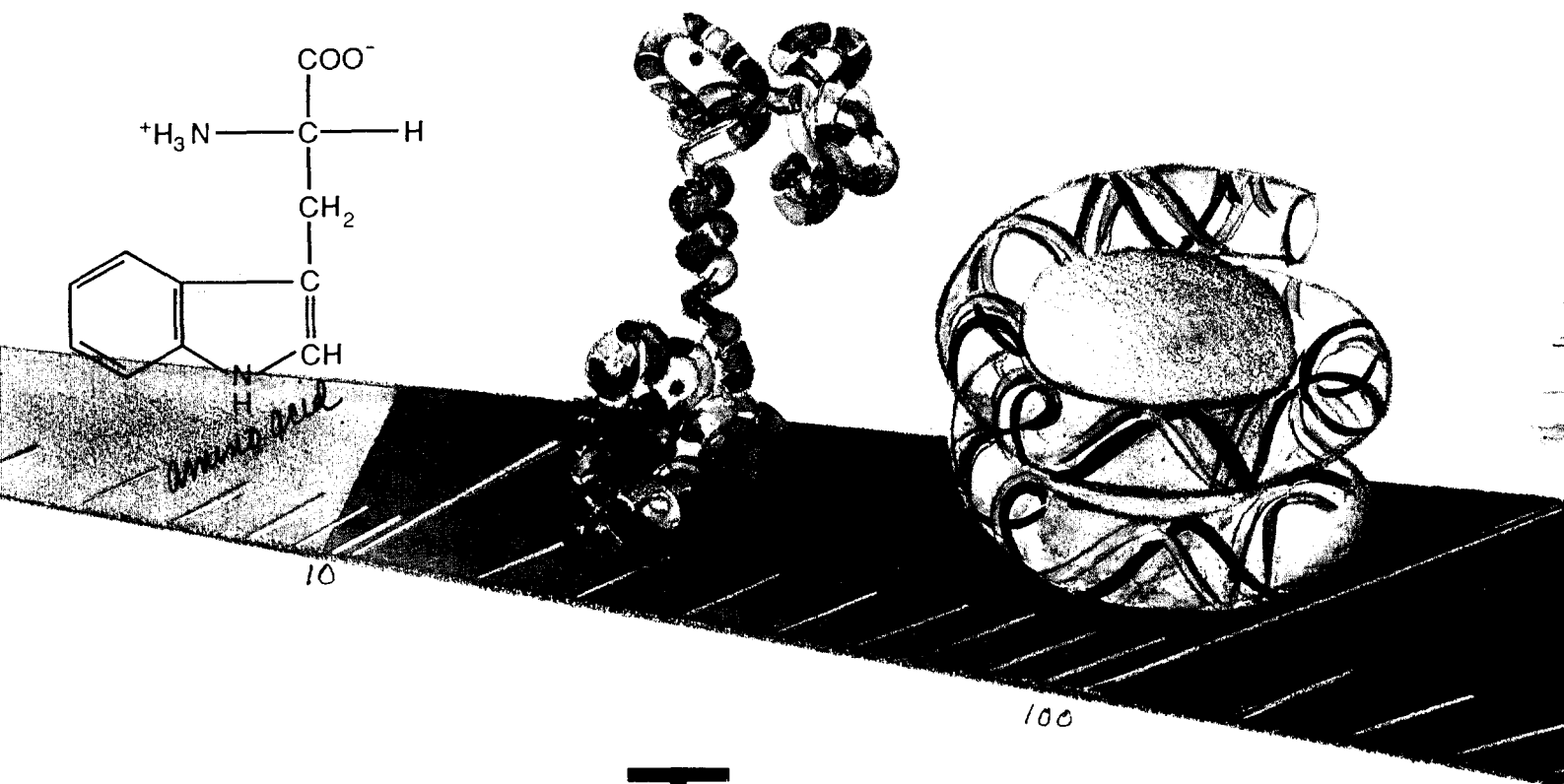


# BIOLOGY

## *on the scale of*



**T**

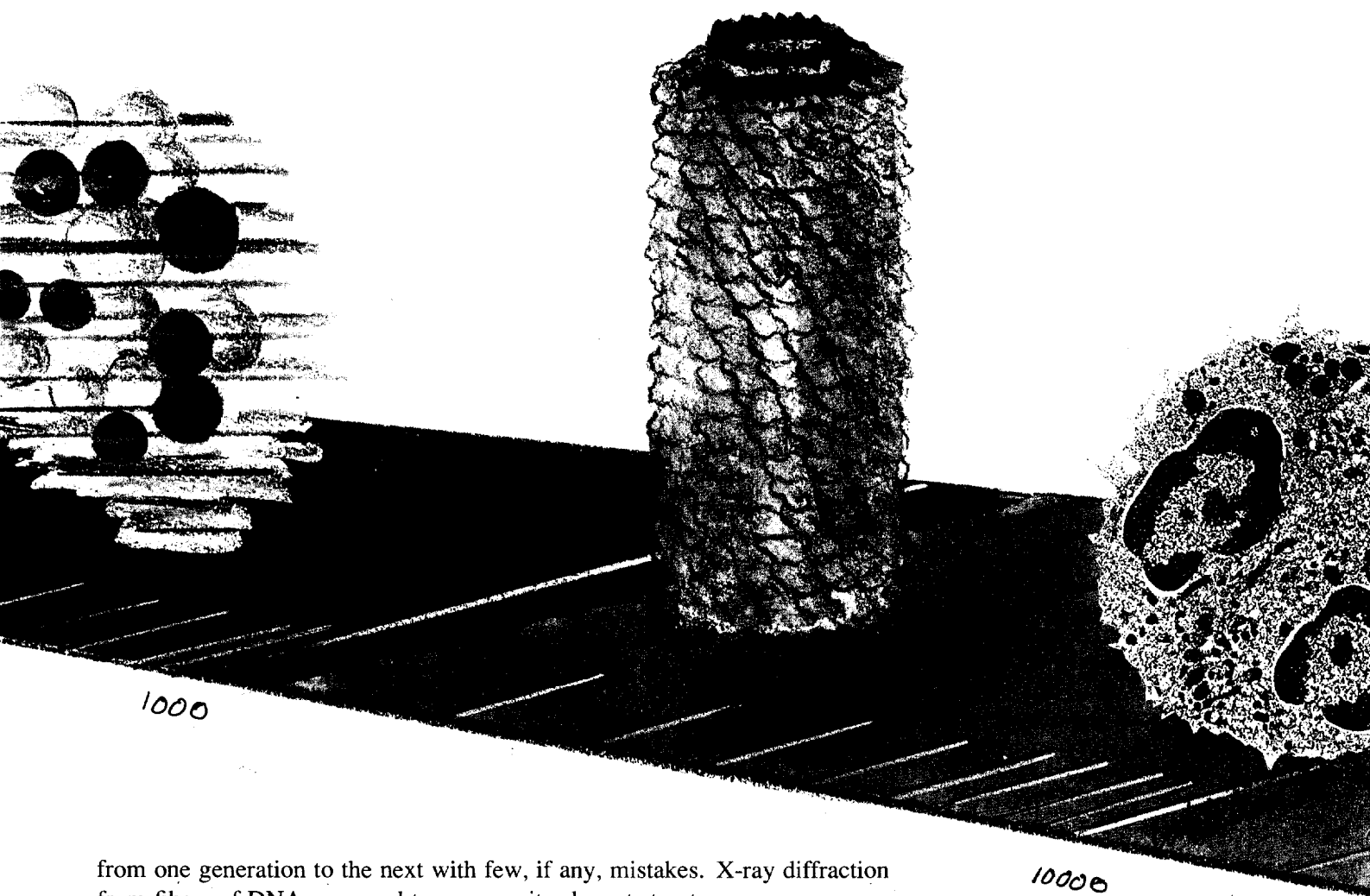
he ultimate challenge of the life scientist is to understand how the molecular constituents of living systems carry out their biological functions. What molecular interactions bring about and control such complicated functions as metabolism, reproduction, defense against invading organisms, and response to external stimuli?

Although none of those functions are fully understood, much of the progress so far has come from uncovering the *structures* of the biomolecules involved. In biology structure and function are intimately related. A most striking illustration is DNA, molecules of which contain all the genetic information for life. The discovery by Watson and Crick of the double-helix structure of DNA led immediately to a molecular picture of how the genetic instructions for making proteins are stored within an organism as well as inherited

The technique of small-angle neutron scattering illuminates length scales from tens to thousands of angstroms. It has been used to locate the different proteins (red and yellow) in the ribosome, to orient the DNA (red) and histone (blue) components of chromatin, and to determine the shape of calmodulin in solution.

# NEGLECTED DIMENSIONS

*by Jill Trehella*



from one generation to the next with few, if any, mistakes. X-ray diffraction from fibers of DNA was used to uncover its elegant structure.

Another major breakthrough in understanding structure-function relationships at the molecular level was the elucidation, through x-ray crystallography, of the structures of various proteins. Proteins are the working agents in all life processes. They catalyze biochemical reactions, they transport molecules and ions from one cellular compartment to another, they convert light to chemical

energy, they generate and transmit nerve impulses, they control cell differentiation and growth, they provide mechanical support, they even form tiny motors that propel microorganisms through their world much as an outboard motor propels a boat.

An organism manufactures a protein by joining together in a linear polypeptide chain a certain combination of some twenty different amino acids. The number of amino acids in the chain and their sequence, or order, along the chain are specified by the organism's DNA. It is the sequence of the amino acids that determines the function of a protein—by determining the way the polypeptide chain folds up into its functional structure, or conformation. The unique conformation of each protein species creates a unique chemical environment. which, in turn, may facilitate some specialized molecular interaction. Hence, each protein is able to carry out its function, or sometimes a set of functions, faithfully.

What tools are available for determining the structures of biological molecules and the functional units they compose? The task is difficult because the molecular weights of the molecules are quite large, 1000 to 1,000,000 daltons. (Such large molecules are often referred to as macromolecules. ) Nonetheless, techniques such as x-ray crystallography and nuclear-magnetic-resonance spectroscopy have given us an exquisitely detailed picture of the locations of individual chemical groups within some macromolecules and, in special cases, within macromolecular assemblies such as small viruses. Ideally, one would like to have that level of structural detail for all biological molecules in all functional states. However, after decades of effort, high-resolution structural data have been obtained for only a few hundred of the thousands of proteins whose amino-acid sequences are known, and those data often pertain to only one functional state.

At the level of the cell, light mi-

croscopy has provided us with a view of the functioning of large assemblies of biological molecules, But between the level of individual chemical groups and the cellular level is another realm, referred to by Wolfgang Oswald in the 1920s as “the world of neglected dimensions.” Populating that realm, which is highlighted in the opening figure, are the biological units responsible for controlling replication and transcription of genetic information. manufacturing proteins, transmitting neural impulses, converting light to neural impulses. contracting muscle fibers, and so on. Even today that structural realm remains relatively mysterious. The functional units are so large, so complex, and so difficult to crystallize that no single technique is adequate to tackle the job of determining their structures. Progress has been made only through the interplay of many different techniques. Among them is neutron scattering.

**The Contributions of Neutron Scattering**

Studies ranging from high-resolution neutron scattering from crystals to low-resolution neutron scattering from particles dispersed in a solvent have made important contributions to the structural database in biology. High-resolution crystallography has provided information, for example, about the locations of hydrogen atoms, and hence of conformation-stabilizing hydrogen bonds. in biological molecules and about the structure of water molecules in the immediate vicinity of the surfaces of biological molecules. The latter information is important because physical or chemical combination of water. the fluid of life, with biological molecules influences their conformations and often facilitates their functions.

At the other extreme, the structure of the basic subunit in the chromosome, namely the nucleosome (consisting of a

**Table 1**

**NEUTRON SCATTERING LENGTHS**

The neutron scattering length of an atom is a measure of the strength of the scattering interaction between neutrons and the atomic nucleus. Listed here are scattering lengths of atoms commonly found in biological molecules. (Deuterium,  $^2_1\text{H}$ , is included for reasons given in the text.) Also listed are values for the analogous parameter in x-ray scattering. Note that the scattering lengths vary unpredictably from one atom to another and that the range of variation is not great. In contrast, the x-ray scattering factors increase monotonically with atomic number.

Atom	Neutron Scattering Length (10 <sup>-12</sup> cm)	X-Ray Scattering Factor (electrons)
$^1_1\text{H}$	-0.38	1.0
$^2_1\text{H}$	0.67	1.0
$^{12}_6\text{C}$	0.66	6.0
$^{14}_7\text{N}$	0.94	7.0
$^{16}_8\text{O}$	0.58	8.0
$^{31}_{15}\text{P}$	0.51	15.0

length of DNA wound around a protein core), was first determined by scattering neutrons from nucleosomes in solution. Low-resolution neutron scattering is also being used to locate the many different protein species within the ribosome, the assembly responsible for translating RNA into proteins. That difficult work has been going on for fifteen years and is not yet completed.

Neutron scattering has also contributed structural information that will help explain how the transmembrane protein bacteriorhodopsin serves as a light-driven pump of protons. or hydrogen

ions. Bacteriorhodopsin is found in certain primitive bacteria and is responsible for their ability to use light to generate metabolic energy when oxygen is scarce. Absorption of light by bacteriorhodopsin creates an electrochemical gradient across the bacterial membrane, which serves as the source of metabolic energy. Neutron diffraction revealed in part how the protein chain folds back and forth across the membrane, presumably bringing together the chemical groups responsible for the proton translocations that create the gradient. but more work is needed to unravel the exact mechanism.

Our studies of the protein calmodulin, presented below, are an example of the use of neutron scattering to study, under physiological conditions, the conformational changes that occur as proteins carry out their functions.

In general, neutron scattering is useful for studying biological units ranging in size from small proteins to animal viruses. Moreover, it is a particularly powerful tool for studying composite entities containing more than one molecular species. That special power arises from a large difference between the scattering of neutrons by hydrogen (a major constituent of all biological molecules) and its neighboring isotope deuterium.

**The Neutron as a Probe of Biomolecules**

In “Neutron Scattering—A Primer,” hereafter referred to as the primer, we learn that unlike x rays, which scatter from the electron cloud surrounding an atomic nucleus, neutrons scatter from the atomic nucleus itself. If the wavelength of the neutron is much larger than the size of the nucleus (as it is in all low-energy neutron-scattering experiments), then the nucleus is considered to be located at a point *r*, and the interaction between the neutron and the

Table 2

**CONTRAST MATCHING**

When low-energy (long-wavelength) neutrons scatter at small angles from biological molecules in solution, the parameter that governs the observed intensity of scattered neutrons is the “contrast” between the biological molecules and the solvent, that is, the difference between their scattering-length densities. Therefore, by matching the scattering-length density of the solvent to the scattering-length density of one component of a two-component biological macromolecule, the contribution of that component to the observed neutron scattering can be eliminated. Listed here are the mean scattering-length densities of a nucleotide (DNA), a protein (histone), and lipids and the mean scattering-length densities of water (the usual solvent) containing various percentages of D<sub>2</sub>O.

Mean Scattering-Length Density (10 <sup>10</sup> cm <sup>-2</sup> )	
DNA	2.01
Histone	3.81
Lipids	between -0.2 and 0.8
0% D <sub>2</sub> O	-0.56
100% D <sub>2</sub> O	6.38
x% D <sub>2</sub> O	0.01[6.38x - 0.56(100 - x)]
37% D <sub>2</sub> O	2.01
63% D <sub>2</sub> O	3.81

nucleus can therefore be described by a delta-function pseudopotential, *bδ(r-r')*. The parameter *b*, the so-called scattering length, is related to the total scattering cross section of the nucleus, 4π*b*<sup>2</sup>. It turns out that the scattering lengths of all elements commonly found in biological macromolecules, except hydrogen, are positive and similar in magnitude. In contrast, the scattering length

of hydrogen is *negative* (because of a nuclear resonance) and equal in magnitude to about half the values of the scattering lengths of the other elements in biomolecules (Table I ). The negative scattering length of hydrogen means that the neutrons scattered from hydrogen are 180° out of phase with the neutrons scattered from elements with positive scattering lengths. Both the sign and the relatively large magnitude of hydrogen scattering length make that element far easier to “see” with neutrons than with x rays.

In this article we will focus on small-angle neutron scattering from macromolecules in solution. That technique is useful for determining overall shapes and sizes, molecular weights, and internal composition variations of the large molecules, rather than the locations of individual atoms. Consequently, a scattering-length density (the sum of the scattering lengths of the nuclei within a volume element *dV* divided by the volume element) is relevant rather than individual scattering lengths. We will denote the scattering-length density at a particular point *r* within a particle by *b<sub>s</sub>(r)* and the mean scattering-length density of the particle by  $\bar{b}$ . Because various types of biological molecules, such as proteins, polynucleotides (DNA and RNA), and lipids (the basic constituents of membranes), contain different proportions of hydrogen relative to other elements and because the mean scattering-length density varies rapidly with the proportion of hydrogen, the mean scattering-length density of those types of molecules differ significantly. Serendipitously, their mean scattering-length densities lie between those for pure H<sub>2</sub>O and pure D<sub>2</sub>O (Table 2). That fact is crucial for the powerful technique called contrast matching, a technique that allows us to see, for example, only the protein component in a biological assembly containing proteins and DNA or proteins and lipids.

## Contrast Matching

At the right is a whimsical example of contrast matching in which one figure blends into the background pattern and the other figure stands out. Another example of contrast matching, more analogous to the situation in neutron scattering, is the viewing of black and white kittens playing on black or white carpets. On a white carpet the black kittens stand out sharply, whereas the white kittens blend into the carpet. In other words, the white kittens are "contrast matched" to the background. The reverse effect is achieved by placing the kittens on a black carpet. The black kittens are then matched to the background, and only the white kittens are easily visible. In neutron scattering from macromolecules in solution, the solvent plays the part of the carpet, the molecules play the part of the kittens, and the contrast is determined by the *difference* between the mean scattering-length densities of the solvent and the molecules.

If the molecules and the solvent have the same scattering-length density, the waves scattered from any point have the same amplitude and phase. When integrated over the entire sample volume, the scattering at any angle (other than zero) cancels because arriving waves have all phases with equal probability. In other words, no small-angle scattering occurs. On the other hand, suppose the dispersed molecules and the solvent have different scattering-length densities,  $b_p$  and  $b_s$ . Each wave scattered from some point within a molecule may be thought of as a sum of two waves, one of amplitude  $b_s$  and one of amplitude  $b_p - b_s$ . The waves of amplitude  $b_s$  cancel with the waves of equal amplitude scattered from the solvent, whereas the waves of amplitude  $b_p - b_s$  interfere constructively or destructively with other waves scattered from other points within the molecule. As described below, con-



When the monster came, Lola, like the peppered moth and the arctic hare, remained motionless and undetected. Harold, of course, was immediately devoured.

structive interference of those waves occurs mostly at small angles about the direction of the incident neutrons. Thus the name small-angle scattering is applied to the phenomenon. The observed intensity of the scattered neutrons is equal to the square of the amplitude of the scattered waves and is thus proportional to  $(b_p - b_s)^2$ .

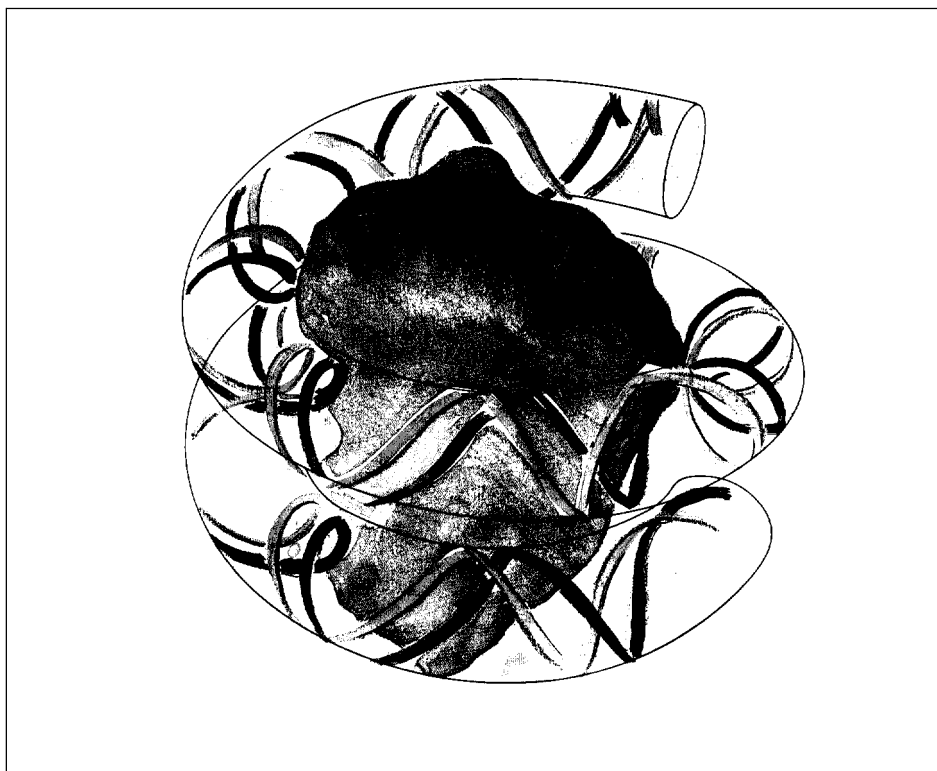
Now consider a particle composed of

two subunits, each with a different mean scattering-length density. By varying the  $D_2O$ -to- $H_2O$  ratio in the solvent, the scattering-length density of the solvent can be matched to that of one of the subunits. Then most of the observed scattering is due to the other subunit. Contrast matching was the technique used in the 1970s to elucidate the structure of the nucleosome, which contains

THE FAR SIDE COPYRIGHT 1988 UNIVERSAL PRESS SYNDICATE. Reprinted with permission. All rights reserved.

DNA and proteins. At that time it was not clear whether the protein component of the nucleosome was outside or inside the DNA component. The DNA component has the same scattering-length density as a solvent containing approximately 63 percent  $D_2O$ , whereas the protein component has the same scattering-length density as a solvent containing approximately 37 percent  $D_2O$ . Neutron-scattering data obtained at each of those match points, as well as at other solvent scattering-length densities, gave information on the relative orientation of the two components and led to the model shown in Fig. 1.

Many biochemical units contain protein species whose mean scattering-length densities are so similar that the technique of contrast matching is not directly applicable. The ribosome in the opening figure is an example. However, the range of variation among protein species can be increased by replacing some of them with “deuterated” versions. A deuterated version of a protein, as you might guess, contains deuterium rather than hydrogen. It can be prepared in various ways. For example, if the protein of interest is a bacterial protein, the bacterium is simply cultured on a medium containing  $D_2O$  rather than  $H_2O$ . As the bacterium multiplies, it produces deuterated versions of various proteins, including the protein of interest. The locations of the different protein species in a bacterial ribosome were determined by small-angle neutron-scattering studies on a set of ribosomes, each of which had been reconstituted with deuterated versions of two protein species prepared in the above manner. (Deuterated versions of nonbacterial proteins can be obtained by using recombinant DNA technology. Clones of the gene for the protein of interest are first combined with the DNA of a bacterium. The bacterium is then cultured as above, and the protein of interest harvested. )



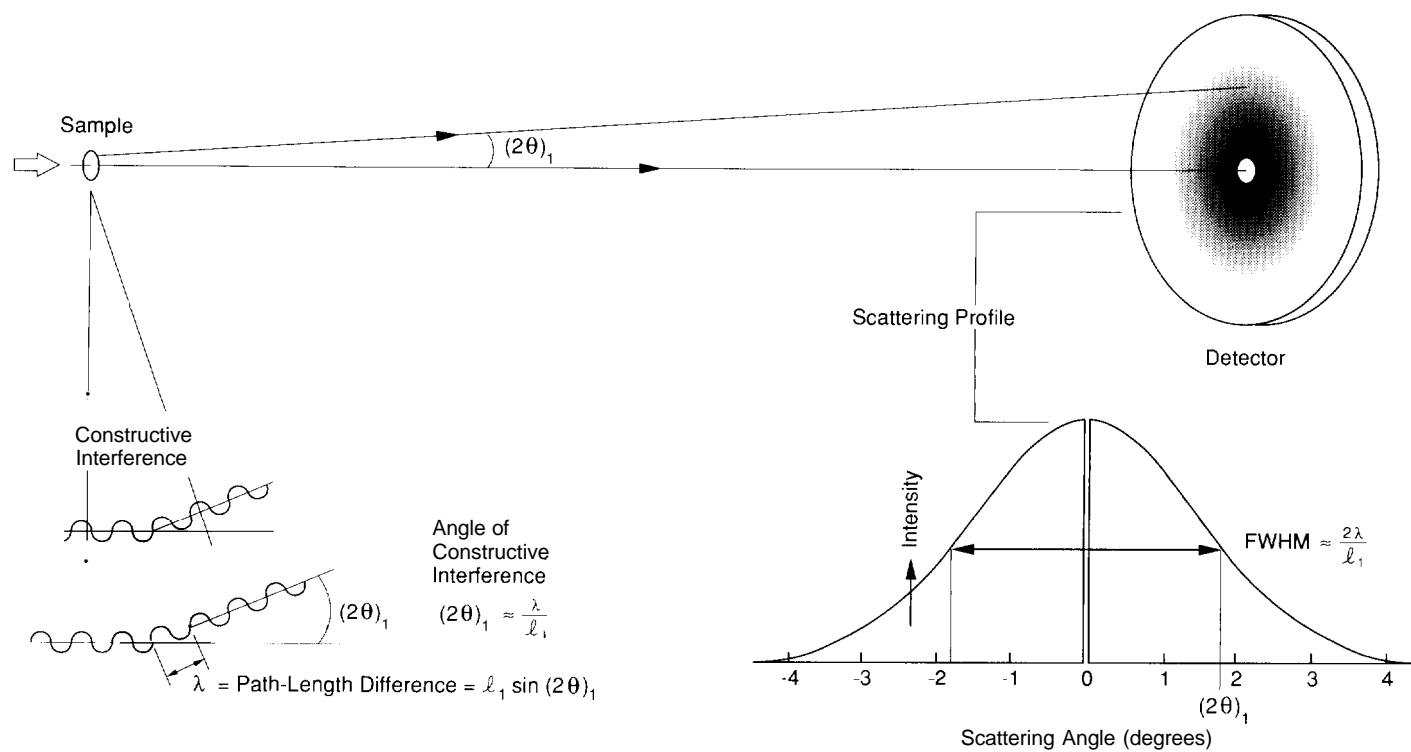
**THE NUCLEOSOME: A TWO-COMPONENT SYSTEM**

**Fig. 1.** The nucleosome, the basic subunit of chromosomes, contains DNA and proteins known as histones. That the DNA was wrapped around the proteins, rather than vice versa, was first deduced by using contrast matching to measure separately the small-angle neutron-scattering profiles of the DNA and the protein components.

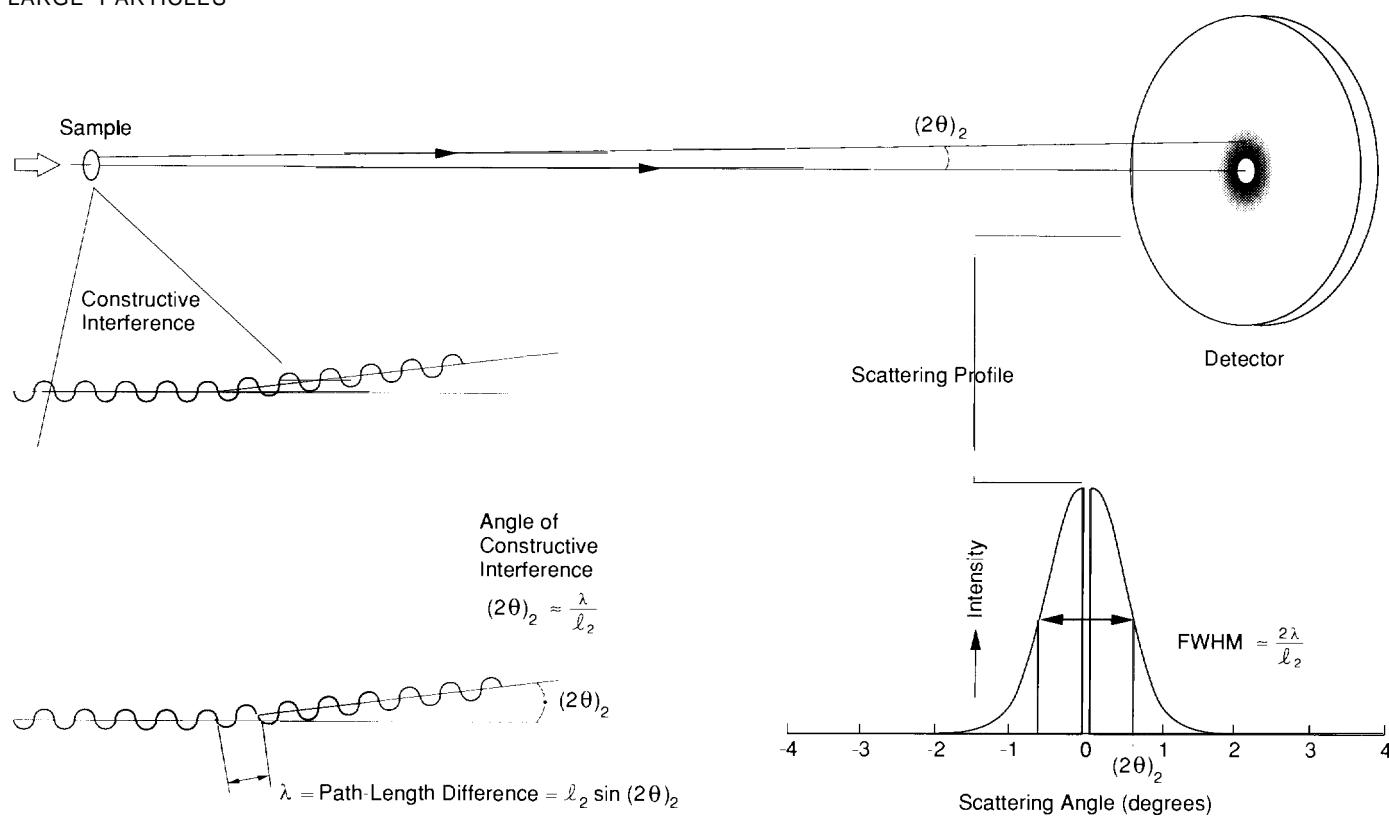
### Small-Angle Neutron Scattering

Figure 2 shows the experimental setup for small-angle neutron-scattering experiments. A sample of particles dispersed in a small volume (of the order of 100 cubic millimeters) of solution is placed in the neutron beam, and a detector located several meters away measures the intensity of the neutrons scattered through various angles  $2\theta$  relative to the direction of the incident beam. To protect the detector, a small beam stop (a few square millimeters) absorbs the unscattered neutrons—as well as those scattered through zero degrees and through angles very close to zero. The intensity of the scattered neutrons

(represented by shades of gray in Fig. 2) varies with the scattering angle and, if the dispersed particles are randomly oriented, is isotropic about the beam axis. Figure 2 shows a peak in intensity centered at zero degrees and extending over a small range of scattering angles. (Note that the beam stop prevents direct observation of the intensity at scattering angles equal and very nearly equal to zero. Instead  $I(0)$  is determined by extrapolation of the observed scattering profile.) Such a peak occurs if the wavelength of the incident neutrons,  $\lambda$ , is small compared with the particle size and large compared with internuclear distances. The peak is produced by constructive interference of waves scat-



# LARGE PARTICLES



## SMALL-ANGLE NEUTRON SCATTERING

**Fig. 2.** In a small-angle neutron-scattering experiment a detector located relatively far from the sample records the intensity  $I$  of the scattered neutrons as a function of scattering angle  $2\theta$ .  $I(0)$  cannot be measured directly because the neutrons scattered through zero degrees are coincident with the unscattered neutrons. Both are absorbed by a small beam stop coaxial with the incident beam. The observed scattering profile (graph of  $I$  versus  $2\theta$ ) at very small scattering angles contains information about the size and shape of the particles in the sample. For example, the width of the scattering profile varies inversely with the size of the scattering particles. In the case of spherical particles, the maximum separation between scattering points for which interference may occur is the particle diameter ( $\ell$ ). Integration over the particle volume yields a scattering profile that is approximately Gaussian (at small angles) with a full width at half maximum (FWHM) of  $(1.14) (2\lambda/\ell)$ . The figure shows the condition for constructive interference between the scattered waves from a pair of nuclei separated by a distance  $\ell$ : the difference between their path lengths,  $\ell \sin 2\theta$ , must equal  $\lambda$ , the wavelength of the incident neutrons. For small angles  $\sin 2\theta \approx 2\theta$ , so the condition for constructive interference becomes  $2\theta \approx \lambda/\ell$ .

tered from pairs of nuclei within each particle, which occurs when the difference in path length of the two scattered waves is equal and nearly equal to an integral multiple of  $\lambda$ . Now the waves scattered at most angles from the many, many pairs of nuclei within the particle will tend to cancel out and produce a very small signal. Only the waves scattered at small angles (for which the path-length differences are also small)

will tend to add up and produce a large peak. The shape of the peak contains information about the size of the particles. For example, as shown in Fig. 2, the width of the peak (at half maximum) varies approximately inversely with the diameter of the particles.

To explain how we extract more exact quantitative information from the small-angle scattering profile, we build on the discussion in the primer. Consider the scattering in vacuum of neutrons from a particle whose scattering-length density is described by  $b_p(\mathbf{r})$ . Equation 4 in the primer states that the scattering profile of the neutrons is given by

$$I(Q) = \left| \int_{V_p} b_p(\mathbf{r}) e^{i\mathbf{Q} \cdot \mathbf{r}} d^3\mathbf{r} \right|^2, \quad (1)$$

where  $\mathbf{Q} = \mathbf{k}_{\text{final}} - \mathbf{k}_{\text{initial}}$  is the scattering vector,  $Q \equiv |\mathbf{Q}| = 2\pi \sin \theta / \lambda$ ,  $d^3\mathbf{r}$  is a volume element, and  $V_p$  is the volume of the particle. Suppose now that the neutrons are scattered from a particle in a matrix of uniform scattering-length density  $b_s$ . Then Eq. 1 becomes

$$I(Q) = \left| \int_{V_p} (b_p(\mathbf{r}) - b_s) e^{i\mathbf{Q} \cdot \mathbf{r}} d^3\mathbf{r} \right|^2. \quad (2)$$

(The quantity  $b_p(\mathbf{r}) - b_s$  in Eq. 2 is called the contrast factor.) If, as is generally the case, the neutrons are being scattered by more than one particle, the integral in Eq. 2 must be multiplied by the number of particles  $N_p$  (provided the particles are identical). For simplicity, that multiplicative constant is not explicitly given in the equations that follow.

By using the fact that, for any complex number  $z$ ,  $|z|^2 = zz^*$  (where  $z^*$  is the complex conjugate of  $z$ ), Eq. 2 becomes

$$I(Q) = \int_{V_p} \int_{V_p} (b_p(\mathbf{r}) - b_s)(b_p(\mathbf{r}') - b_s) \times e^{i\mathbf{Q} \cdot (\mathbf{r} - \mathbf{r}')} d^3\mathbf{r} d^3\mathbf{r}'. \quad (3)$$

The scattering of neutrons from biological molecules in solution is generally spherically averaged because the molecules are generally randomly oriented. The spherical average of Eq. 3 can be expressed in terms of a so-called pair-distribution function  $P(R)$ , where  $R \equiv |\mathbf{r} - \mathbf{r}'|$ :

$$I(Q) = 4\pi \int_0^{d_{\max}} P(R) \frac{\sin QR}{QR} dR, \quad (4)$$

where  $d_{\max}$  is the maximum dimension of the particles. (The transition from Eq. 3 to Eq. 4 is more easily comprehended by recalling that the spherical average of  $e^{i\mathbf{x} \cdot \mathbf{x}}$  is  $\sin x/x$ .) The pair-distribution function  $P(R)$  in Eq. 4 is a weighted frequency distribution of the distances  $|\mathbf{r} - \mathbf{r}'|$  between all pairs of scattering centers (that is, nuclei); each frequency is weighted by the factor  $(b_p(\mathbf{r}) - b_s)(b_p(\mathbf{r}') - b_s)$ .

Implicit in the derivation of Eq. 4 are several important assumptions: (1) all the particles in the solution are identical; (2) the particles are randomly oriented; (3) the concentration of the particles is sufficiently low that neutrons scattered by different particles do not interfere with each other; and (4) thermodynamically, the solution is a two-component system consisting of particles dissolved in a homogeneous solvent.

Structural information about the particles is deduced from an analysis of  $P(R)$ , which is determined by measuring  $I(Q)$  and calculating the inverse Fourier transform of Eq. 4:

$$P(R) = \frac{1}{2\pi^2} \int_0^\infty I(Q) QR \sin(QR) dQ. \quad (5)$$

(Although Eq. 5 indicates that the integration should extend from zero to infinity,  $I(Q)$  can be measured only over a finite  $Q$  range. That experimental limitation can introduce truncation errors into the calculated  $P(R)$ . Also note that



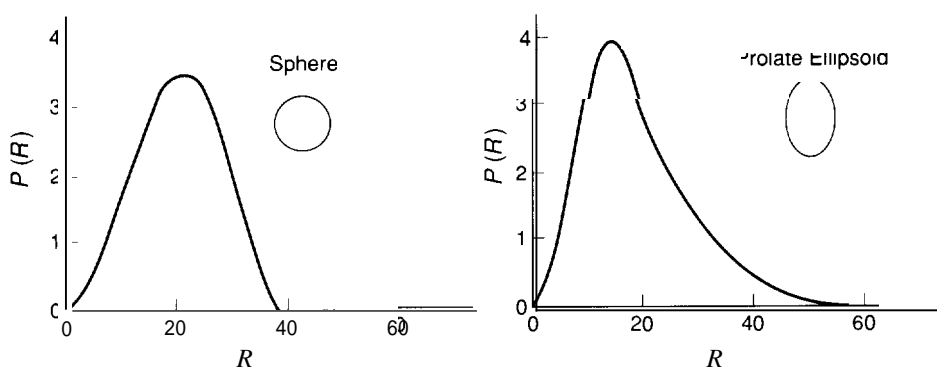
$I(Q)$  is measurable only if  $N_p$  is relatively large.) One piece of information obtained from  $P(R)$  is the overall shape of the scattering particle. For a spherical particle of uniform scattering-length density,  $P(R)$  is symmetric (Fig. 3a). It has a maximum near a value of  $R$  equal to the radius of the sphere and falls to zero at  $R = 0$  and at an  $R$  value equal to the sphere's diameter. As the particle becomes more asymmetric, so does  $P(R)$ . Figure 3b shows pair-distribution functions for several two-lobed particles with uniform scattering-length density. Note that those functions are quite sensitive to the relative dispositions of the two lobes. In all cases  $P(R) = 0$  for all  $R$  values greater than the maximum distance between scattering centers within the particle, and hence one can estimate the maximum dimension of the molecule,  $d_{max}$ .

Although  $P(R)$  is a complicated function of  $R$ , its integral over  $R$  is simply  $[(\bar{b}_p - b_s)V_p]^2$ , the square of the mean contrast factor times the particle volume. Moreover, according to Eq. 3 the intensity of the neutrons scattered through zero degrees,  $I(0)$ , is also equal to  $[(\bar{b}_p - b_s)V_p]^2$ :

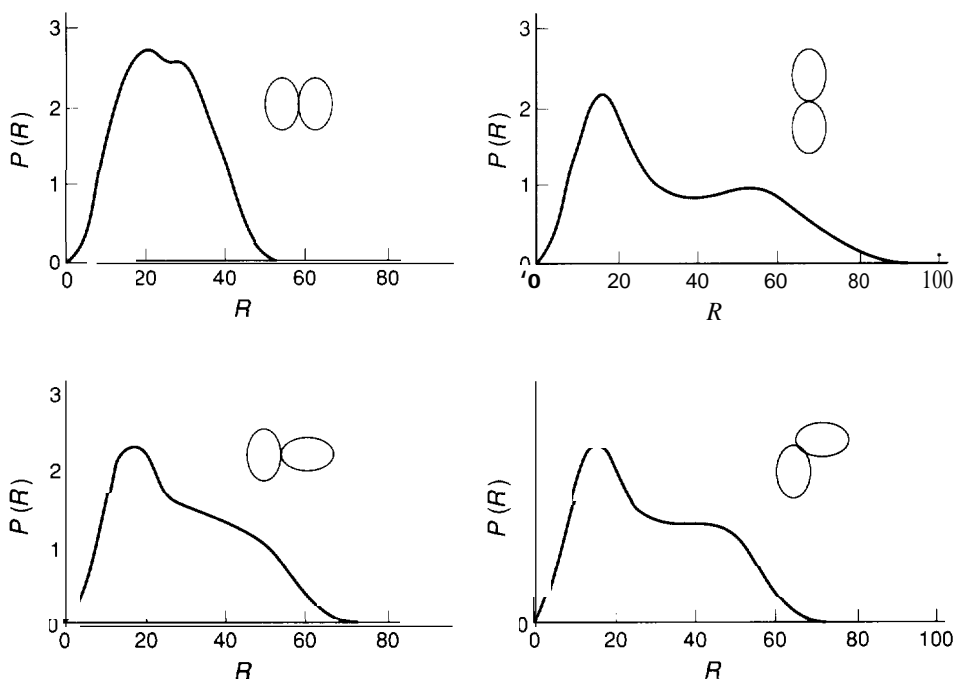
$$\begin{aligned}
 I(0) &= \int_{V_p} \int_{V_p} (b_p(\mathbf{r}) - b_s)(b_p(\mathbf{r}') - b_s) d^3\mathbf{r} d^3\mathbf{r}' \\
 &= [(\bar{b}_p - b_s)V_p]^2 \\
 &= \int_0^{R_{max}} P(R) dR.
 \end{aligned}
 \tag{6}$$

$I(0)$  is one of the most commonly cited small-angle neutron-scattering parameters. Since the neutrons scattered through zero degrees are coincident with the unscattered beam,  $I(0)$  cannot be measured directly. Instead,  $I(0)$  is determined by extrapolating the measured scattering profile to  $2\theta = 0$  (see below). By determining  $I(0)$  at different values of  $b_s$ , we can determine  $\bar{b}_p$  and  $V_p$ . In particular, according to Eq. 6 a plot of

(a) Single-Lobed Objects



(b) Two-Lobed Objects



#### PAIR-DISTRIBUTION FUNCTIONS FOR OBJECTS OF DIFFERENT SHAPE

Fig. 3. General structural features of an object are reflected in its pair-distribution function  $P(R)$ , which is determined from its small-angle neutron-scattering profile. If the object has a uniform scattering-length density,  $P(R)$  is proportional to the frequency distribution of the distances  $R$  between pairs of scattering centers (nuclei). Shown here are pair-distribution functions for objects of uniform scattering-length density and various shapes. (a) The pair-distribution functions for a sphere and a prolate ellipsoid show that the asymmetry of  $P(R)$  increases with the asymmetry of the object's shape. (b) The pair-distribution functions for various two-lobed objects demonstrate the sensitivity of  $P(R)$  to the relative orientation of the two lobes.

$\sqrt{I(0)}$  versus  $b_s$  is a straight line with a slope of  $-V_p$  and an intercept with the  $b_s$  axis of  $\bar{b}_p$ . Also, because the molecular weights of proteins are approximately proportional to their volumes,  $I(0)$  serves as a check that proteins in a sample are dispersed as single particles rather than as aggregates.

Another commonly cited parameter is  $R_g$ , the radius of gyration, which is defined by the equation

$$R_g^2 = \frac{\int_{V_p} (b_p(\mathbf{r}) - b_s)^2 d^3\mathbf{r}}{\int_{V_p} (b_p(\mathbf{r}) - b_s) d^3\mathbf{r}}. \quad (7)$$

The radius of gyration of neutron scattering is analogous to the radius of gyration of classical mechanics, which is defined by an equation exactly like Eq. 7 except that a mass density replaces the contrast factor. (The numerator and denominator in Eq. 7 are analogous, respectively, to the moment of inertia and the mass of classical mechanics.) The radius of gyration can also be determined as the second moment of the pair-distribution function:

$$R_g^2 = \frac{\int_0^{R_{\max}} P(R) R^2 dR}{\int_0^{R_{\max}} P(R) dR}. \quad (8)$$

Using  $P(R)$  to calculate  $R_g$  makes use of scattering data over the entire measurable  $Q$  range.

An alternative method for calculating  $R_g$  was developed by Guinier. He showed that the innermost portion of the scattering profile can be approximated well by a Gaussian curve:

$$I(Q) = I(0)e^{-\frac{1}{3}R_g^2Q^2}. \quad (9)$$

The value of  $R_gQ$  below which the Guinier approximation is valid depends on the particle shape; typically, however, it is valid for  $R_gQ < 1$ . Equation 9 implies that a plot of  $\ln I(Q)$  versus  $Q^2$  is a straight line with a slope of  $-\frac{1}{3}R_g^2$  and an extrapolated intercept

with the  $\ln I(Q)$  axis of  $\ln I(0)$ . Guinier analysis is a common method for determining  $I(0)$  as well as  $R_g$ . Later, when we present neutron-scattering data for a two-protein system at various contrast factors, we will see how an analysis of the variation of  $R_g$  with contrast factor allows us to deduce the relative positions of the two components, provided they have different mean scattering-length densities.

The theory for small-angle x-ray scattering is essentially the same as the theory presented here for small-angle neutron scattering except that x-ray scattering factors are substituted for neutron scattering lengths. X-ray scattering factors increase monotonically with the number of electrons in the electron cloud surrounding the atomic nucleus (that is, with the atomic number of the atom), and, because the dimensions of the electron cloud are of the same order as the wavelength of the x rays, they decrease with increasing scattering angle. Isotopes of the same element scatter x rays identically (because the number of electrons remains the same), and there are no negative x-ray scattering factors. Hence, although the scattering factor of a solvent can be changed by, for example, adding glucose or sucrose, the mean scattering factor of one component of a two-component biological assembly cannot be changed by isotopic substitution. However, x rays are easier and cheaper to produce than neutrons, and, when feasible, x-ray scattering is used to complement neutron scattering.

One final point. We mentioned above that neutron scattering from particles randomly oriented in solution yields only spherically averaged information about particle structure. One way to increase the information content of the scattering profile is to orient the particles along some direction in space. The sidebar "Ferrofluids—A New Alignment Technique" discusses a new method for orienting biological assemblies in so-

lution, which promises to yield more detailed structural information.

All the neutron-scattering experiments to be described below were performed at the Laboratory's pulsed neutron source (see "LANSCE-A Facility for Users") with the Low-Q Diffractometer (LQD), an instrument optimized for biological applications. Because the detector must be placed far from the sample to measure the neutron-scattering profile with high angular resolution and because biological molecules are inherently weak neutron scatterers, the flux of neutrons incident on the sample must be as large as possible to achieve a measurable intensity of scattered neutrons at the detector. The LQD is designed to maximize the incident flux and yet permit intensity measurements at very small scattering angles. The smallest angle at which useful intensity data can be obtained determines in part the largest particles that can be studied with the instrument.

The  $Q$  range measured with the LQD (0.002–0.5 angstrom<sup>-1</sup>) covers small-angle scattering from objects with dimensions from tens to thousands of angstroms. The wide  $Q$  range also allows use of the LQD for low-resolution diffraction measurements. (Diffraction from plant viruses is discussed in "Ferrofluids—A New Alignment Technique.") Moreover, unlike the case at steady-state neutron sources, all of those measurements can be made without reconfiguring the instrument.

## Calmodulin and Biochemical Regulation

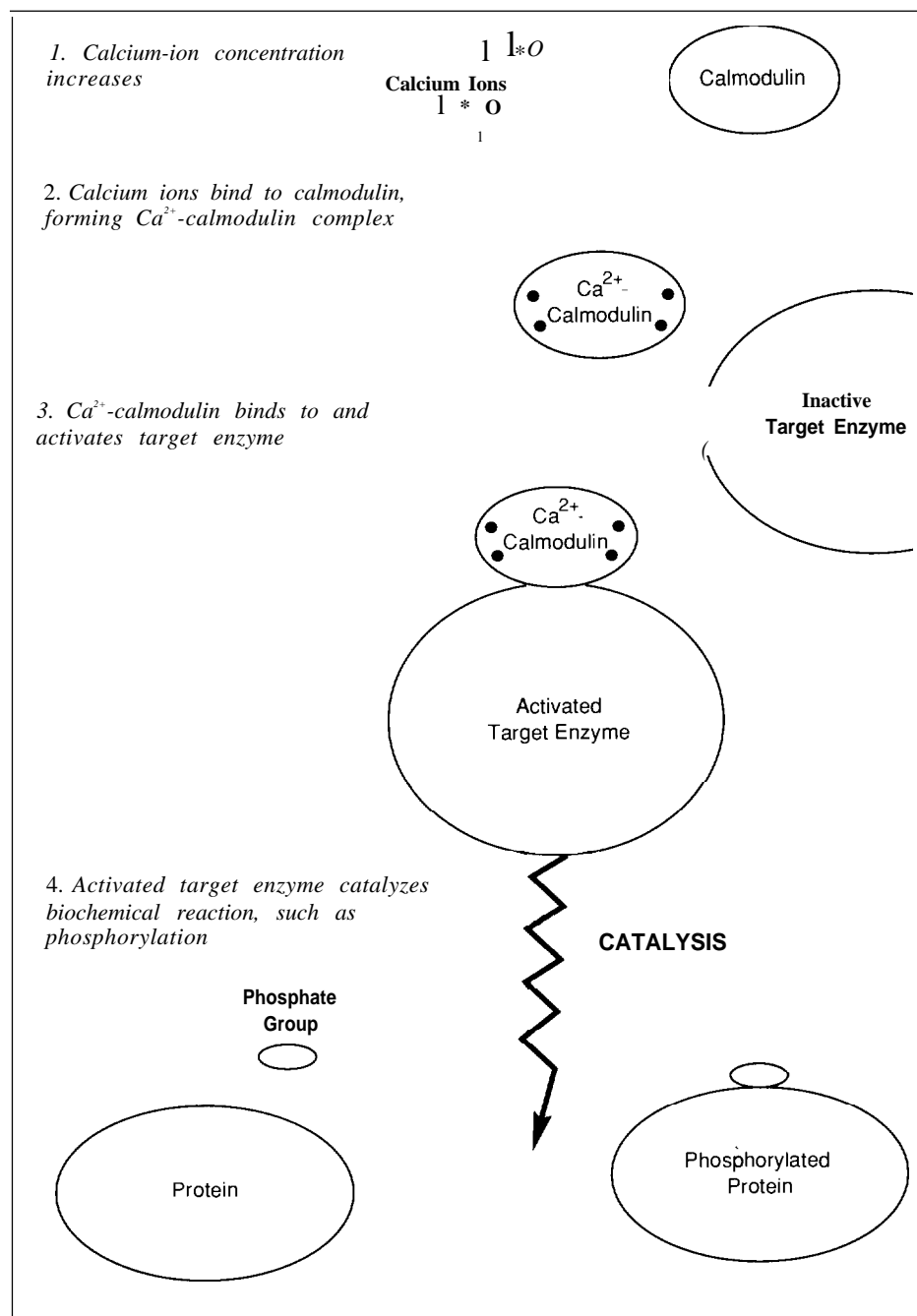
In the remainder of this article, we will focus on structural studies of the protein calmodulin, which mediates the regulatory effect of calcium ions on many biochemical processes. Through a combination of small-angle neutron and x-ray scattering, we have begun to study the conformational changes that

## CALCIUM REGULATION MEDIATED BY CALMODULIN

**Fig. 4.** Calcium ions,  $\text{Ca}^{2+}$ , regulate many intracellular processes, and many of the regulatory functions of  $\text{Ca}^{2+}$  are mediated by the protein calmodulin. That is, a local increase in  $\text{Ca}^{2+}$  concentration causes  $\text{Ca}^{2+}$  to bind to calmodulin, and the  $\text{Ca}^{2+}$ -calmodulin complex binds to a target enzyme and thereby activates a catalytic site on the enzyme. The activated target enzyme then catalyzes, or speeds up, a biochemical reaction, such as the addition of phosphate groups to some protein. Regulation by  $\text{Ca}^{2+}$  through calmodulin is involved in such processes as muscle contraction, transmission of nerve impulses, the degradation of glycogen to glucose, and cell division.

calmodulin undergoes as it carries out its function as mediator. But before we get ahead of the story, let's focus on the process of regulation.

Living systems must regulate a myriad of interdependent biochemical processes. In other words, processes must be turned on and off or slowed down and speeded up as needed. Regulation generally involves the interaction of small messenger molecules or ions with proteins. The messengers fall naturally into three major classes defined by the time scale of the processes they control: **hormones regulate intercellular** processes that occur on time scales of days to hours; small organic molecules known as cyclic nucleomonomophosphates (cyclic AMP, for example) regulate intracellular processes that occur on time scales of minutes to seconds; and the divalent (doubly charged) calcium ion,  $\text{Ca}^{2+}$ , regulates a large number of diverse intracellular processes that occur on time scales of seconds to milliseconds. Among the  $\text{Ca}^{2+}$ -regulated processes are mitosis (cell division), the degradation of glycogen (the storage



form of glucose) to glucose, muscle contraction, and transmission of nerve impulses. All those processes are regulated by localized changes in the concentration of calcium ions,

The question of how the simple calcium ion can regulate so many types of processes has attracted an enormous amount of attention. Many of the regulatory effects of  $\text{Ca}^{2+}$  are mediated by the protein calmodulin. Binding of  $\text{Ca}^{2+}$  to calmodulin is thought to induce some kind of conformational change that enables  $\text{Ca}^{2+}$ -calmodulin to bind to and thereby activate a "target" en-

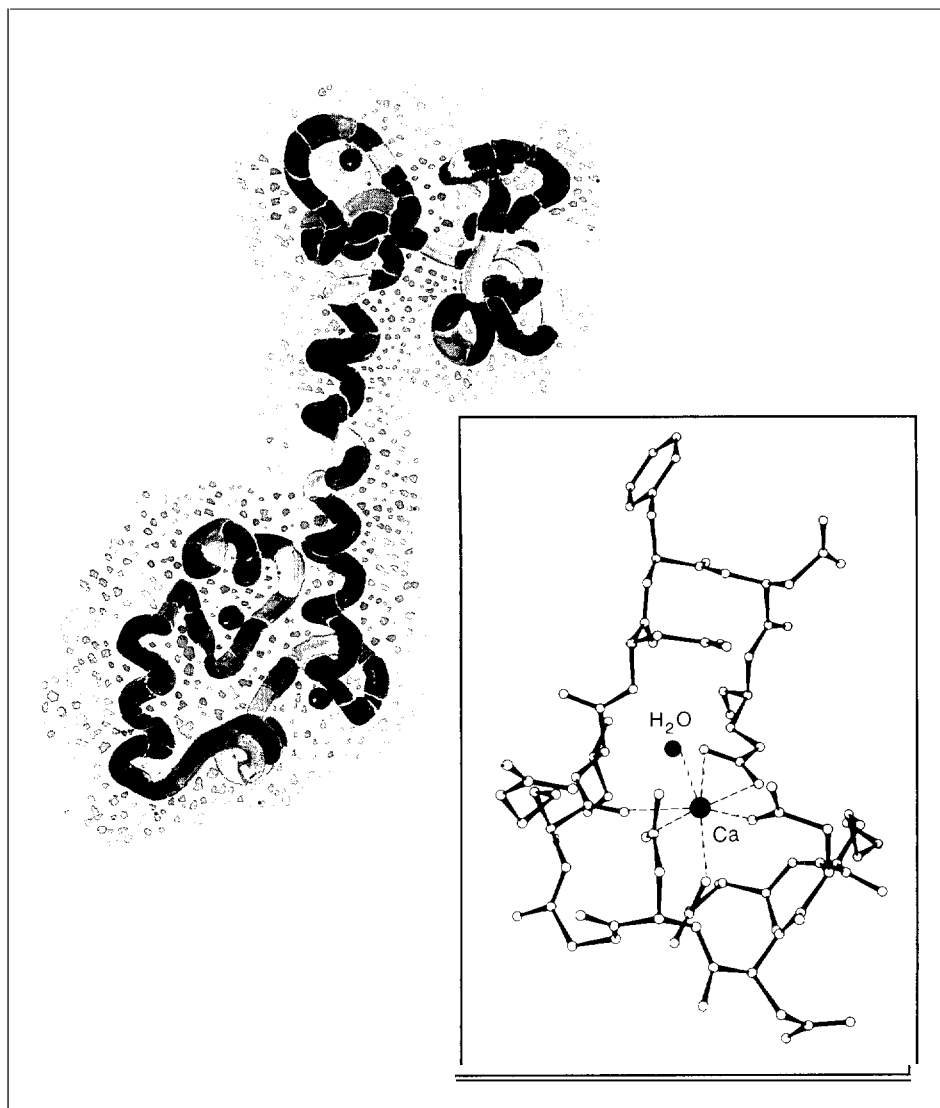
zyme. The activated target enzyme then catalyzes a biochemical reaction. The activated catalytic site on the target enzyme is sometimes far from the  $\text{Ca}^{2+}$ -calmodulin-binding site, and the mechanism of such "activation at a distance" is unknown (Fig. 4). We do know that  $\text{Ca}^{2+}$ -calmodulin can activate many different target enzymes (one at a time). Moreover, the constancy of the amino-acid sequence of calmodulin throughout evolution suggests that sequence changes almost anywhere in calmodulin are fatal to the organism and must therefore destroy key functions. Thus, the

sites for binding to different target enzymes may be distributed over the entire calmodulin molecule, and the activation mechanism may vary from one target enzyme to another.

What features of calmodulin's structure allow it to support such diverse activities? Calmodulin is a water-soluble protein, and most such proteins fold into globular structures that minimize the number of hydrophobic amino-acid residues exposed to water. But x-ray-diffraction data for  $\text{Ca}^{2+}$ -calmodulin (calmodulin without  $\text{Ca}^{2+}$  has not yet been crystallized) show that the polypeptide chain folds into two globular lobes connected by an extended  $\alpha$  helix of about eight turns (Fig. 5). Two  $\text{Ca}^{2+}$ -binding sites are found in each globular lobe, and all four  $\text{Ca}^{2+}$ -binding sites have a high degree of structural similarity with each other and with the  $\text{Ca}^{2+}$ -binding domains of other  $\text{Ca}^{2+}$ -activated regulatory proteins, such as parvalbumin and troponin C.

The  $\alpha$  helix connecting the two globular domains of calmodulin is presumed to be important in its interactions with target enzymes. Consequently, it has been the subject of considerable interest. Helical structures in more compact proteins are stabilized by intramolecular or sometimes intermolecular interactions. However, the crystal packing of  $\text{Ca}^{2+}$ -calmodulin shows no evidence of any such stabilizing interactions. So what happens to the  $\alpha$  helix under physiological conditions? Is it stable in the extended conformation of the crystal form, or is it flexible and able to twist or bend?

The two-lobed structure of  $\text{Ca}^{2+}$ -calmodulin is an ideal subject for study by both x-ray and neutron scattering because the scattering data are quite sensitive to the relationship between the two globular domains. Moreover, when coupled with the techniques of selective deuteration and contrast matching, neutron scattering allows us to see the



#### CRYSTAL STRUCTURE OF $\text{Ca}^{2+}$ -CALMODULIN

**Fig. 5.** The structure of crystalline  $\text{Ca}^{2+}$ -calmodulin was deduced from x-ray diffraction data for calmodulin crystallized at low pH and in the presence of  $\text{Ca}^{2+}$ . The structure is depicted here in a model of calmodulin's backbone superimposed on a space-filling model (left). The structure consists of two globular lobes connected by an extended  $\alpha$  helix of about eight turns. Located on each globular lobe are two  $\text{Ca}^{2+}$ -binding sites (red). The basic structure of all four  $\text{Ca}^{2+}$ -binding sites is the same (right).

shape of  $\text{Ca}^{2+}$ -calmodulin when it is bound to a target enzyme.

#### X-Ray Studies— The Solution Structure

My colleagues and I have gathered small-angle x-ray-scattering data that reveal differences between the crystal and solution structures of  $\text{Ca}^{2+}$ -calmodulin and between the solution structures of calmodulin and  $\text{Ca}^{2+}$ -calmodulin.

First we gathered small-angle x-ray-scattering data for calmodulin dispersed in a solution containing enough calcium

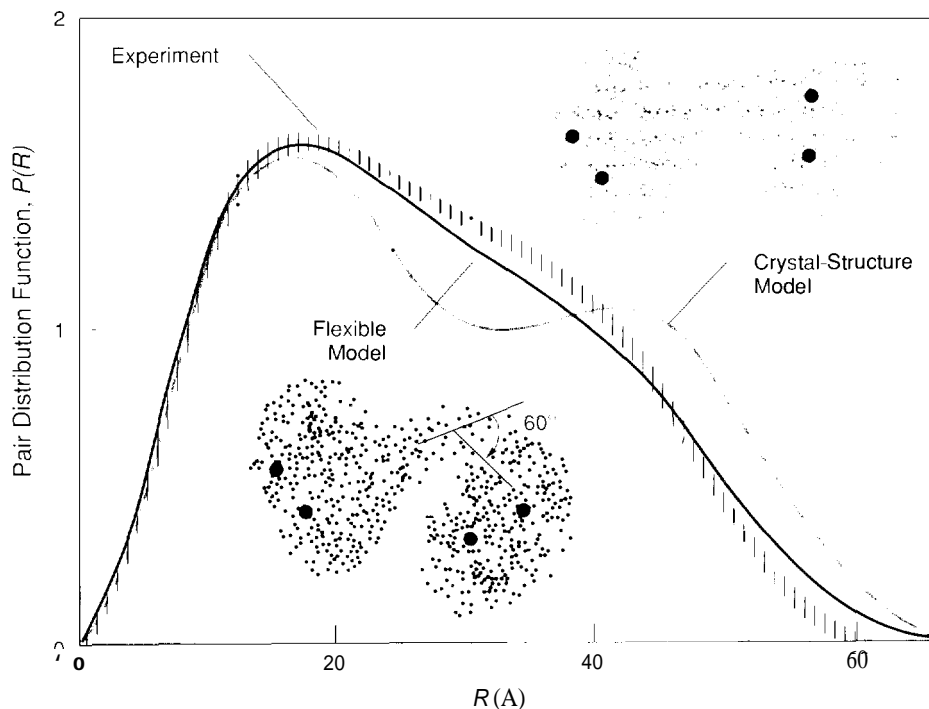
ions to saturate its  $\text{Ca}^{2+}$ -binding sites. Figure 6 compares the pair-distribution function derived from those data with two theoretical pair-distribution functions. Both theoretical pair-distribution functions were based on the same pattern of scattering centers, namely the pattern deduced from the reported crystal structure of  $\text{Ca}^{2+}$ -calmodulin. However, in one case the portion of the pattern representing the  $\alpha$  helix was allowed to be flexible so that on average the two globular lobes are closer together than in the crystal form. It should be pointed out that a uniform

density of scattering centers within calmodulin would yield similar theoretical pair-distribution functions. Thus calmodulin is fairly well approximated as a molecule of uniform scattering-length density.

As expected, all the pair-distribution functions in Fig. 6 bear some resemblance to the pair-distribution functions for double-lobed, uniform-density structures shown in Fig. 3b. Note that the experimental  $F'(R)$  falls to zero at a smaller  $R$  value than does the  $P(R)$  based on the (inflexible) crystal structure. Therefore the solution structure of  $\text{Ca}^{2+}$ -calmodulin has a smaller maximum dimension than does the crystal structure. More important, the experimental  $P(R)$  lies below the  $F'$  (1?) for crystalline  $\text{Ca}^{2+}$ -calmodulin at large  $R$  values and has no saddle at intermediate  $R$  values. In other words  $\text{Ca}^{2+}$ -calmodulin in solution has fewer widely separated scattering centers and more intermediately separated scattering centers than does crystalline  $\text{Ca}^{2+}$ -calmodulin. On the other hand, the  $P(R)$  based on a flexible  $\alpha$  helix agrees well with the experimental  $P(R)$ . Thus the x-ray-scattering studies indicate that in solution the interconnecting helix in  $\text{Ca}^{2+}$ -calmodulin is flexible and the globular lobes are, on average, about 10 angstroms closer together.

We repeated the x-ray studies, this time eliminating  $\text{Ca}^{2+}$  from the solution. Using both the pair-distribution function and Guinier analysis (Eqs. 8 and 9), we found that the radius of gyration of calmodulin is 5 percent smaller than that of  $\text{Ca}^{2+}$ -calmodulin. From complementary chemical studies we know that the increase in  $R_g$  due to binding of  $\text{Ca}^{2+}$  results in the exposure of hydrophobic regions in each globular lobe. Those nonpolar regions are thought to be important in the interactions of  $\text{Ca}^{2+}$ -calmodulin with target enzymes.

The issue of the flexibility of calmodulin's  $\alpha$  helix remains controversial,

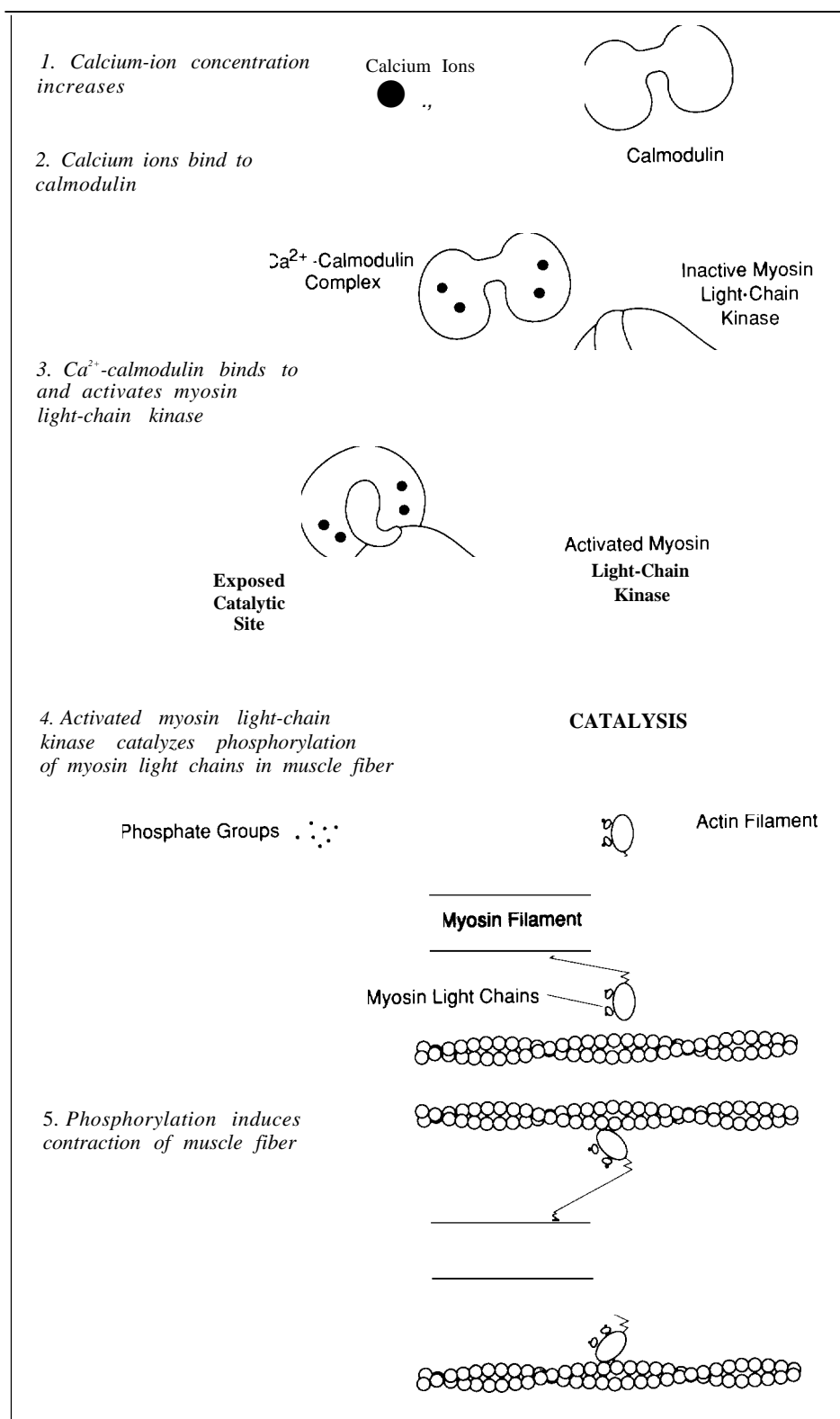


#### X-RAY RESULTS FOR $\text{Ca}^{2+}$ -CALMODULIN IN SOLUTION

Fig. 6. Shown here is the pair-distribution function calculated from small-angle x-ray-scattering data for calmodulin in a solution containing sufficient  $\text{Ca}^{2+}$  to saturate calmodulin's  $\text{Ca}^{2+}$ -binding sites. Also shown are two theoretical pair-distribution functions and the pattern of scattering centers on which each is based. The crystal-structure scattering-center pattern was deduced from the crystal structure of  $\text{Ca}^{2+}$ -calmodulin; the flexible-model scattering-center pattern is identical except that the portion corresponding to calmodulin's interconnecting helix is allowed to be flexible rather than rigid. The fact that the pair-distribution function based on the flexible model more closely matches the experimental pair-distribution function suggests that in solution calmodulin's interconnecting helix is flexible.

in part because more than one structure may be consistent with the pair-distribution function obtained from the spherically averaged scattering data. To pin down the solution structure more precisely, we have attempted a novel experiment in which  $^{240}\text{Pu}^{3+}$  is bound to the  $\text{Ca}^{2+}$ -binding sites in calmodulin. The triply charged ion of plutonium-240 has the same ionic radius as  $\text{Ca}^{2+}$ , and we have shown that it can substitute for  $\text{Ca}^{2+}$  in studies of biochemical processes dependent on  $\text{Ca}^{2+}$ -calmodulin. Plutonium-240 was chosen as a substitute for  $\text{Ca}^{2+}$  because 0.28-angstrom

neutrons form a short-lived bound state with plutonium-240 and thus have a much greater probability of scattering from plutonium-240 than from the other nuclei in calmodulin. That "resonant" scattering gave rise to a measurable interference pattern that directly reflects the distances between the  $\text{Ca}^{2+}$ -binding sites. With some further refinement the experiment will provide crucial information about the relationship between the two lobes. In addition, the small-angle scattering from  $\text{Pu}^{3+}$ -calmodulin showed that it has the same structure as  $\text{Ca}^{2+}$ -calmodulin.



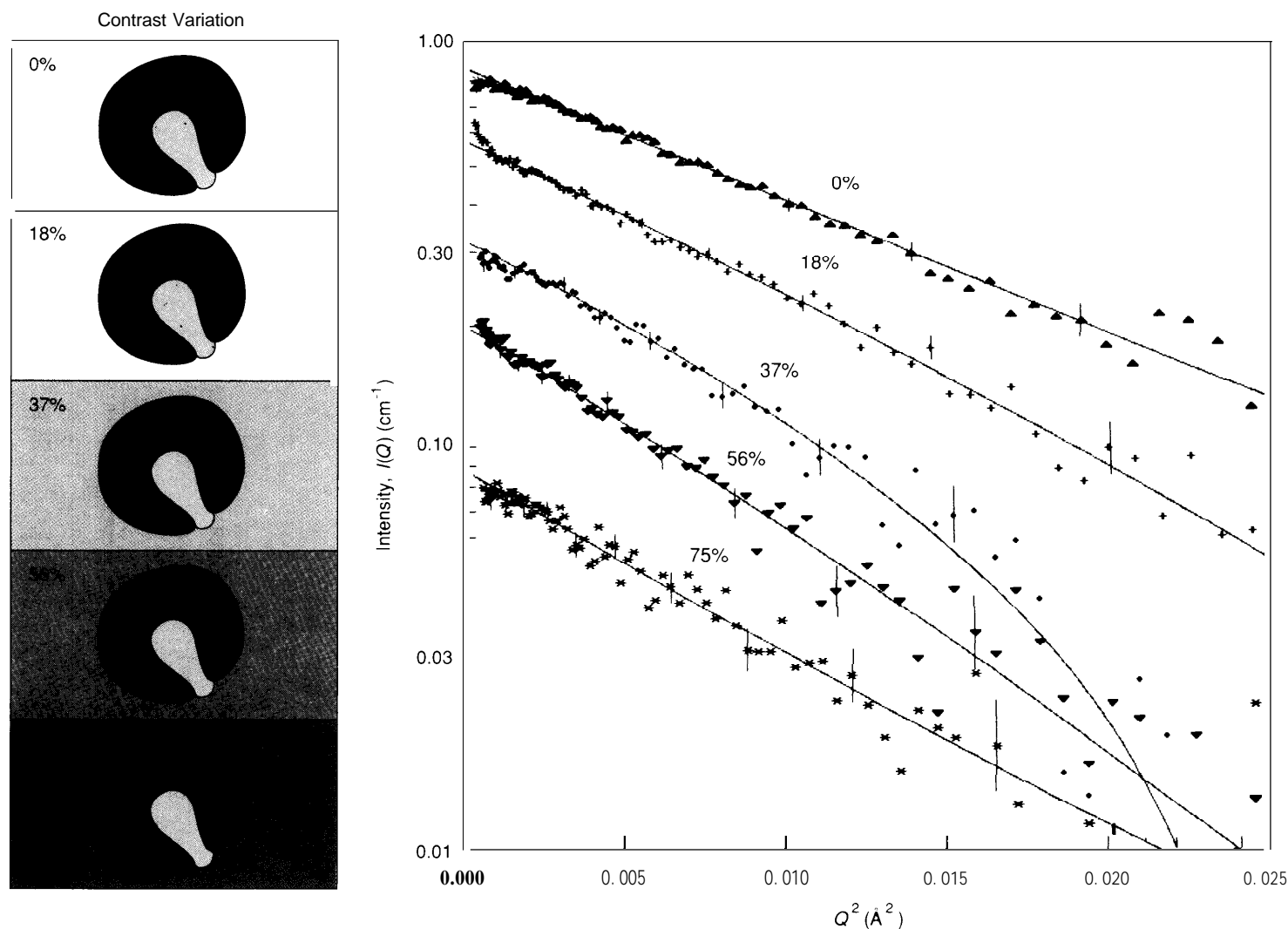
### REGULATION OF MUSCLE CONTRACTION BY $\text{Ca}^{2+}$ , CALMODULIN, AND MLCK

Fig. 7. Muscle contracts (and relaxes) as filaments of myosin and actin slide past each other. One molecular mechanism that enables the sliding motion involves calmodulin and the enzyme myosin light-chain kinase. A local increase in  $\text{Ca}^{2+}$  concentration causes binding of  $\text{Ca}^{2+}$  to calmodulin, and the resulting  $\text{Ca}^{2+}$ -calmodulin binds to MLCK and thereby activates a catalytic site on the enzyme. The activated MLCK then catalyzes the phosphorylation of a subunit of myosin, which enables the relative motion of myosin and actin.

### Neutron Studies—The Interaction with Target Enzymes

Although x-ray and chemical studies by us and others have yielded knowledge about the structural changes in calmodulin that result from binding of calcium ions, much less is known about the molecular interactions of  $\text{Ca}^{2+}$ -calmodulin with its various target enzymes. The difficulties associated with isolating large quantities of purified target protein and the very high molecular weights of most target enzymes are major obstacles to structural studies.

Myosin light-chain kinase (MLCK) from rabbit skeletal muscle is one of the best characterized of the enzymes activated by  $\text{Ca}^{2+}$ -calmodulin. When  $\text{Ca}^{2+}$ -calmodulin binds to MLCK, the enzyme catalyzes the binding of phosphate groups to a subunit of the muscle protein myosin. The "phosphorylation" of myosin is thought to induce a twitch response in muscle cells (Fig. 7). The domain of MLCK to which  $\text{Ca}^{2+}$ -calmodulin binds was the first such domain to be identified and sequenced. It consists of twenty-seven amino-acid residues located toward one end of the polypeptide chain. A synthetic polypeptide with an amino-acid sequence iden-



#### NEUTRON CONTRAST-VARIATION DATA FOR $\text{Ca}^{2+}$ -CALMODULIN-MLCK-I

**Fig. 8.** Small-angle neutron-scattering data for  $\text{Ca}^{2+}$ -calmodulin complexed with the synthetic peptide MLCK-I were obtained at different contrast factors by varying the percentage of  $\text{D}_2\text{O}$  in the solvent. Since the calmodulin was deuterated, the contrast factor was greatest when the solvent was pure  $\text{H}_2\text{O}$  and progressively decreased as the percentage of  $\text{D}_2\text{O}$  was increased. Consequently, the maximum intensity decreases as the  $\text{D}_2\text{O}$  content of the solvent increases. MLCK-I is contrast matched to a solvent containing 37 percent  $\text{D}_2\text{O}$ ; therefore the observed scattering at that  $\text{D}_2\text{O}$  content is dominated by scattering from calmodulin. The figures alongside the data represent by shades of gray the scattering contrast among solvent,  $\text{Ca}^{2+}$ -calmodulin, and MLCK-I.

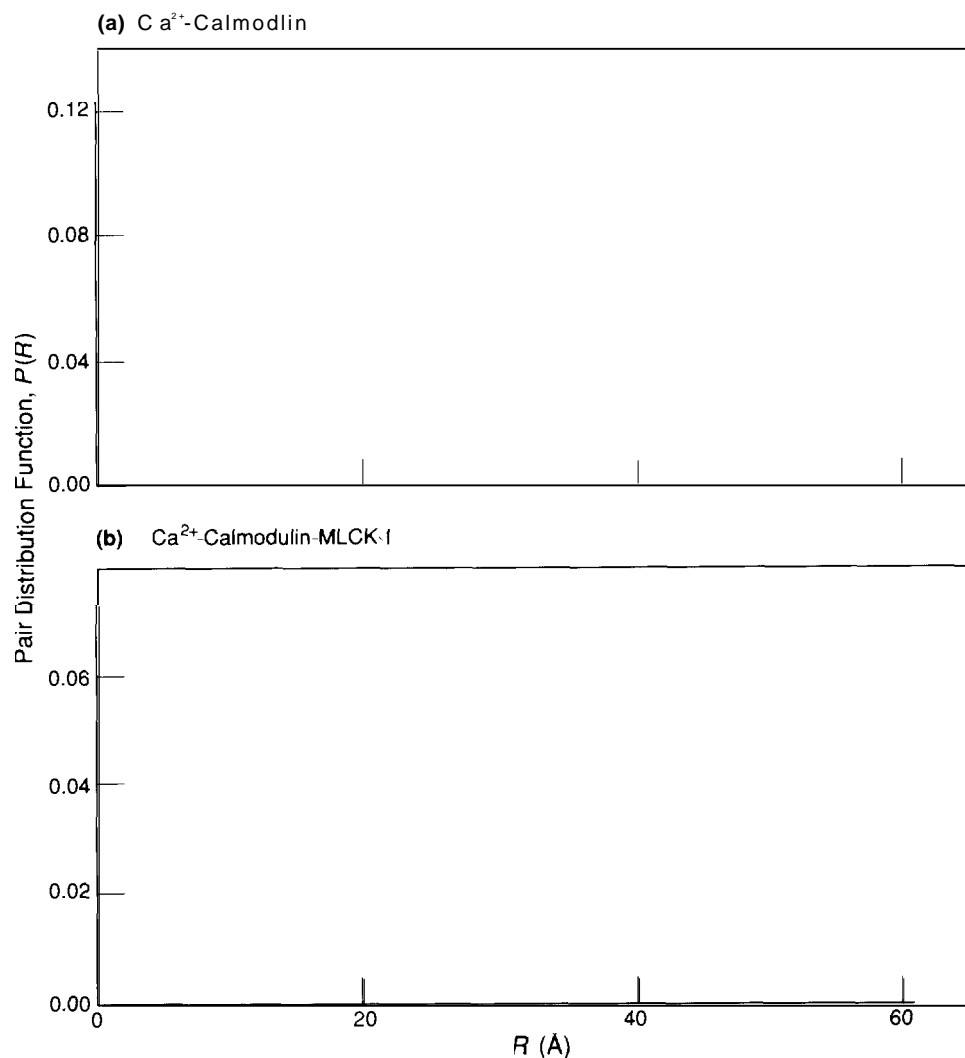
tical to that of the MLCK binding domain has been used in several instances as a model for studying the interaction between MLCK and  $\text{Ca}^{2+}$ -calmodulin. The synthetic polypeptide, referred to as MLCK-I, has been shown to bind one-to-one and with high affinity to  $\text{Ca}^{2+}$ -calmodulin. Moreover, binding of the synthetic polypeptide inhibits  $\text{Ca}^{2+}$ -calmodulin from activating MLCK and other target enzymes. Thus the binding of MLCK-I appears to be a good model for the binding of the entire enzyme. We knew, from circular-dichroism (selective absorption of circularly polarized

light) and nuclear-magnetic-resonance studies of calmodulin and MLCK-I, that upon binding, MLCK-I assumes an  $\alpha$ -helical conformation and  $\text{Ca}^{2+}$ -calmodulin undergoes global structural changes. However, those studies did not provide details of the final structures of the two components of the complex.

We decided to do some small-angle x-ray- and neutron-scattering experiments to determine the relative positions of the two components in the complex as well as the overall shape changes that occur on binding. Our initial guess was that the structure of  $\text{Ca}^{2+}$ -calmodulin in the complex might be more like its crystal structure because of stabilizing interactions between the  $\alpha$  helix of MLCK-I and the interconnecting  $\alpha$  helix of calmodulin.

We collected neutron-scattering data for the complex formed by MLCK-I and  $\text{Ca}^{2+}$ -calmodulin in five aqueous solvents containing various percentages of  $\text{D}_2\text{O}$ . By using deuterated calmodulin, we ensured that  $\text{Ca}^{2+}$ -calmodulin and MLCK-I had different mean scattering-length densities. Figure 8 shows plots of  $I(Q)$  versus  $Q^*$  for the five solvents and the Fourier transforms of the pair-distribution functions fitted to the data (solid lines). Because the 37-percent  $\text{D}_2\text{O}$  solvent has approximately the same scattering-length density as MLCK-I, the solid line for the 37-percent  $\text{D}_2\text{O}$  solvent is a first-order approximation to the Fourier transform of the pair-distribution function for the complexed  $\text{Ca}^{2+}$ -calmodulin alone.

We also collected neutron-scattering data for deuterated  $\text{Ca}^{2+}$ -calmodulin in pure  $\text{H}_2\text{O}$  in the absence of MLCK-I. As shown in Figure 9a, the pair-distribution function for  $\text{Ca}^{2+}$ -calmodulin deduced from the neutron-scattering data is asymmetric and indicates an elongated structure consistent with the flexible model. Note also that Figure 9a is similar the pair-distribution function for  $\text{Ca}^{2+}$ -calmodulin in solution deduced



#### NEUTRON RESULTS FOR $\text{Ca}^{2+}$ -CALMODULIN AND $\text{Ca}^{2+}$ -CALMODULIN-MLCK-I

**Fig. 9.** The pair-distribution function for  $\text{Ca}^{2+}$ -calmodulin (a) resembles that of a double-lobed structure, whereas the pair-distribution function for  $\text{Ca}^{2+}$ -calmodulin-MLCK-I (b) resembles that of a spherical structure. Apparently when  $\text{Ca}^{2+}$ -calmodulin binds to the peptide, the lobes of calmodulin bend around the peptide. Both pair-distribution functions were deduced from small-angle neutron-scattering data. The solvent in both cases was pure  $\text{H}_2\text{O}$ .

from x-ray-scattering data (see Fig. 6). The differences that do exist can be attributed mostly to differences in the  $Q$  range measured but also to differences between the neutron scattering-length density and the electron density of  $\text{Ca}^{2+}$ -calmodulin. Both x-ray and neutron data support very similar

structures for  $\text{Ca}^{2+}$ -calmodulin in solution.

Now compare Figure 9a with Fig. 9b, the pair-distribution function for  $\text{Ca}^{2+}$ -calmodulin when it is complexed with MLCK-I in pure  $\text{H}_2\text{O}$ . The result is a surprise. Rather than becoming more like the crystal form, the structure of the



complexed  $\text{Ca}^{2+}$ -calmodulin becomes much more like a sphere, and its maximum dimension decreases by about 30 percent. Moreover, the radius of gyration of the  $\text{Ca}^{2+}$ -calmodulin-MLCK-I complex is much smaller than that of  $\text{Ca}^{2+}$ -calmodulin.

The data set in Fig. 8 contains much more structural information than do data obtained from a single solvent. In particular, through a method of analysis developed by Stuhrmann, the relative dispositions of the two proteins in the  $\text{Ca}^{2+}$ -calmodulin-MLCK-I complex can be deduced. We begin by determining values for  $R_g^2$  and  $d_{\max}$  from each pair-distribution function (Table 3). Note that the values of  $R_g^2$  and  $d_{\max}$  vary as the percentage of  $\text{D}_2\text{O}$  in the solvent varies. The smallest  $d_{\max}$  at 37-percent  $\text{D}_2\text{O}$  (where MLCK-I does not contribute to the scattering), characterizes the maximum dimension of  $\text{Ca}^{2+}$ -calmodulin when it is bound to MLCK-I. The variation of  $d_{\max}$  among the results for the other solvents (to which both  $\text{Ca}^{2+}$ -calmodulin and MLCK-I contribute) reflects relatively large uncertainties arising from the finite range of  $Q$  for which we can measure  $Z(Q)$ .

The variation in  $R_g$  is more interesting. As the solvent is varied, the contrast factor for each component in the complex changes, and so does the relative contribution of each component to  $R_g$ . Changes in  $R_g$  thus reflect variations in scattering-length density within the complex. Stuhrmann's method involves separating  $R_g^2$  into a term that is independent of variations in scattering-length density,  $R_V^2$ , and terms that are not. That is,

$$R_g^2 = R_V^2 + \frac{\alpha}{b_c} - \frac{\beta}{b_c^2}, \quad (10)$$

where  $b_c \equiv \bar{b}_p - b_s$ , the mean contrast factor of the whole complex.  $R_V$  can be interpreted as the radius of gyra-

**Table 3**

### VARIATION OF STRUCTURAL PARAMETERS WITH CONTRAST FACTOR

Analysis of the pair-distribution function derived from small-angle neutron-scattering data provides the structural parameters  $I(0)$ ,  $R_g$ , and  $d_{\max}$ . Listed here are values of those parameters for  $\text{Ca}^{2+}$ -calmodulin dispersed in  $\text{H}_2\text{O}$  and for  $\text{Ca}^{2+}$ -deuterated calmodulin-MLCK-I dispersed in water containing various percentages of  $\text{D}_2\text{O}$ . As discussed in the text, the variation of  $R_g$  with contrast factor contains further structural information.

	Percentage $\text{D}_2\text{O}$	$I(0)$ ( $\text{cm}^{-1}$ )	$R_g$ ( $\text{\AA}$ )	$d_{\max}$ ( $\text{\AA}$ )
$\text{Ca}^{2+}$ -calmodulin	0	$0.965 \pm 0.006$	$19.4 \pm 0.2$	$59 \pm 5$
$\text{Ca}^{2+}$ -( $^2\text{H}_1$ ) calmodulin-MLCK-I	0	$0.846 \pm 0.005$	$15.1 \pm 0.5$	$45 \pm 5$
	18	$0.564 \pm 0.004$	$16.0 \pm 0.2$	$45 \pm 5$
	37	$0.322 \pm 0.004$	$16.7 \pm 0.2$	$41 \pm 5$
	56	$0.199 \pm 0.006$	$18.2 \pm 0.6$	$49 \pm 5$
	75	$0.088 \pm 0.002$	$18.0 \pm 0.5$	$53 \pm 10$

tion of the complex at infinite contrast, that is, when the complex has a uniform scattering-length density, or no variation in scattering-length density. Thus,  $R_V$  is a measure of the overall shape of the whole complex and remains constant as the mean contrast factor is varied. The coefficient  $\alpha$  is related to the second moment of  $(b_p(\mathbf{r}) - b_s)$  about the mean and therefore measures the change of scattering-length density with distance from the center of scattering power. The coefficient  $\beta$  is related to the square of the first moment of  $(b_p(\mathbf{r}) - b_s)$  and measures the separation between the centers of scattering power of the two scattering components.

To determine  $R_V$ ,  $\alpha$ , and  $\beta$ , we fit a parabola to a plot of  $R_g^2$  versus  $b_c^{-1}$ . Obviously we must know  $b_c$ , and therefore  $\bar{b}_p$ , to perform the Stuhrmann analysis. We mentioned earlier that  $\bar{b}_p$  can be determined from a linear fit to a plot of  $\sqrt{I(0)}$  versus  $b_s$ .

We applied the Stuhrmann analysis to the values of  $R_g$  listed in Table 3. We obtained a positive value for

$\alpha$ , which indicates that the component with the higher scattering-length density (deuterated  $\text{Ca}^{2+}$ -calmodulin) is situated more toward the outside of the complex and the component with the lower scattering-length density (MLCK-I) is situated more toward the inside of the complex. The nonzero value derived for  $\beta$  indicates that the centers of scattering power of the two proteins are not coincident.

The values we derived for  $\alpha$ ,  $\beta$ , and  $R_V$  from the Stuhrmann analysis had large errors. The mean contrast factors of all the samples we studied were positive because  $\bar{b}_p > b_s$ , the mean scattering-length density of the complex containing deuterated  $\text{Ca}^{2+}$ -calmodulin and MLCK-I exceeded the scattering-length density of pure  $\text{D}_2\text{O}$ . As a result, we calculated  $R_V$  by extrapolation rather than by interpolation. The resulting uncertainty in  $R_V$  results in large errors in  $\alpha$  and  $\beta$ .

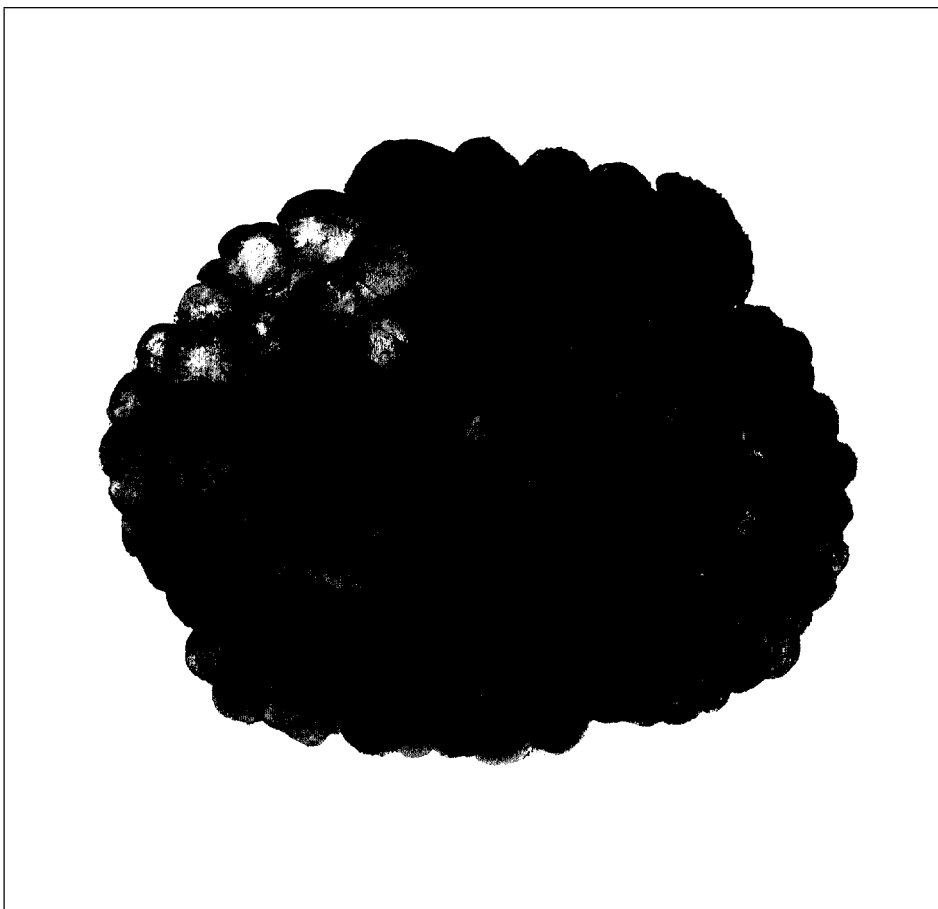
Nevertheless, our qualitative results for  $\alpha$  and  $\beta$  are consistent with the idea that, when MLCK-I interacts with  $\text{Ca}^{2+}$ -calmodulin, MLCK-I binds near

calmodulin's interconnecting helix and the two globular lobes of calmodulin form additional contacts with each other by closing in around at least part of the peptide. An artist's conception of the configuration is shown in Fig. 10. Thus our neutron-scattering experiments have yielded a surprising view of the interaction between  $\text{Ca}^{2+}$ -calmodulin and the binding domain of one target enzyme.

That view is not necessarily applicable to calmodulin's interaction with all of its target enzymes. In particular, our experiments on phosphorylase kinase (PhK) show rather different results. PhK, a large enzyme containing many subunits, catalyzes the transfer of a phosphate group from ATP to glycogen phosphorylase, thereby activating glycogen phosphorylase as a catalyst in the conversion of glycogen to glucose (Fig. 11). PhK is made up of four identical copies each of four different polypeptides, denoted  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ . The  $\gamma$  subunit contains the catalytic site; the  $\alpha$ ,  $\beta$ , and  $\delta$  subunits play regulatory roles in response to various physiological signals. The protein composing the  $\delta$  subunit is calmodulin. When  $\text{Ca}^{2+}$  binds to those intrinsic calmodulin molecules, the catalytic activity of the  $\gamma$  subunit increases twentyfold. Calmodulin's activation of PhK is fundamentally different from its activation of MLCK in that the calmodulin remains associated with the  $\delta$  subunit in the absence of  $\text{Ca}^{2+}$ .

It has recently been shown that  $\text{Ca}^{2+}$ -calmodulin binds to the  $\gamma$  subunit at two noncontiguous sites, each containing twenty-five amino-acid residues. The binding between  $\text{Ca}^{2+}$ -calmodulin and the  $\gamma$  subunit in PhK apparently encompasses a more extensive region than does the binding of  $\text{Ca}^{2+}$ -calmodulin to MLCK.

We have recently completed x-ray studies of  $\text{Ca}^{2+}$ -calmodulin complexed with one or the other or both of two synthetic polypeptides corresponding to



#### STRUCTURE OF $\text{Ca}^{2+}$ -CALMODULIN BOUND TO MLCK-I

**Fig. 10.** This artist's conception of  $\text{Ca}^{2+}$ -calmodulin bound to MLCK-I is based on small-angle neutron-scattering data. Those data suggest that the  $\alpha$ -helical peptide is in contact with calmodulin's interconnecting  $\alpha$  helix and that the two globular lobes of calmodulin bend around the peptide to form additional contacts.

the two binding domains on the  $\gamma$  subunit of PhK. The two synthetic polypeptides are designated PhK5 and PhK13. In addition, we have gathered neutron-scattering data for deuterated  $\text{Ca}^{2+}$ -calmodulin complexed with both synthetic polypeptides. The results are again a surprise. The binding of PhK5 alone results in a contraction of  $\text{Ca}^{2+}$ -calmodulin quite similar to that induced by MLCK-I. The x-ray-scattering data for the complex give values of 17.1 angstroms and 50 angstroms for

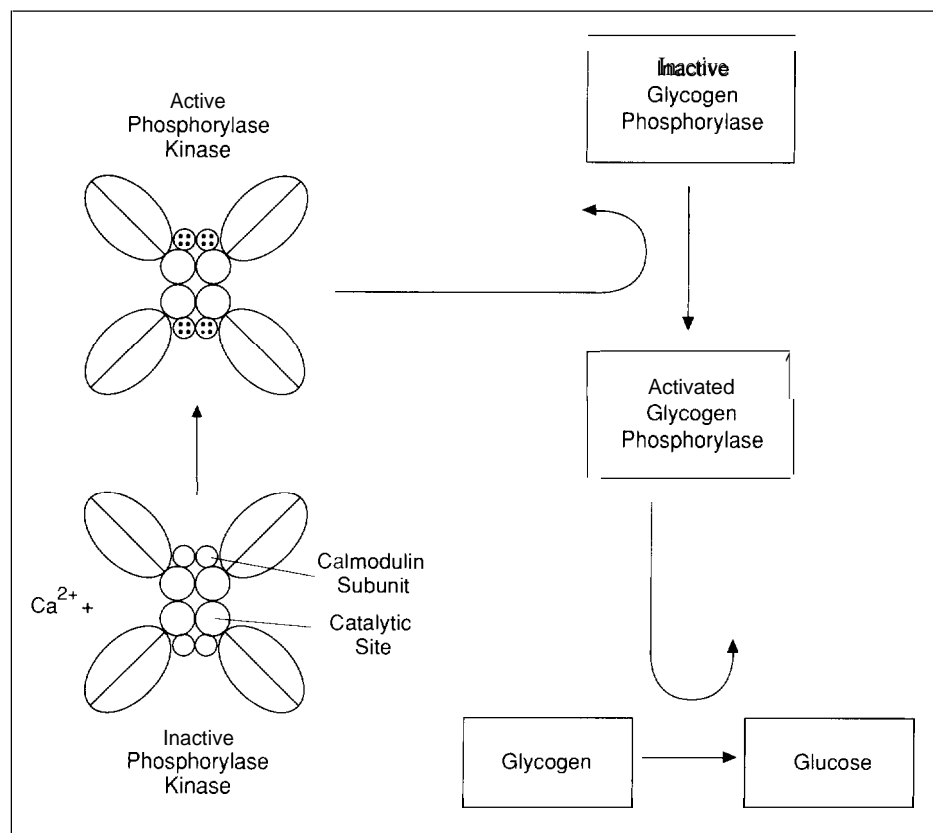
$R_g$  and  $d_{max}$ , respectively. Apparently PhK5 and MLCK-I bind similarly to calmodulin. The amino-acid sequences of PhK5 and MLCK-I are such that both peptides are predicted to have a high propensity for forming an amphipathic helix, a configuration that may facilitate interaction with the interconnecting helix of calmodulin. The sequence of PhK13 is such that the peptide is predicted to have little or no  $\alpha$  helix but significant  $\beta$  structure. The x-ray data for  $\text{Ca}^{2+}$ -calmodulin com-

# REGULATION OF GLYCOGEN DEGRADATION BY $\text{Ca}^{2+}$ THROUGH CALMODULIN INTRINSIC TO PhK

Fig. 11. Another target enzyme of  $\text{Ca}^{2+}$ -calmodulin is phosphorylase kinase (PhK), a large (1.3 million daltons) enzyme consisting of four identical copies of each of four different polypeptide subunits, one of which is calmodulin. PhK is involved in the cascade of interactions, initiated by a hormone binding to a cell, that results in the degradation of glycogen to glucose. Metabolism of the glucose then provides energy for use by the cell in various processes, including muscle contraction. The activation of PhK by  $\text{Ca}^{2+}$ -calmodulin differs from that of most of  $\text{Ca}^{2+}$ -calmodulin's target enzymes in that the calmodulin is an intrinsic subunit of PhK. Binding of  $\text{Ca}^{2+}$  to those intrinsic calmodulin subunits activates the subunits containing the sites at which catalysis of the activation of glycogen phosphorylase occurs.

plexed to PhK 13 show no contraction at all, but rather an overall extension of the structure ( $R_g \approx 28$  angstroms and  $d_{\max} \approx 90$  angstroms). The complex with both PhK5 and PhK13 is also extended, as indicated by  $R_g$  and  $d_{\max}$  values not significantly different from those of  $\text{Ca}^{2+}$ -calmodulin-PhK 13.

Neutron-scattering data for deuterated  $\text{Ca}^{2+}$ -calmodulin complexed with both PhK5 and PhK13 add further insight into the nature of the complex. The scattering-length densities of both peptides are approximately equal to that of a solvent containing 40 percent  $\text{D}_2\text{O}$ . The pair-distribution function determined from scattering data for the complex dispersed in such a solvent shows that calmodulin itself is extended ( $R_g \approx 26$  angstroms and  $d_{\max} \approx 90$  angstroms). That result indicates that PhK13 influences the way in which  $\text{Ca}^{2+}$ -calmodulin binds to PhK5. In particular, PhK13 apparently prevents PhK5 from inducing a conformational change in  $\text{Ca}^{2+}$ -calmodulin similar to the contraction



induced by MLCK-I. The differences between the conformational changes induced in  $\text{Ca}^{2+}$ -calmodulin by the two target enzymes (phosphorylase kinase and myosin light-chain kinase) are likely to be due to differences in the nature of its interactions with the two enzymes, and they provide a clue as to the purpose of the unusual interconnecting helix in calmodulin. That flexible domain clearly facilitates changes in the relationship between the globular domains of calmodulin that allow its binding to a diversity of target enzymes and its activation of a wide variety of biochemical processes. Our goal is to use neutron-scattering data, in conjunction with other structural data, to build up a complete picture of the molecular basis for  $\text{Ca}^{2+}$ -dependent activation of target enzymes by calmodulin. Future studies will focus on developing methods

for preparing complexes of deuterated calmodulin with intact target enzymes. In addition, the interactions of  $\text{Ca}^{2+}$ -calmodulin with other target enzymes will be investigated.

As is apparent from this discussion, unraveling the molecular basis of a biological function is a long and arduous task. Neutron-scattering studies, although difficult, are providing a view of biological molecules in functional states that were heretofore unobservable. Their role in studies of biological function should continue to grow and to yield important clues concerning the structural dynamics of complex biological units. □

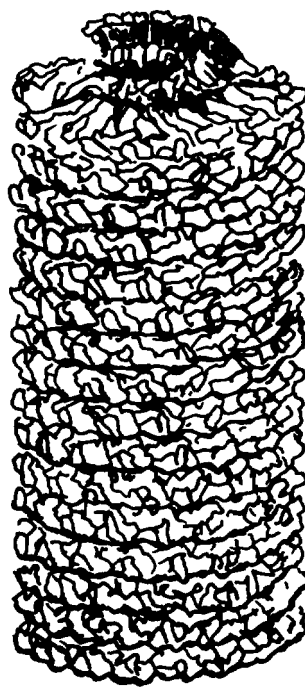
# FERROFLUIDS

## *a new alignment technique*

Typical neutron-scattering experiments on particles in solution yield only one-dimensional (spherically averaged) data because the scattering particles are randomly oriented. Such data can provide a model for the general shape and the boundaries of the scattering particles, but often more than one model is consistent with the data. However, if the particles are partially ordered by being given a definite orientation, the resulting scattering data contain more detailed structural information.

Alignment in one dimension has been achieved for the rod-shaped tobacco mosaic virus by applying shearing forces to an aqueous gel of the viruses. Two-dimensional x-ray-diffraction data for the aligned viruses led to a complete three-dimensional structure at a resolution of 3.6 angstroms (Fig. 1). Although that result is very impressive, the technique is not applicable to many biological structures.

Here we present preliminary studies of a new technique for aligning elongated biological assemblies in solution. The technique involves dispersing the assemblies in a ferrofluid (a fluid in which magnetic particles are suspended) and applying a moderate magnetic field. (Note that because the magnetic particles generally contain iron, which absorbs x rays strongly, the ferrofluid alignment technique is applicable to neutron-scattering experiments but not to x-ray-scattering experiments.) Magnetic forces cause the moments, or spins, of the magnetic particles to align along the direction of the mag-



**STRUCTURE OF TMV**

**Fig. 1.** The rod-shaped tobacco mosaic virus is about 3000 angstroms long and 180 angstroms in diameter. Its structure at a resolution of 3.6 angstroms was deduced from x-ray diffraction data for an aqueous gel of viruses. The viruses were oriented in one direction by shearing forces. The viral genetic material (RNA) is enclosed within a protein coat consisting of a helical array of many copies of a single protein species. The distance between the turns of the helix is 23 angstroms.

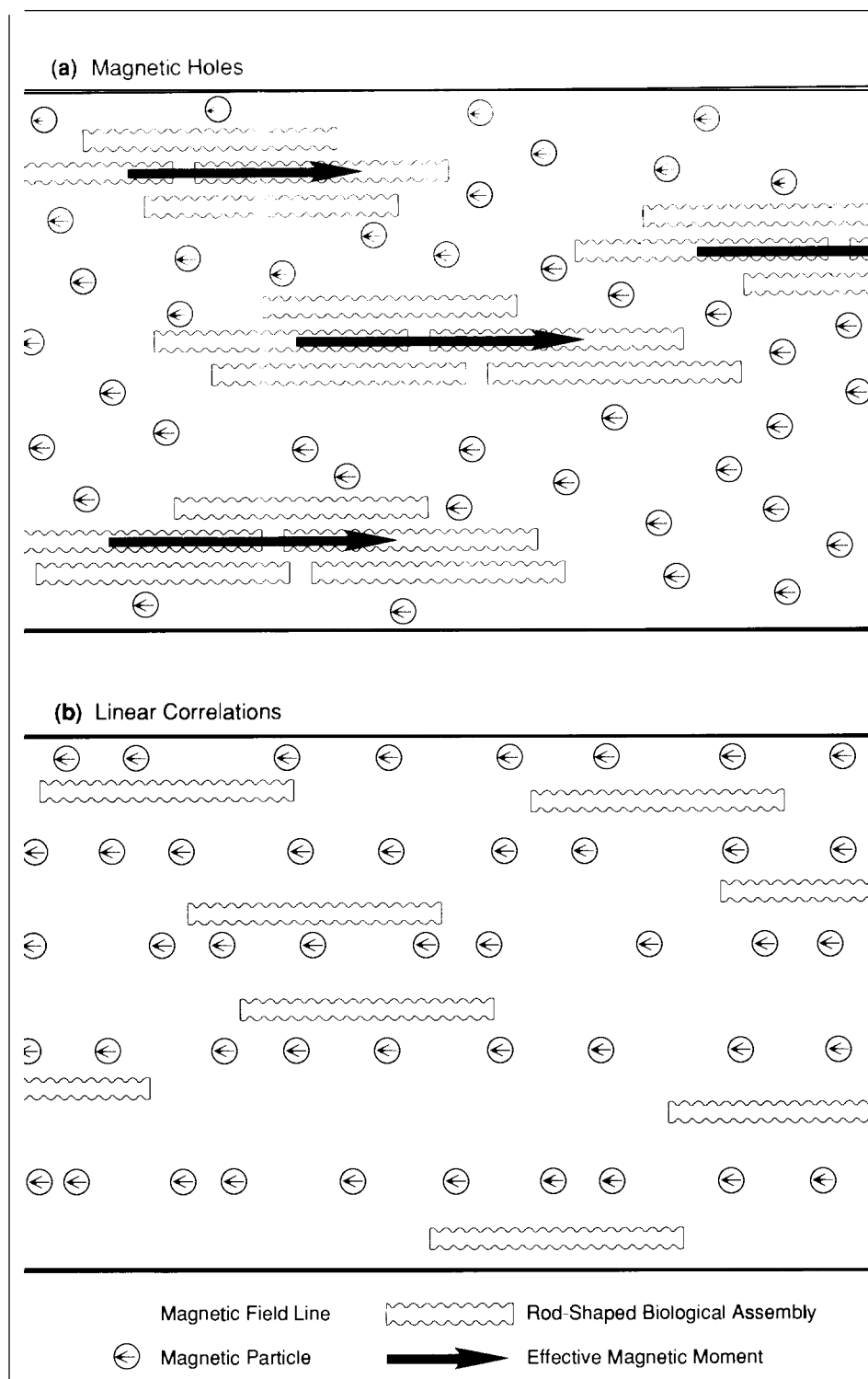
netic field, and their alignment, in turn, causes alignment of the elongated biological assemblies. Some biological assemblies have intrinsic diamagnetic moments and will therefore align along a magnetic field in the absence of mag-

netic particles. However, very strong magnetic fields are usually required. In contrast, the ferrofluid technique requires no intrinsic magnetic properties of the biological assemblies and only moderate magnetic fields.

We have tested the ferrofluid technique on two viruses, the tobacco mosaic virus (TMV) and the tobacco rattle virus (TRV), and have obtained neutron-diffraction data that testify to its success. Our tests focused not only on obtaining a high degree of alignment but also on understanding how the alignment comes about. Figure 2 shows two possible mechanisms. One possibility is that the biological assemblies act like "magnetic holes." That is, by displacing the ferrofluid, they acquire effective magnetic moments equal in magnitude and opposite in sign to the sum of the moments of the magnetic particles in the displaced ferrofluid. The effective moments then align along the applied magnetic field. In our experiments, however, the effective moments would be large enough to cause alignment of the assemblies in moderate magnetic fields only if the assemblies existed as ordered domains, or aggregates. Then each domain becomes a magnetic hole oriented along the applied field (Fig. 2a). The other possible mechanism of alignment invokes linear correlations among the magnetic particles, that is, a tendency for them to line up in rows along the applied field. Because disrupting those linear correlations would require energy, the non-magnetic elongated assemblies also tend to line up, with their long axes along the field, in rows between the rows of magnetic particles (Fig. 2b). It seems likely at this time that both alignment mechanisms operate to different degrees depending on the relative concentrations of magnetic particles and biological assemblies and whether the solution conditions favor the formation of ordered domains of the biological assemblies.

## POSSIBLE MECHANISMS OF ALIGNMENT IN A FERROFLUID

**Fig. 2.** Elongated biological assemblies dispersed in a ferrofluid align along a moderate applied magnetic field even if they lack intrinsic magnetic properties. Shown here are two possible mechanisms of alignment. (a) Ordered domains of the biological assemblies act like magnetic holes. In other words, the domains acquire effective magnetic moments equal in magnitude and opposite in direction to the sum of the magnetic moments of the magnetic particles they displace. The effective magnetic moments cause the biological assemblies to align along the field. (b) Long-range magnetic correlations among the magnetic particles cause them to align in rows along the applied field. Lacking the energy to disrupt the linear correlations, the elongated biological assemblies also align along the field.

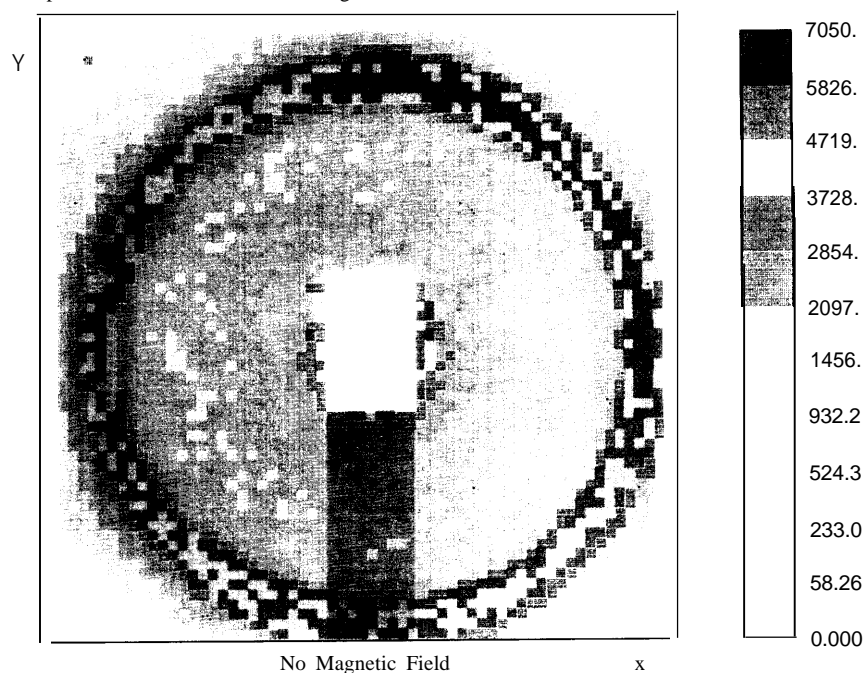


The magnetic particles in the ferrofluid would normally contribute to the observed scattering intensity. However, that contribution can be eliminated by matching the scattering-length density of the magnetic particles to that of the solvent. By a happy coincidence of nature, the neutron scattering-length density of the ferromagnetic material magnetite is quite close to that of  $D_2O$ . As is well known, particles of magnetite align with an applied magnetic field. To achieve a

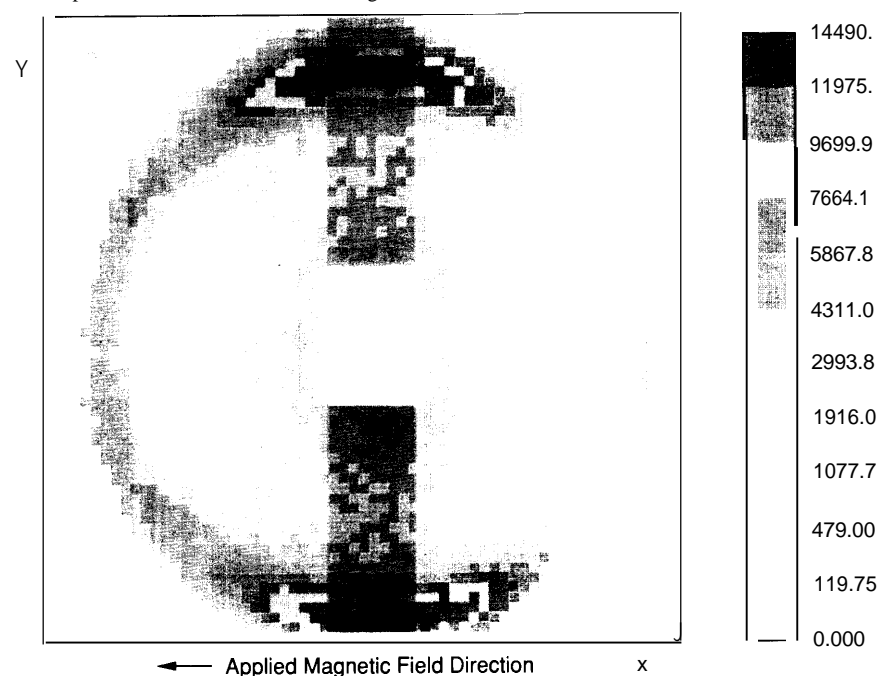
reasonably homogeneous suspension of magnetite particles in  $D_2O$ , the particles must be coated with a detergent surfactant. The detergent must be carefully chosen, however, to avoid its degrading the biological assemblies. Through extensive electron microscopy we identified a detergent that did not degrade TMV or TRV. We then deuterated the detergent to match the scattering-length densities of magnetite and  $D_2O$ .

To find out how well our ferrofluid

(a) Isotropic Diffraction Peak at 363 angstroms



(b) Anisotropic Diffraction Peak at 363 angstroms



### SMALL-ANGLE NEUTRON-SCATTERING DATA FOR TMV IN A FERROFLUID

Fig. 3. Color-coded intensity data for neutrons scattered from tobacco mosaic virus dispersed in a ferrofluid. The intensity corresponding to the color of a detector pixel is related to the number of neutrons detected in the pixel. (a) In the absence of a magnetic field, the scattering is isotropic and a diffraction peak (a circle of high intensity) appears at a scattering angle corresponding to a distance of 363 angstroms. The diffraction peak reflects the packing distance between individual viruses in ordered domains. (b) When a magnetic field of 0.3 tesla is applied along the x direction, the diffraction peak appears only in the y direction, that is, perpendicular to the applied field. That result means that the packing distances are aligned perpendicular to the field and the long axes of the viruses are aligned along the field. The degree of alignment cannot be explained by intrinsic diamagnetism of the viruses and must therefore be caused by the ferrofluid.

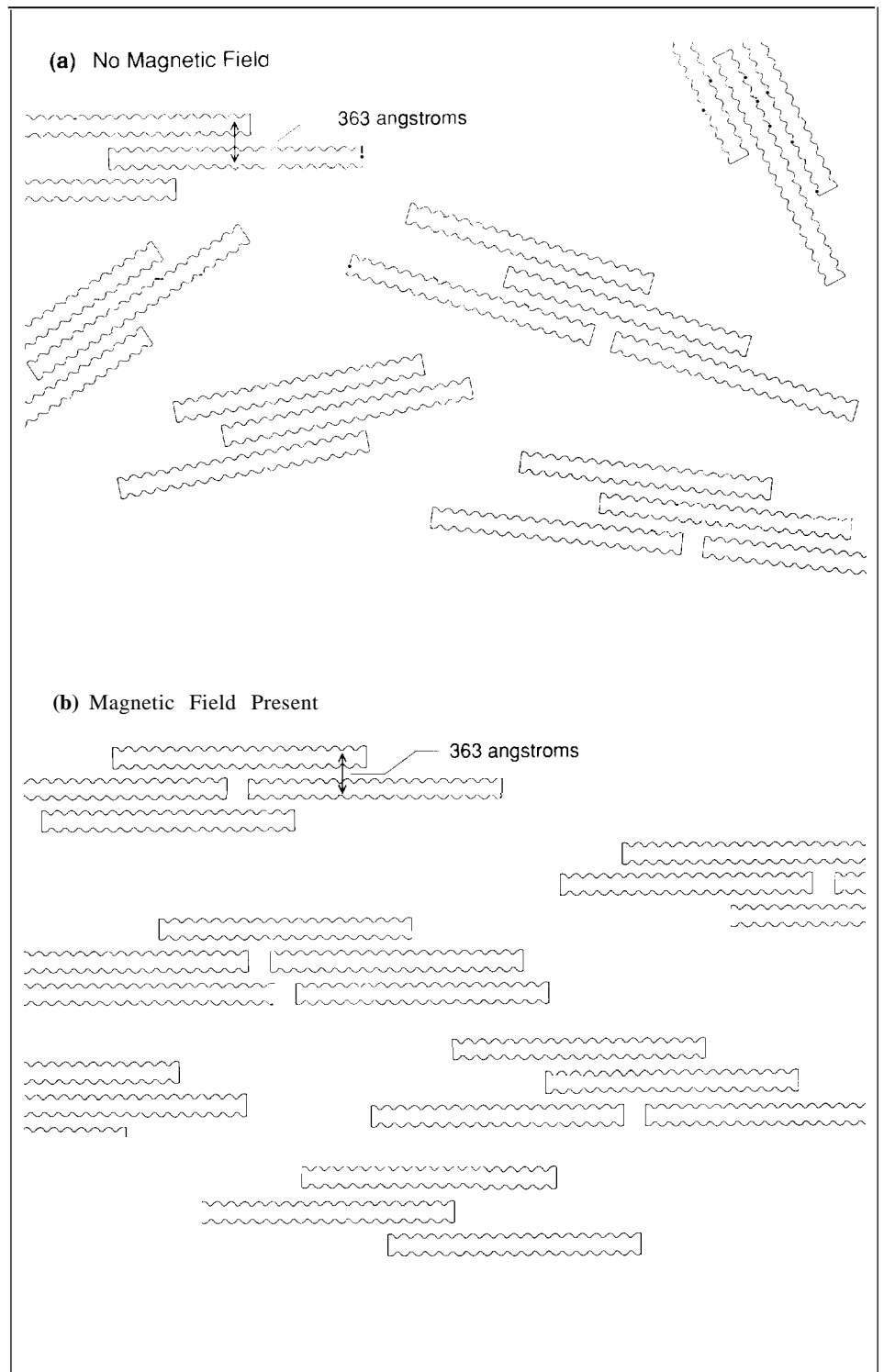
worked, we obtained neutron-scattering data for 0.06-volume-fraction dispersions of each virus in the ferrofluid. In the absence of a magnetic field, the scattering from TMV was isotropic, as expected (Fig. 3a). We also saw a diffraction peak corresponding to a distance of 363 angstroms. That diffraction peak arises because the viruses formed ordered domains in which the packing distance was 363 angstroms (Fig. 4a). (The packing distance is determined by

a repulsive electrostatic force between the viruses and varies with the ionic strength of the solution.) In a modest magnetic field of 0.3 tesla, the scattering was anisotropic, and the diffraction peak at 363 angstroms appeared only in the direction perpendicular to the field (Fig. 3b). Those results suggest that the TMV domains were aligned in such a way that the long axes of the viruses were parallel to the applied field.

If the TMV domains do so align,

### ORDERED AGGREGATES OF TMV IN A FERROFLUID

**Fig. 4.** Tobacco mosaic viruses tend to form ordered domains in solution, as shown in the electron micrograph below. The viruses within a domain are kept at a certain distance by electrostatic forces. That distance depends on the pH and ionic strength of the solution and in our experiments was 363 angstroms. (a) In the absence of a magnetic field, the domains are randomly oriented in a ferrofluid. (b) An applied magnetic field aligns the magnetic particles in the ferrofluid, which, in turn, causes the domains to align along the field. Thus the 363-angstrom spacing between the viruses is perpendicular to the field and produces diffraction peaks only in that direction.

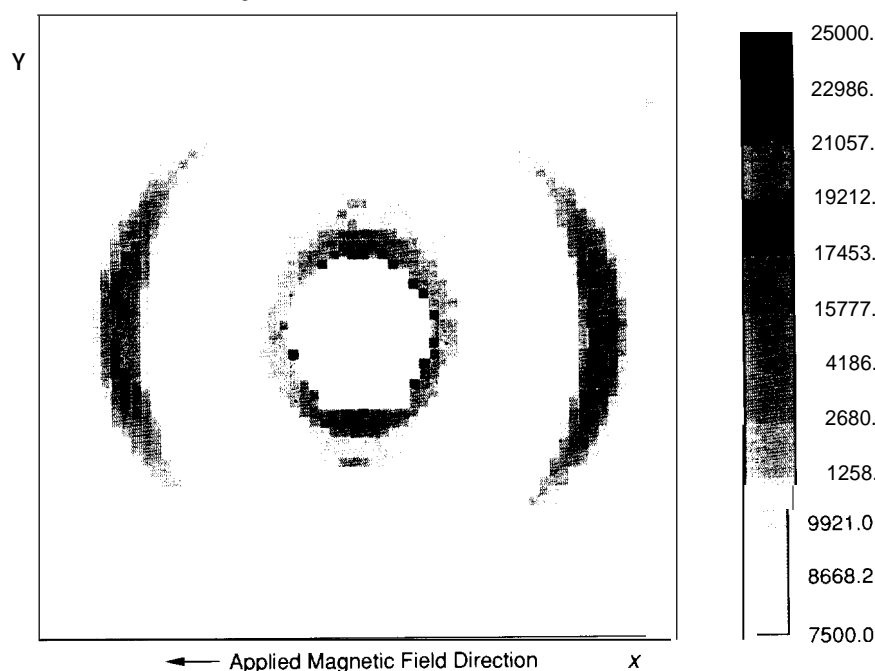


another diffraction peak should be seen, one produced by the 23-angstrom distance between the turns of the helical viral coat protein (see Fig. 1). That peak should occur at scattering angles much greater than those included in the data of Fig. 3 and at a direction perpendicular to the direction at which the diffraction peak due to the packing distance is observed. We moved the detector closer to the sample and did indeed see a diffraction peak corresponding to a

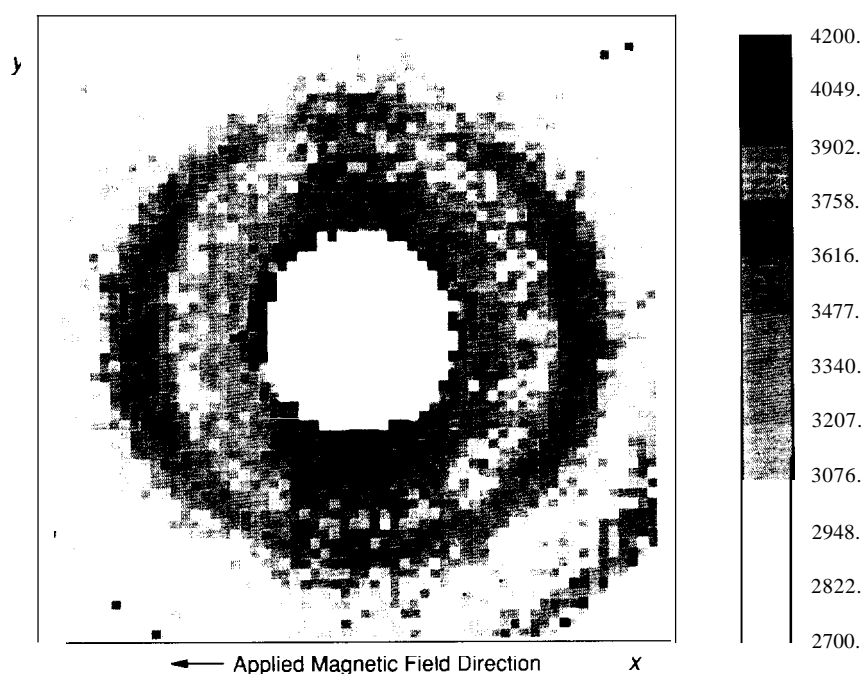
23-angstrom helical pitch (Fig. 5a). The existence of the peak confirms the structural integrity of the viruses in the ferrofluid, and its orientation parallel to the magnetic field is consistent with alignment of the long axes of the viruses parallel to the magnetic field (Fig. 4b).

Although our experiments demonstrated alignment, we wondered whether it was caused by intrinsic magnetic properties of TMV or by one of the mechanisms depicted in Fig. 2. The

(a) Diffraction Peak at 23 angstroms



(b) Diffraction Peak at 25 angstroms



## NEUTRON DIFFRACTION PEAKS FROM TMV AND TRV

Fig. 5. Color-coded intensity data for (a) tobacco mosaic virus and (b) tobacco rattle virus dispersed in a ferrofluid in the presence of a 0.3-tesla applied magnetic field. Note the diffraction peaks at 23 and 25 angstroms, which correspond to the regular spacing between the turns of the helical protein coats of TMV and TRV, respectively. The diffraction peaks appear only in the x direction, that is, along the direction of the applied field. Thus, the ferrofluid has aligned the viruses with their long axes along the applied magnetic field.

trated samples of TRV did not align in a normal solution even in a high applied magnetic field (7 teslas). However, when TRV was dispersed in the ferrofluid, a field of 0.5 tesla was sufficient to produce alignment comparable to that obtained for TMV in phosphate buffer, Figure 5b shows a diffraction peak corresponding to the 25-angstrom pitch of the helical TRV protein coat, similar to the peak obtained for TMV in Fig. 5a. The position of the peak indicates that the viruses are aligned parallel to the applied magnetic field. Moreover, like TMV in phosphate buffer, TRV showed no evidence of any domains.

Our preliminary studies demonstrated that the ferrofluid technique can orient elongated biological assemblies irrespective of their intrinsic magnetic properties and without disrupting their structural integrity. Furthermore, the orientation is sufficiently great to facilitate measurement of internal structural parameters. Future studies will concentrate on increasing the degree of alignment so that higher-resolution data can be obtained. ■

viruses could have been ordered by diamagnetic moments; however, the magnetic field strength was smaller than is generally needed to achieve the degree of alignment indicated by the data. Such alignment often requires the presence of ordered domains in the sample. So we repeated the experiments, this time adding phosphate buffer, the ions of which are known to disrupt domain structure. Thus alignment based on the intrinsic diamagnetism of the viruses is

not expected in the presence of phosphate buffer. We found that the viruses were still aligned but to a lesser degree. That result suggests not only that the ordering is due to the presence of the ferrofluid but also that the mechanism depicted in Fig. 2b may enhance the alignment of TMV.

The tobacco rattle virus is genetically unrelated but morphologically quite similar to TMV. TRV shows only poor orientation under shear, and concen-



## Acknowledgments

I am pleased to express my gratitude to the people whose collaboration made possible the experiments described in this article. In particular, the calcium-regulation experiment required the expertise and resources of Sue Rokop and Anthony Means in molecular biology and biochemistry, of Philip A. Seeger, Douglas Heidorn, and Stephen Henderson in neutron and x-ray scattering, of David Hobart and Phillip Palmer in plutonium chemistry, of Donald Blumenthal in target-enzyme activation and peptide synthesis, and of Henri Crespi in preparing deuterated materials. The experiments on ferrofluids were a collaborative effort with Stuart Charles, Gerald Stubbs, Tobin Sosnick, Peter Timmins, Roger Pynn, and John Hayter.

## Further Reading

Benno P. Schoenborn, editor. 1976. *Neutron Scattering for the Analysis of Biological Structures: Report of Symposium Held June 2-6, 1975*. Brookhaven National Laboratory report 50543.

B. Jacrot. 1976. The study of biological structures by neutron scattering from solution. *Reports on Progress in Physics* 39:911-953.

O. Glatter and O. Kratky, editors. 1982. *Small Angle X-ray Scattering*. London: Academic Press.

Jill Trehwella, J.-L. Popot, D. M. Engelman, and G. Zaccai, 1986. Neutron diffraction studies of bacteriorhodopsin. *Physics* 136B: 249.

Jill Trehwella, Jean-Luc Popot, Giuseppe Zaccai, and Donald M. Engelman. 1986. Localization of two chymotryptic fragments in the structure of renatured bacteriorhodopsin by neutron diffraction. *The EMBO Journal* 5: 3045.

Philip Cohen and Claude B. Klee, editors. 1988. "Calmodulin." In *Molecular Aspects of Cellular Regulation*, edited by Philip Cohen. Amsterdam: Elsevier.

Douglas B. Heidorn and Jill Trehwella. 1988. Comparison of the crystal and solution structures of calmodulin and troponin C. *Biochemistry* 27: 909.

D. B. Heidorn, P. A. Seeger, S. E. Rokop, D. K. Blumenthal, A. R. Means, H. Crespi, and J. Trehwella. 1989. Changes in the structure of calmodulin induced by a peptide based on the calmodulin-binding domain of myosin light chain kinase. *Biochemistry* 28: 6757.

R. A. Kowluru, D. B. Heidorn, S. P. Edmondson, M. W. Bitensky, A. Kowluru, N. W. Downer, T. W. Whaley, and J. Trehwella. 1989. Glycation of calmodulin: Chemistry and structural and functional consequences. *Biochemistry* 28: 2220.

D. B. Heidorn and J. Trehwella. Low-resolution structural studies of proteins in solution: Calmodulin. Accepted for publication in *Comments on Molecular and Cellular Biophysics: A Journal of Critical Discussion of the Current Literature (Comments on Modern Biology A)*.

J. Trehwella, D. B. Heidorn, and P. A. Seeger. Solution structures of calcium-binding proteins: A small-angle-scattering study. *Journal of Molecular Crystals and Liquid Crystals* 180A: 45-53.

J. Trehwella, D. K. Blumenthal, S. Rokop, and P. A. Seeger. The solution structure of calmodulin complexed with two peptides based on the regulatory domain of the catalytic subunit of phosphorylase kinase. Submitted to *Biochemistry*.

Keiichi Namba and Gerald Stubbs. 1986. Structure of tobacco mosaic virus at 3.6 Å resolution: Implications for assembly. *Science* 231: 1401.

John B. Hayter, Roger Pynn, Stuart Charles, Arne T. Skjeltorp, Jill Trehwella, Gerald Stubbs, and Peter Timmins. 1989. Ordered macromolecular structures in ferrofluid mixtures. *Physical Review Letters* 62: 1667.



**Jill Trehwella** is an Australian-born biophysicist. She received her B. SC. and M. SC. degrees in physics at the University of New South Wales. Her studies there in molecular structure included an x-ray-crystallographic analysis of an anti-cancer agent that interacts with DNA to inhibit replication. That work directed her interests toward structural biology. Her Ph.D. research at the University of Sydney utilized physical techniques, in particular NMR spectroscopy, to study the structure and function of heme proteins. She came to the United States in the fall of 1980 as a postdoctoral fellow at Yale University, where

she used neutron diffraction to study the structure of the membrane protein bacteriorhodopsin. She came to the Laboratory's Life Sciences Division in 1984 to develop a structural biology program in association with the then newly developing neutron-scattering center. Since that time her research program in several areas of structural biology has utilized neutron and x-ray scattering, as well as other biophysical techniques. One of her major interests is in the molecular mechanisms of calcium regulation. She is also actively involved in structural studies of chromatin and molecules of the immune system.