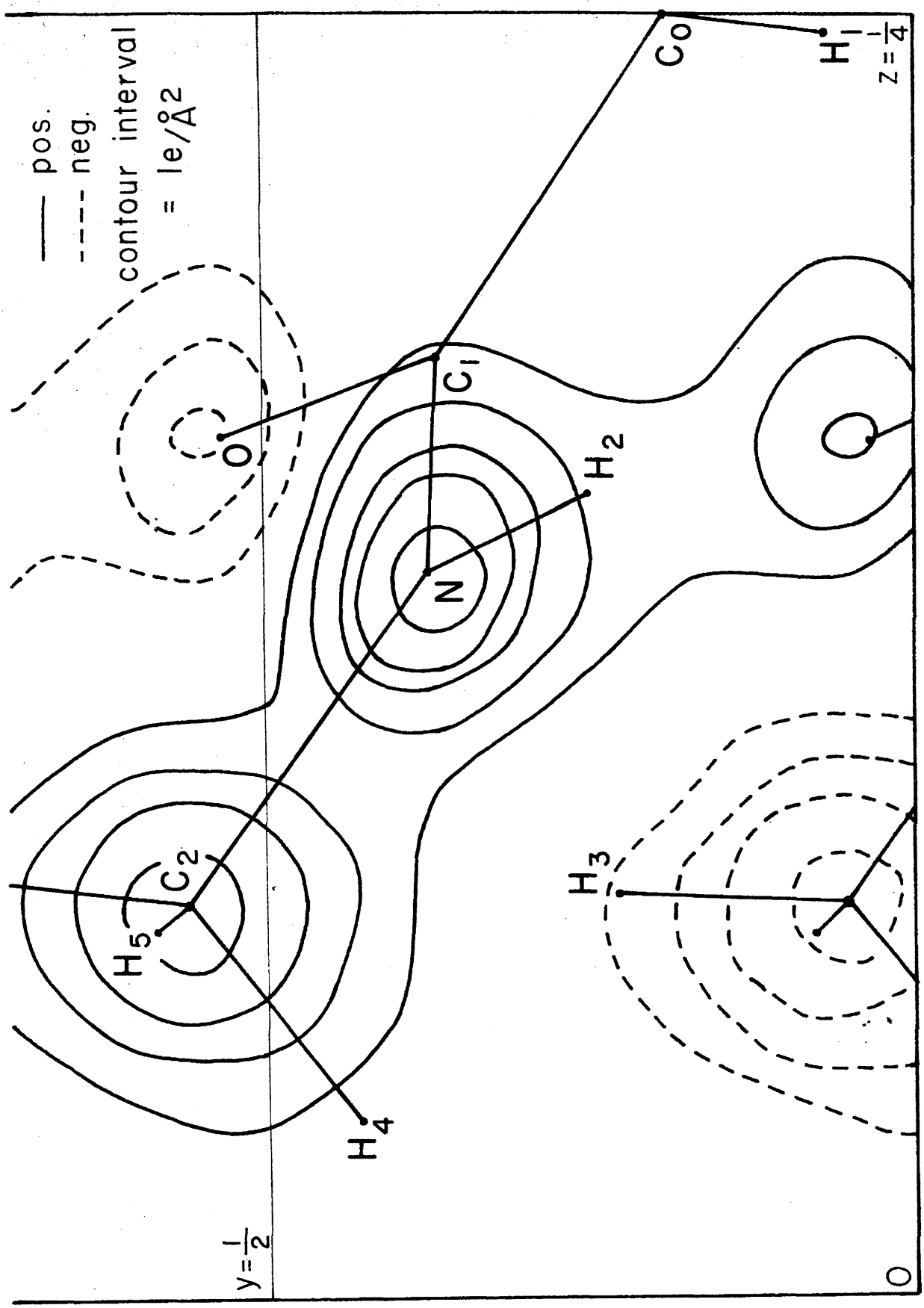



Fig. 7 |k| Sine Projection

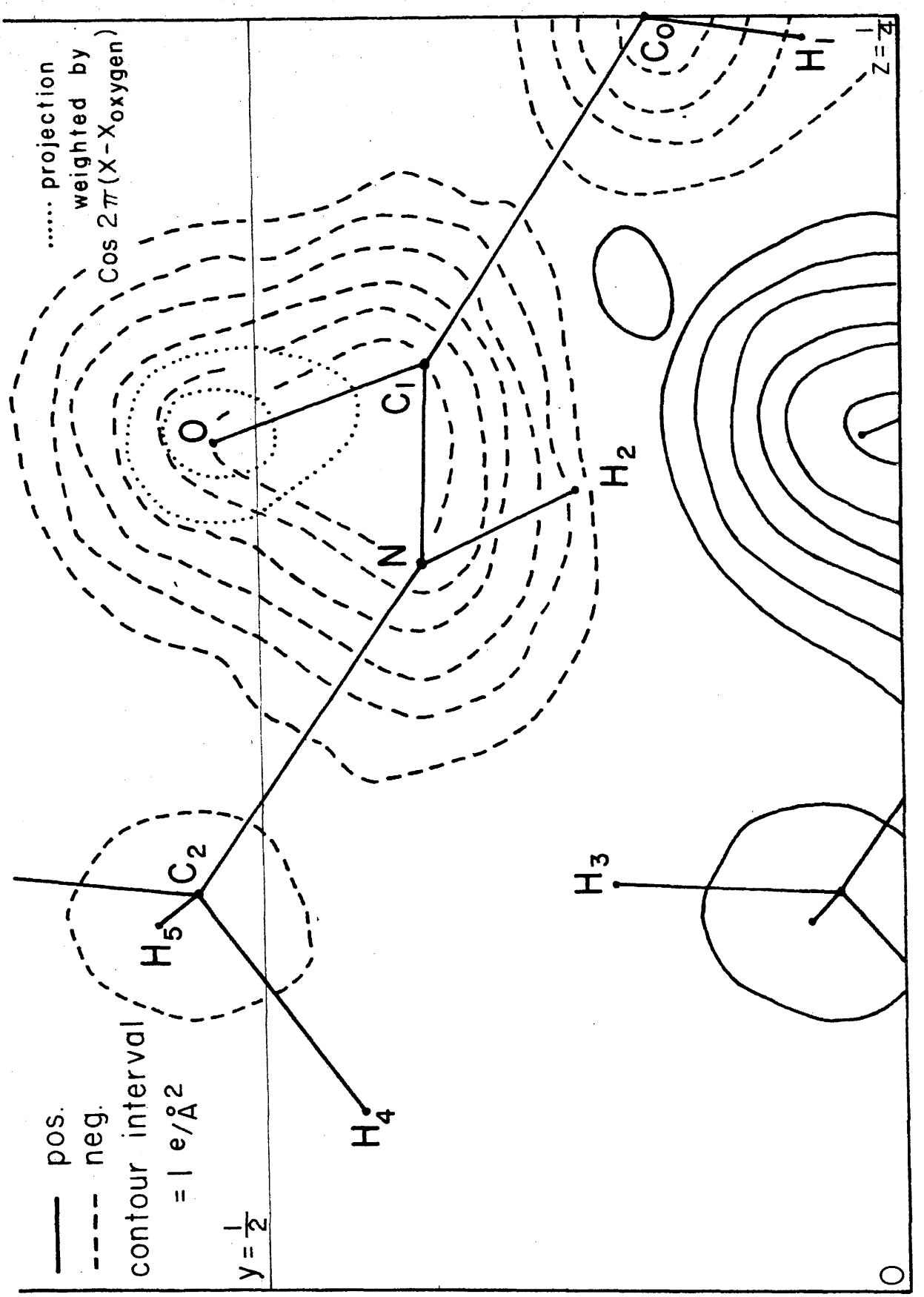



Fig. 8 1kl COSINE PROJECTION

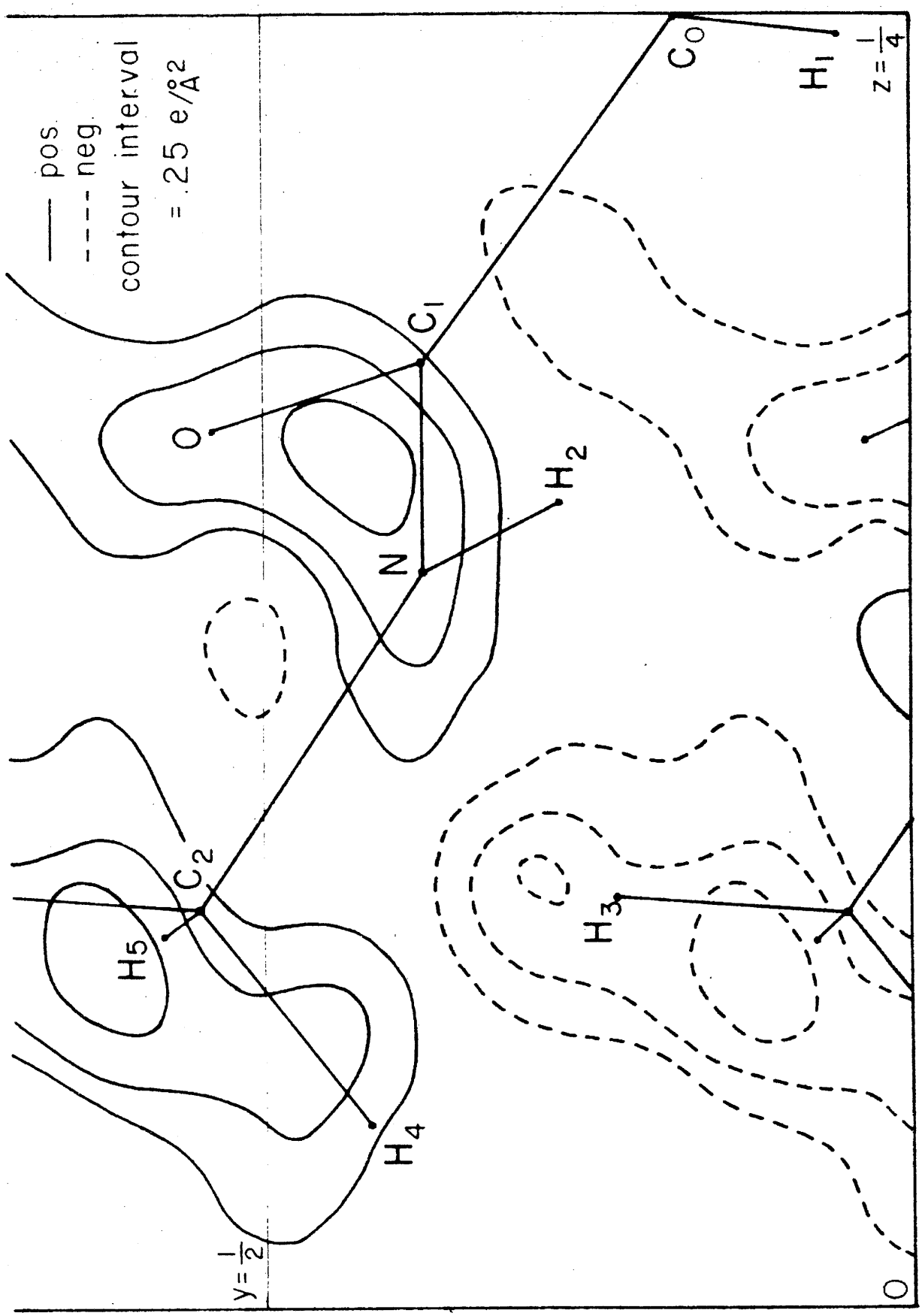



Fig. 9 |K| DIFFERENCE SINE PROJECTION

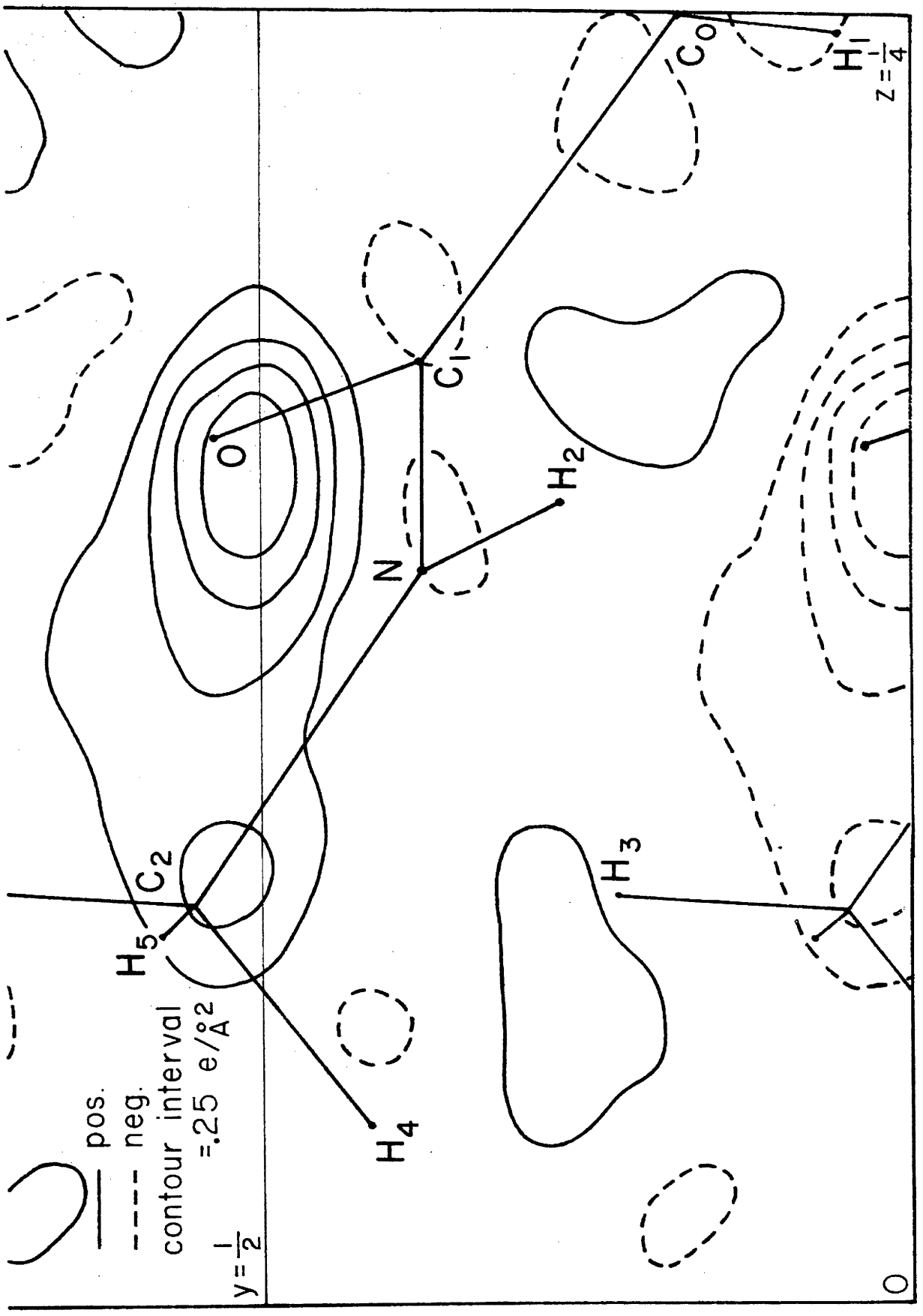
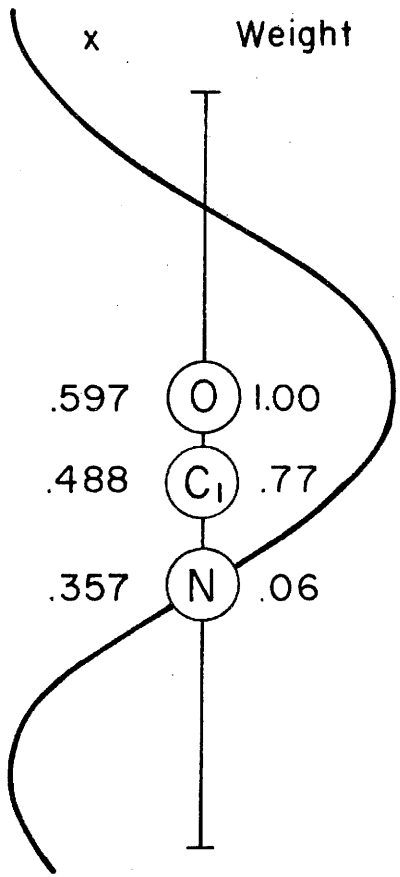
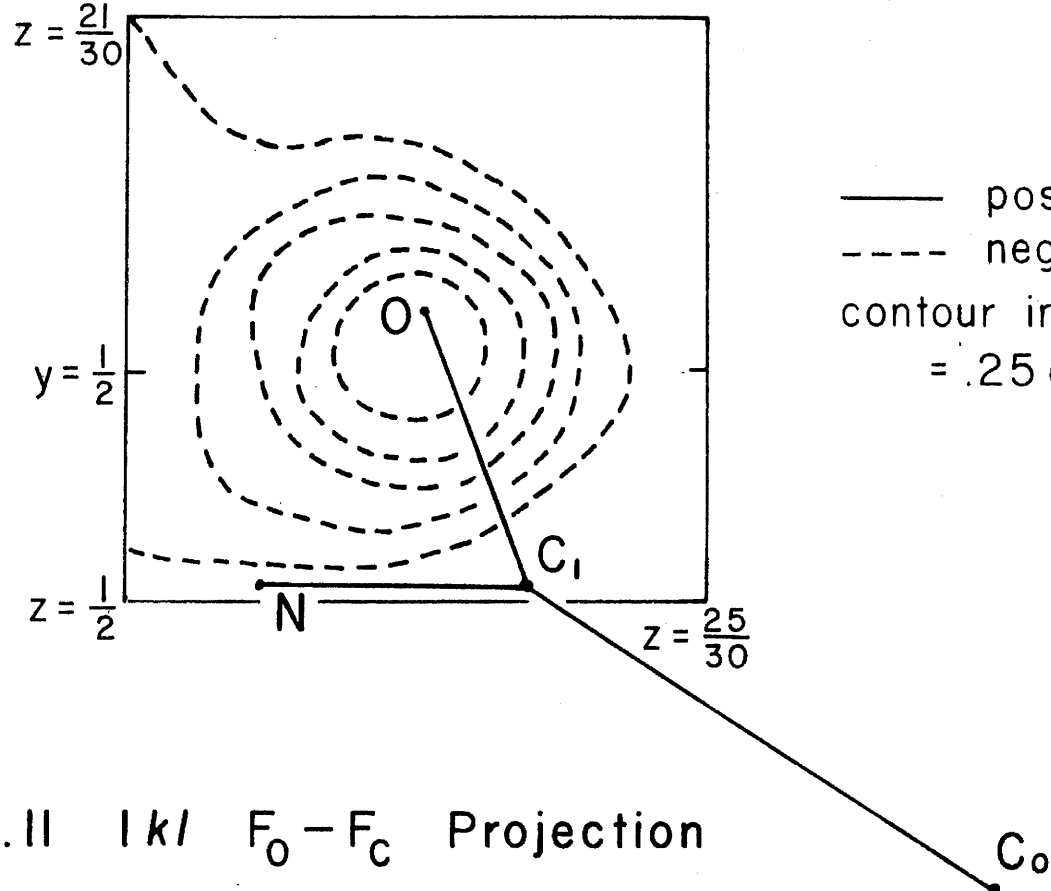



Fig. 10 $|k|$ DIFFERENCE COSINE PROJECTION




Weights for accompanying projection




—— pos.
---- neg.
contour interval
= .25 e/Å²

Fig. II $|kl|$ $F_o - F_c$ Projection weighted by Cos $2\pi(X - X_{oxygen})$

or the analogous difference projections (figs. 9 and 10) were prepared and then combined in the ratio $\frac{\text{Cos } 2\pi\delta}{\text{Sin } 2\pi\delta}$ in order to place the sliding weighting function Cos $2\pi(\underline{x}-\delta)$ in the desired position. The necessary labor was kept to a minimum by combining the cosine and sine projections only in the neighborhood of the atom under examination.

Using $\underline{x}$ and $\underline{z}$ parameters from the refined $\underline{h}0\underline{l}$ projection and $\underline{y}$ parameters from the trial structure, $1\underline{k}\underline{l}$ structure factors were calculated; after adjustment of scale and temperature factors, R was 0.30. Signs were assigned to about 3/4 of the observed $1\underline{k}\underline{l}$ reflections which were used as coefficients in the calculation of the sine and cosine weighted projections (equations 4a and 4b). These two projections were combined in the region of each heavy atom in the manner described above. Peaks were located by parabolic interpolation and new structure factors were calculated with the inclusion of contributions from the hydrogens $H_1$ and $H_2$. The discrepancy factor dropped to 0.24. Next, a set of weighted difference projections was prepared (figs. 9 and 10). These indicated several shifts of about 0.02 Å in the $\underline{y}$ and $\underline{z}$ parameters. No anisotropic thermal vibration in the plane of projection was indicated but the isotropic temperature factor for the oxygen atom was found to be greater by about one B-unit than that for the other atoms. Peaks found in positions where methyl hydrogens might be expected were utilized along with packing considerations and the final $\underline{h}0\underline{l}$ difference synthesis to fix the positions of the hydrogens $H_3$, $H_4$, and $H_5$. Fig. 11 is a weighted difference projection along $\underline{a}$ in which the oxygen atom has been given maximum weight in order to reduce overlap by $C_1$ and N. Corrections to thermal and

positional parameters were computed from this map by the method of Appendix II.

Structure factors were calculated incorporating the positional and thermal corrections indicated by the difference maps and containing all hydrogen contributions. The scale factor and the over-all temperature factor were adjusted by the method of least squares. For atoms other than oxygen, the temperature factor used in the calculation of the final structure factors was $\exp-\left[\frac{5.5\,\mathrm{Sin}^2\theta}{\lambda^2}\right]$ for $F_{0k\underline{l}}$ and $\exp-\left[\frac{6.0\,\mathrm{Sin}^2}{\lambda^2}\right]$ for $F_{1k\underline{l}}$. For both layers an additional temperature factor $\exp-\left[\frac{0.8\,\mathrm{Sin}^2}{\lambda^2}\right]$ was applied to the oxygen atom. The final discrepancy factors were $R(0k\underline{l}) = 0.152$ and $R(1k\underline{l}) = 0.158$ (neglecting unobserved reflections and the two most intense reflections for which considerable extinction was apparent).

An encouraging check on the use of the weighted projections is found in the comparison of the z coordinates resulting from the h0l and the 1kl refinements (cf. Table 1); the maximum difference is 0.006 Å.

Table 1

Comparison of z Coordinates

| Atom | z from 1kl refinement (cell units) | z from h0l refinement (cell units) | Difference (Å) |
|---|---|---|---|
| $C_o$ | 0.2500 | 0.2500 | - |
| $C_1$ | .1832 | .1832 | 0.000 |
| $C_2$ | .0768 | .0766 | .004 |
| N | .1436 | .1434 | .004 |
| O | .1674 | .1677 | .006 |

The averaged values of the z coordinates and the other positional and thermal parameters are given in Table 2. The probable error in the length of a bond between heavy atoms is .012 Å; the probable error in a bond angle of 120 degrees is .9°. These estimates of error are based on the relation (7)

$$p.e. = 0.67\left(\frac{2\pi\sqrt{\sum_{h,l} h^2[F_o(hol) - F_c(hol)]^2}}{a A C_n}\right)$$

Here, $C_n$ is the curvature of the nth atom calculated for the well-resolved and nearly orthogonal h0l projection. The area A is equal to ac sin $\beta$. Comparison with Table 1 indicates this estimate of probable error is perhaps conservative.

Table 2

Atomic Coordinates

| Atom | x | y | z |
|---|---|---|---|
| $C_o$ | .5000 | .1836 | .2500 |
| $C_1$ | .4880 | .3770 | .1832 |
| $C_2$ | .3262 | .5503 | .0767 |
| N | .3570 | .3742 | .1435 |
| O | .5968 | .5350 | .1676 |
| $H_1$ | .595 | .058 | .249 |
| $H_2$ | .267 | .270 | .157 |
| $H_3$ | .395 | .732 | .077 |
| $H_4$ | .348 | .414 | .035 |
| $H_5$ | .211 | .605 | .071 |

Table 3

Intermolecular Distances and Angles

| Distance | N,N'-Dimethyl malonamide | Selected Values for Peptides (8) |
|---|---|---|
| $C_o$-$C_1$ | 1.52 Å | 1.53 Å |
| $C_1$-O | 1.24 | 1.24 |
| $C_1$-N | $1.30_5$ | 1.32 |
| $C_2$-N | 1.48 | 1.47 |
| N-H...O | $2.79_5$ | 2.79 ± .12 |
| $C_o$-$H_1$ | 1.00 | -- |
| N-$H_2$ | .96 | -- |
| $C_2$-$H_3$ | 1.02 | -- |
| $C_2$-$H_4$ | 1.03 | -- |
| $C_2$-$H_5$ | 1.03 | -- |
| Angle | | |
| $C_1$-$C_o$-$C_1'$ | 109° | -- |
| $C_o$-$C_1$-O | 117 | 121° |
| $C_o$-$C_1$-N | 117 | 114 |
| O-$C_1$-N | 126 | 125 |
| $C_1$-N-$C_2$ | 124 | 123 |
| $C_2$-N...O | 110 | 110 |

G. Discussion of the Structure

Interatomic distances and angles are listed in Table 3 and shown in fig. 12. The molecular configuration may be simply described by considering the two planar peptide groups to be rotated out of a common plane about the $C_o$-$C_1$ bonds while preserving two-fold rotational symmetry about

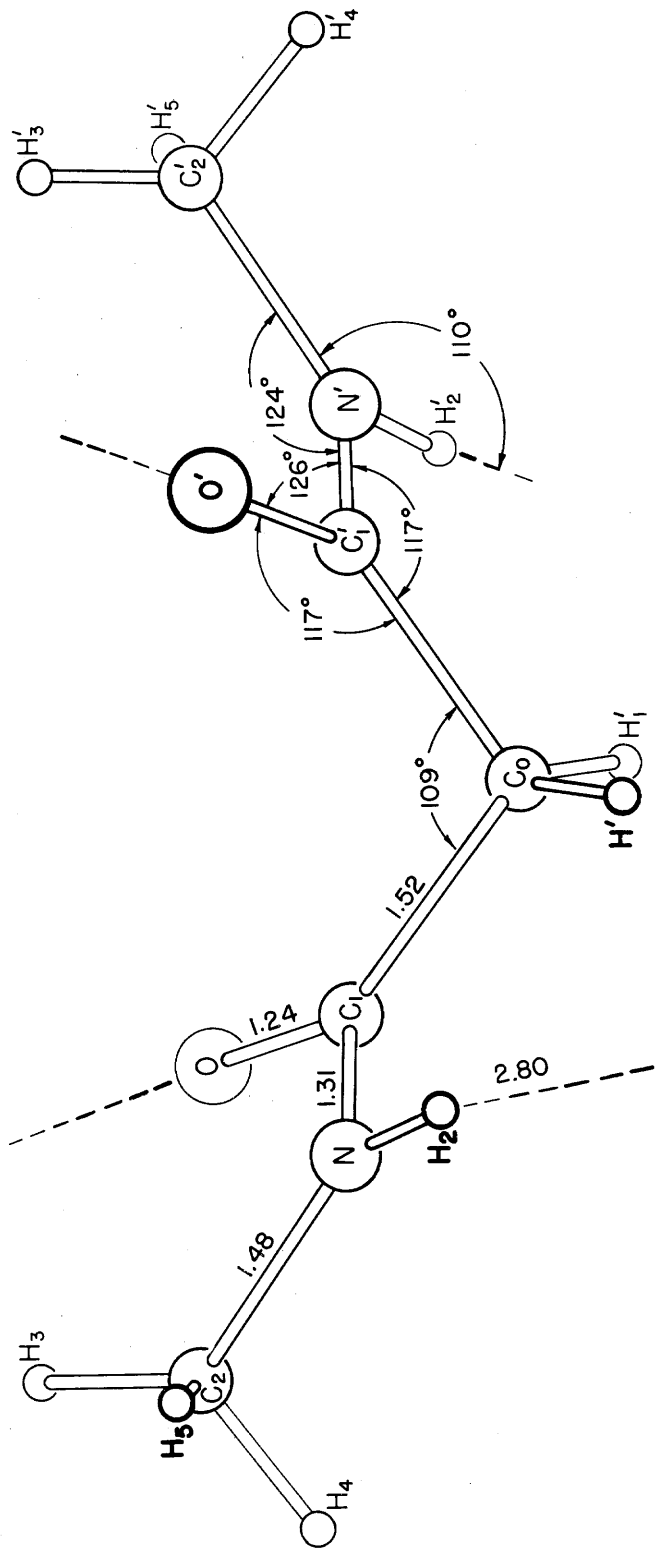



Fig.12

BOND LENGTHS AND ANGLES IN N,N'-DIMETHYL MALONAMIDE

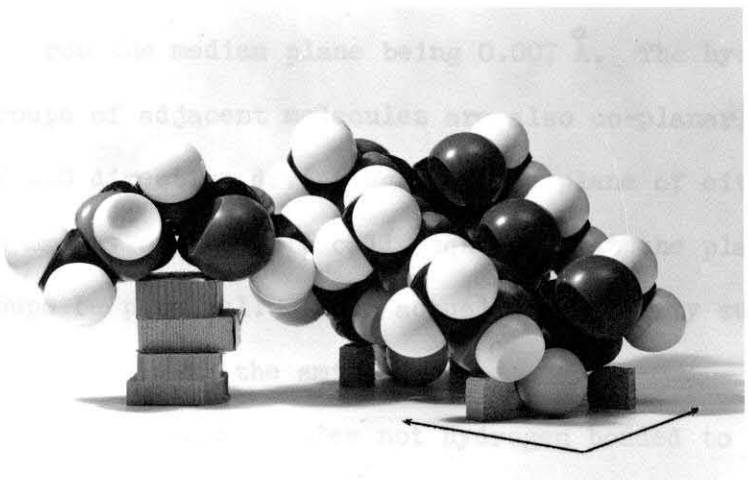


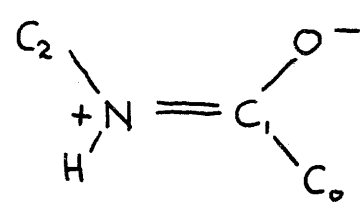
FIG. 13 THE PACKING OF MOLECULES IN A HYDROGEN-BONDED LAYER WITH A MOLECULE OF AN ADJOINING LAYER

the central carbon until the angle between the normals of the two planes is 94° (fig. 1). Each molecule is hydrogen bonded to the four neighbors related to it by the C-center translations (fig. 2). These hydrogen-bonded molecules form a layer parallel to the cleavage plane and successive layers are related by the c-glide (fig. 13).

The five atoms of the trans peptide group are co-planar, the maximum deviation from the median plane being 0.007 Å. The hydrogen-bonded peptide groups of adjacent molecules are also co-planar; the hydrogen-bonded N...O direction deviating from the plane of either by only 0.3°. The symmetry of the unit cell requires that the planes of these peptide groups be parallel. Their actual co-planarity suggests that $H_2$ is also co-planar with the amide group.

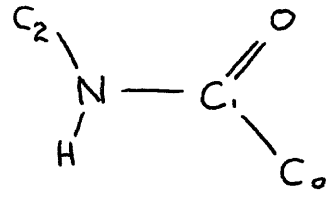
The packing between molecules not hydrogen bonded to each other (fig. 13) is determined by interatomic contact distances in good agreement with conventional van der Waals radii (9). No unusually short van der Waals contacts occur in the structure.

The value 117° found for the angle $C_o$-$C_1$-O is very significantly less than has been found for the corresponding angle in related compounds and the angle $C_o$-$C_1$-O is accordingly greater. Comparison of the observed values of the angles about $C_1$ with the selected peptide values given in Table 3 suggests that the resonance structure




I

makes an unusually large contribution to the amide group in N,N'-dimethyl malonamide. The observed $C_1$-O and $C_1$-N bond lengths in peptide amide groups studied to date agree well (8) with the assumption of 40 per cent contribution of structure I and 60 per cent contribution from the structure




II

If structure I makes more than a 40 per cent contribution to the amide group of N,N'-dimethyl malonamide, the $C_1$-N bond should be shorter than 1.32 Å and the $C_1$-O bond should be longer than 1.24 Å. The observed length (1.30$_5$ Å) of the $C_1$-N bond corresponds (10) to a 50 per cent contribution each from structures I and II. If this assignment is correct, the $C_1$-O bond length would be expected to be 1.255 Å, significantly longer than the value 1.24 Å found in this investigation. It will be shown, however, in the following section, that the actual $C_1$-O bond length has probably been underestimated by approximately this amount. It seems likely, then, that N,N'-dimethyl malonamide presents an extreme case of amide group resonance.

The oxygen atom appears to be vibrating strongly and isotropically in directions normal to the $C_1$-O bond. It is estimated from the indications of the difference maps that the temperature factor in these directions is about one B-unit greater for the oxygen than for the carbon. Because the value of the force constant for bond bending is less than for bond stretching, it is reasonable to suppose that the electron density distribution of the oxygen atom is concave toward the carbon atom. In projection this would result in an apparent shortening of the $C_1$-O distance. An estimate of the amount of contraction to be expected in such cases may be obtained under the assumptions that the bond length remains constant during the vibration and that the probability of a given displacement is a Gaussian function of its magnitude. Then, if the bond lies approximately parallel to the projection plane and executes small vibrations with mean square amplitude $\overline{u^2}$ in a plane normal to the plane of projection, the apparent contraction of the bond length is given by the expression

$$5) \qquad \Delta d = \frac{\overline{u^2}}{d} = \frac{B}{8\pi^2 d}$$

where $\underline{d}$ is the actual bond length. In the present case, with $B = 1$, this correction is 0.01 Å.

## Appendix I

Observed and Calculated Structure Factors for N,N'-Dimethyl Malonamide*

h0l reflections

| h | l | $F_{obs}$ | $F_{calc}$ | h | l | $F_{obs}$ | $F_{calc}$ |
|---|---|---|---|---|---|---|---|
| | 2 | 767 | -809 | 10 | 4 | 16 | -33 |
| | 4 | 811 | -779 | 2 | -2 | 978 | -1234 |
| | 6 | 472 | 491 | 2 | -4 | 275 | -286 |
| | 8 | 361 | -334 | 2 | -6 | 580 | 518 |
| | 10 | 339 | -332 | 2 | -8 | | 69 |
| | 12 | 312 | 335 | 2 | -10 | 96 | -147 |
| | 14 | 63 | -11 | 2 | -12 | 257 | 250 |
| | 16 | 22 | 25 | 2 | -14 | 158 | -113 |
| | 18 | 61 | -31 | 2 | -16 | 75 | -86 |
| | 20 | 18 | -16 | 2 | -18 | | 10 |
| 2 | | 480 | 421 | 2 | -20 | 28 | 40 |
| 2 | 2 | 362 | -371 | 4 | -2 | 414 | 384 |
| 2 | 4 | 674 | 763 | 4 | -4 | 402 | 414 |
| 2 | 6 | 125 | 129 | 4 | -6 | 379 | -313 |
| 2 | 8 | 201 | -198 | 4 | -8 | 21 | 97 |
| 2 | 10 | 82 | 122 | 4 | -10 | 79 | 32 |
| 2 | 12 | 110 | -51 | 4 | -12 | 114 | -136 |
| 2 | 14 | 126 | -171 | 4 | -14 | 72 | -47 |
| 2 | 16 | 135 | 140 | 4 | -16 | 115 | 116 |
| 4 | | 324 | -364 | 4 | -18 | 24 | -11 |
| 4 | 2 | 562 | -557 | 6 | -2 | 68 | -7 |
| 4 | 4 | 465 | 410 | 6 | -4 | 91 | 89 |
| 4 | 6 | 87 | -134 | 6 | -6 | 248 | -280 |
| 4 | 8 | 144 | -113 | 6 | -8 | 43 | 30 |
| 4 | 10 | 191 | 211 | 6 | -10 | 144 | 123 |
| 4 | 12 | 60 | 40 | 6 | -12 | 95 | -60 |
| 4 | 14 | 114 | -120 | 6 | -14 | 21 | 23 |
| 6 | | 207 | 194 | 6 | -16 | 81 | 101 |
| 6 | 2 | 26 | 37 | 6 | -18 | 72 | -96 |
| 6 | 4 | 69 | -52 | 8 | -2 | 320 | -330 |
| 6 | 6 | 15 | -37 | 8 | -4 | 133 | 102 |
| 6 | 8 | 15 | -11 | 8 | -6 | 116 | 114 |
| 6 | 10 | 59 | -75 | 8 | -8 | 116 | -55 |
| 6 | 12 | 51 | 48 | 8 | -10 | 49 | 50 |
| 8 | | 139 | 98 | 8 | -12 | 23 | 28 |
| 8 | 2 | 88 | 75 | 8 | -14 | 39 | -79 |
| 8 | 4 | 113 | -82 | 10 | -2 | 101 | -64 |
| 8 | 6 | 42 | 57 | 10 | -4 | 89 | 101 |
| 8 | 8 | 49 | 66 | 10 | -6 | 35 | 34 |
| 8 | 10 | 31 | -65 | 10 | -8 | 121 | -89 |
| 10 | | 58 | 46 | 10 | -10 | 47 | -5 |
| 10 | 2 | 20 | -32 | | | | |

* Structure factors refer to the half molecule and have been multiplied by 100.

Observed and Calculated Structure Factors for N,N'-Dimethyl Malonamide*

0kl reflections

| k | l | $F_{obs}$ | $F_{calc}$ | k | l | $F_{obs}$ | $F_{calc}$ |
|---|---|---|---|---|---|---|---|
| | 2 | 723 | -813 | 2 | 14 | | 22 |
| | 4 | 779 | -791 | 2 | 15 | | -4 |
| | 6 | 488 | 498 | 2 | 16 | | -24 |
| | 8 | 376 | -349 | 2 | 17 | | 17 |
| | 10 | 329 | -361 | 2 | 18 | | 17 |
| | 12 | 306 | 374 | 2 | 19 | 23 | -25 |
| | 14 | 63 | -4 | 2 | 20 | 19 | -30 |
| | 16 | 21 | 39 | 4 | | 122 | -67 |
| | 18 | 64 | -57 | 4 | 1 | 13 | -41 |
| | 20 | 21 | -20 | 4 | 2 | 64 | 43 |
| 2 | | 460 | 468 | 4 | 3 | 70 | -78 |
| 2 | 1 | 216 | 155 | 4 | 4 | 66 | 36 |
| 2 | 2 | 64 | 4 | 4 | 5 | 48 | 30 |
| 2 | 3 | 72 | 14 | 4 | 6 | 94 | -52 |
| 2 | 4 | 269 | -222 | 4 | 7 | 55 | -43 |
| 2 | 5 | 454 | -407 | 4 | 8 | 18 | -15 |
| 2 | 6 | 160 | 155 | 4 | 9 | 47 | 49 |
| 2 | 7 | 182 | 169 | 4 | 10 | 20 | 11 |
| 2 | 8 | 231 | -219 | 4 | 11 | 13 | 27 |
| 2 | 9 | 87 | 89 | 4 | 12 | 15 | 18 |
| 2 | 10 | 22 | -17 | 4 | 13 | 16 | -1 |
| 2 | 11 | 87 | 89 | 4 | 14 | | -3 |
| 2 | 12 | 89 | 120 | 4 | 15 | 21 | -22 |
| 2 | 13 | 53 | -70 | 4 | 16 | 15 | -13 |

1kl reflections

| k | l | $F_{obs}$ | $F_{calc}$ | k | l | $F_{obs}$ | $F_{calc}$ |
|---|---|---|---|---|---|---|---|
| 1 | | 929 | 1152 | 1 | 11 | 79 | 50 |
| 1 | 1 | 458 | 433 | 1 | 12 | | 26 |
| 1 | 2 | 150 | 123 | 1 | 13 | | 12 |
| 1 | 3 | 150 | 132 | 1 | 14 | 20 | -22 |
| 1 | 4 | 116 | -82 | 1 | 15 | 21 | -53 |
| 1 | 5 | 203 | -209 | 1 | 16 | 84 | 101 |
| 1 | 6 | 489 | 514 | 1 | 17 | 37 | 65 |
| 1 | 7 | 352 | -296 | 1 | 18 | 16 | 11 |
| 1 | 8 | 440 | -457 | 1 | 19 | 21 | -35 |
| 1 | 9 | 183 | 185 | 1 | 20 | 26 | -52 |
| 1 | 10 | 53 | -40 | 1 | 21 | 9 | -14 |

Observed and Calculated Structure Factors for N,N'-Dimethyl Malonamide*

1kl reflections (cont'd)

| k | l | $F_{obs}$ | $F_{calc}$ | k | l | $F_{obs}$ | $F_{calc}$ |
|---|---|---|---|---|---|---|---|
| 3 | | | 48 | 1 | -12 | 250 | 266 |
| 3 | 1 | 14 | 17 | 1 | -13 | 29 | -45 |
| 3 | 2 | 125 | -132 | 1 | -14 | | 7 |
| 3 | 3 | 173 | 113 | 1 | -15 | 31 | 46 |
| 3 | 4 | 244 | 196 | 1 | -16 | 79 | -111 |
| 3 | 5 | 124 | -122 | 1 | -17 | | -33 |
| 3 | 6 | 65 | -8 | 1 | -18 | 33 | -7 |
| 3 | 7 | 78 | -100 | 1 | -19 | 9 | 28 |
| 3 | 8 | 78 | -95 | 1 | -20 | | -1 |
| 3 | 9 | 18 | 37 | 1 | -21 | 11 | -16 |
| 3 | 10 | 37 | -53 | 1 | -22 | 11 | 10 |
| 3 | 11 | 74 | 78 | 3 | -1 | 81 | -36 |
| 3 | 12 | 76 | 77 | 3 | -2 | 116 | 87 |
| 3 | 13 | 15 | -14 | 3 | -3 | 143 | 145 |
| 3 | 14 | 17 | -23 | 3 | -4 | 80 | -98 |
| 3 | 15 | 16 | -1 | 3 | -5 | 53 | -19 |
| 3 | 16 | | 3 | 3 | -6 | 172 | -121 |
| 3 | 17 | 12 | 13 | 3 | -7 | 171 | -161 |
| 5 | | 81 | -36 | 3 | -8 | 131 | 121 |
| 5 | 1 | 43 | -29 | 3 | -9 | 19 | 16 |
| 5 | 2 | 31 | 1 | 3 | -10 | 15 | -24 |
| 5 | 3 | | -7 | 3 | -11 | | -5 |
| 5 | 4 | | 11 | 3 | -12 | 20 | 48 |
| 5 | 5 | 11 | -4 | 3 | -13 | 24 | 30 |
| 5 | 6 | 8 | 6 | 3 | -14 | 23 | -13 |
| 5 | 7 | 22 | -29 | 3 | -15 | 23 | 27 |
| 5 | 8 | 8 | -16 | 5 | -1 | 69 | 82 |
| 1 | -1 | 678 | -678 | 5 | -2 | 84 | 46 |
| 1 | -2 | 644 | -606 | 5 | -3 | | 5 |
| 1 | -3 | 627 | 674 | 5 | -4 | 19 | -16 |
| 1 | -4 | 885 | -1056 | 5 | -5 | 43 | -45 |
| 1 | -5 | 20 | -23 | 5 | -6 | 25 | -13 |
| 1 | -6 | 627 | 653 | 5 | -7 | 27 | 9 |
| 1 | -7 | 338 | -342 | 5 | -8 | 14 | 10 |
| 1 | -8 | 244 | -207 | 5 | -9 | 11 | -7 |
| 1 | -9 | 75 | -70 | 5 | -10 | 12 | 6 |
| 1 | -10 | 58 | -89 | 5 | -11 | 9 | -24 |
| 1 | -11 | 130 | 126 | 5 | -12 | 15 | -11 |

## Appendix II

## Calculation of Scale, Positional, and Temperature Factor Corrections from Difference Maps

The method used to obtain corrections from difference maps starts with the supposition* that the projected difference density in the neighborhood of an atomic center may be expressed as

$$D(x_1,x_2) = (1+\Delta K)\int f(H)\exp\left[\tfrac{-1}{4}B_0H^2\right]\exp\left[\tfrac{-1}{4}(\Delta B_1H_1^2+\Delta B_2H_2^2)\right]\cdot \cos 2\pi\left[(x_1-\Delta x_1)H_1+(x_2-\Delta x_2)H_2\right]\,d\Omega_H$$

$$-\int f(H)\exp\left[\tfrac{-1}{4}B_0H^2\right]\cos 2\pi(H_1x_1+H_2x_2)\,d\Omega_H$$

where $\underline{x}_1$ and $\underline{x}_2$ are measured in Å from the assumed atomic center

$B_o$ is the uncorrected isotropic temperature factor in $Å^2$

k, $\underline{x}_1$, $\underline{x}_2$, $B_1$, $B_2$ are the scale, positional, and thermal corrections to be added to the uncorrected values

f(h) is the atomic form factor

$H_1$ and $H_2$ are the components along $\underline{x}_1$ and $\underline{x}_2$ of the continuous reciprocal vector $\underline{H}$

---

* This representation of D is correct if
1) The contribution of overlapping atoms is negligible.
2) The atomic scattering factor in three-dimensional reciprocal space is of the form $f(\underline{H})\exp\left[\tfrac{-1}{4}\sum_{i,j}H_iH_jB_{ij}\right]$ if the projection is a normal one or of the more restrictive form $f(\underline{H})\exp\left[\tfrac{-1}{4}(H_1^2B_1+H_2^2B_2+H_3^2B_3)\right]$ in case $H_3$ is constant but non-zero (i.e. for projections calculated from non-zero layer data).

$A_H$ is area in reciprocal space

$$H^2 = 4\,\frac{\sin^2\theta}{\lambda^2}$$

The integrations cover that part of reciprocal space from which data have been drawn for the computation of the projection.

Then, by straightforward elimination of terms of order higher than the first in k, $\underline{x}_1$, $\underline{x}_2$, $B_1$, $B_2$ from the expressions for D, $\frac{\partial D}{\partial \underline{x}_1}$, $\frac{\partial D}{\partial \underline{x}_2}$, $\frac{\partial^2 D}{\partial \underline{x}_1^2}$, and $\frac{\partial^2 D}{\partial \underline{x}_2^2}$, there result the equations

$$\Delta k = \frac{4\pi^2 I_5 D(0) + I_3\left(\frac{\partial^2 D}{\partial x_1^2}\right)_0 + I_3\left(\frac{\partial^2 D}{\partial x_2^2}\right)_0}{8\pi^3 (I_1 I_5 - I_3^2)}$$

$$\Delta \underline{x}_1 = \left(\frac{\partial D}{\partial x_1}\right)_0 \Big/ 4\pi^3 I_3$$

$$\Delta \underline{x}_2 = \left(\frac{\partial D}{\partial x_1}\right)_0 \Big/ 4\pi^3 I_3$$

$$\Delta B_1 = \frac{4\pi^2 I_3 I_5 D(0) + (3 I_1 I_5 - 2 I_3^2)\left(\frac{\partial^2 D}{\partial x_1^2}\right)_0 + (2 I_3^2 - I_1 I_5)\left(\frac{\partial^2 D}{\partial x_2^2}\right)_0}{2\pi^3 I_5 (I_1 I_5 - I_3^2)}$$

$$\Delta B_2 = \frac{4\pi^2 I_3 I_5 D(0) + (2 I_3^2 - I_1 I_5)\left(\frac{\partial^2 D}{\partial x_1^2}\right)_0 + (3 I_1 I_5 - 2 I_3^2)\left(\frac{\partial^2 D}{\partial x_2^2}\right)_0}{2\pi^3 I_5 (I_1 I_5 - I_3^2)}$$

where $$I_n = \int f(H) \exp\left[\tfrac{-1}{4} B_0 H^2\right] H^n \, dH$$

The integrals $I_n$ are easily found by graphical integration using tabulated atomic scattering factors.

For projections based on zero-layer data out to a limit of H = 1.30 values of the $I_n$ for $B_o = 6$ are given in Table A. The atomic form factor values of McWeeny (11) have been used in their evaluation.

Table A

| | Carbon | Nitrogen | Oxygen |
|---|---|---|---|
| $I_1$ | 0.836 | 1.044 | 1.304 |
| $I_2$ | 0.317 | 0.382 | 0.462 |
| $I_3$ | 0.238 | 0.277 | 0.329 |

## Appendix III

### Note on the Modular Projection

The composite projection given by

$$\rho_n(yz) = \sqrt{\left(\int_0^1 \rho(xyz)\cos 2\pi nx\,dx\right)^2 + \left(\int_0^1 \rho(xyz)\sin 2\pi nx\,dx\right)^2}$$

has been used in connection with a direct structure determination and refinement (12). For the refinement of atomic parameters $\rho_n$ resembles a normal projection in that all levels of $\underline{x}$ are significantly weighted in one projection. In this "modular" projection an isolated centrosymmetric atom is projected without peak shift. When overlap is present, however, there occurs a rather complicated distortion which in general does result in peak shift. This distortion may be understood in terms of the difference between the square of the simple projection and the square of the modular projection. This difference is a convenient measure of the distortion introduced by the modular projection and is given by the expression

$$\rho_n^2(yz) - \left(\int_0^1 \rho(xyz)\,dx\right)^2 =$$

$$\int_0^1\int_0^1 \rho(xyz)\rho(x'yz)[\cos 2\pi n(x-x') - 1]\,dx\,dx'$$

From the right hand side it is seen that an isolated centrosymmetric atom will be projected without peak shift but that in case of overlap a complicated interaction occurs between the electron densities of the overlapping atoms with the general result that atomic peaks appear shifted from their correct positions.

References

1) M. A. Whiteley, J. Chem. Soc. Lon. 83, 33 (1903).

2) W. P. Mason, "Piezoelectric Crystals", D. Van Nostrand Co., New York, 1950, p. 53.

3) W. Cochran and H. B. Dyer, Acta. Cryst. 5, 635 (1952).

4) R. M. Curtis and R. A. Pasternak, Acta Cryst 8, 675 (1955).

5) E. W. Hughes, Acta Cryst. 7, 526 (1954).

6) D. C. Phillips, Acta Cryst. 7, 221 (1954).

7) D. W. Cruickshank, Acta Cryst. 2, 65 (1949).

8) R. B. Corey and L. Pauling, Proc. Roy. Soc. Lon., B141, 10 (1953).

9) L. Pauling, "Nature of the Chemical Bond", Cornell, Ithaca, 1948, p. 189.

10) L. Pauling, Ibid. p. 175.

11) R. McWeeny, Acta Cryst. 4, 513 (1951).

12) J. Fridrichsons and A. McL. Mathieson, Acta Cryst. 8, 761 (1955).

# PROPOSITIONS

1) A specific molecular mechanism is proposed for the powerful mutagenic action of 5-bromouracil on bacteriophage. Some critical experiments are suggested.

2) Hughes (1) has stated that in 'generalized projections' unless the projection axis is perpendicular to the other two axes the generalized projection of a spherically symmetrical atom does not show an extremum in a position corresponding to its true coordinates. It is proposed that there exists an important exception to the above restriction.

3) A simple expression is proposed for the estimation of errors in bond lengths calculated from projections of vibrating atoms.

4) Svedberg (2), in deriving an expression for the sedimentation of charged molecules, employs an argument involving the electrical potential which is unnecessarily complex and physically misleading. An alternative derivation is proposed.

5) Kraemer (3) is lead by an equation of the form zero equals zero to the false conclusion that for colloidal solutes in concentrated salt solutions, combination between the sedimenting solute and some other component of the solution may give rise to quite appreciable errors in molecular weight values determined by the method of sedimentation equilibrium.

6) The measurement of small concentrations of substitution defects in crystals is seriously hindered by non-substitutive contamination. A method is proposed by which this difficulty may be overcome in the special but biologically interesting case of racemic crystals of optical enantiamorphs.

7) Levinthal (4), on the basis of auto-radiography of DNA containing radioactive phosphorus, has concluded that DNA prepared from bacteriophage by osmotic shock consists of at least two very different molecular weight species. Another explanation is proposed for Levinthal's findings which does not require molecular weight heterogeneity of the DNA. A simple experiment is proposed which should distinguish between the two explanations.

8) The apparent monodispersity (5) of DNA from bacteriophage T4 suggests that the molecules of this material are in some sense also discrete in vivo. This speculation is made more plausible by the molecular weight value itself which corresponds to about 12 molecules of DNA per phage particle, a number consistent with the apparent symmetry of the virus. A method is proposed for directly determining the size of the heritable sub-units.

9) The arrangement of the DNA in bacteriophage particles may be investigated by measurements of the survival of orientated phage after irradiation with polarized ultraviolet light. A detailed experimental approach is offerred.

10) It is proposed that the technique of density gradient centrifugation of DNA lends itself to experiments able to test critically some present models of the mechanism of DNA replication.

11) At room temperature the scattering of X-rays by crystals of N,N'-dimethyl malonamide indicates the presence of disorder (6). At very low temperatures, a transition to another structure takes place and no evidence of disorder is found. A single crystal can be made to undergo this transition repeatedly. The curious distribution of the diffuse scattering suggests an explanation for this behavior. It is proposed that this system be investigated in detail and specific procedures are suggested.

References to Propositions

1) E. W. Hughes, Acta Cryst. 7 526 (1954).

2) T. Svedberg and K. O. Pedersen, "The Ultracentrifuge", Oxford, 1940, p. 12.

3) E. O. Kraemer, Ibid, p. 62.

4) C. L. Levinthal and C. A. Thomas, Jr., Biochim. et Biophys. Acta, 23 453 (1957).

5) This Thesis, p. 35.

6) This Thesis, p. 43.