# **JOHN A. ROTHFUS, Northern Regional Research Laboratory, Peoria, Illinois 61604**

# **Abstract**

This review summarizes theory, experimental procedures and applications of current research methodology in each stage of protein purification from the disruption of seed structures and isolation of protein bodies to cleavage of intermolecular bonds and the purification of individual proteins. Methods for characterizing proteins are discussed in terms of their utility in monitoring protein isolations, disclosing structural features and predicting protein behavior. Specific topics include: the macrostrueture of seeds, scanning electron microscopy, particle fraetionation by density gradient eentrifugation, filtration, association-dissociation of seed proteins, disulfide bonds and other forms of intermolecular bonding, high pressure gel filtration, immunospeeific adsorbents or precipitating agents, immunochemical techniques, isoelectric focusing, physical and chemical methods for characterizing proteins, differential thermal analysis and the determination of polypeptide sequences.

## **Introduction**

Investigation of detailed protein structure, which began with Sanger and Tuppy's studies on insulin (1) and which reached a high point less than 20 years later in the complete chemical synthesis of the enzyme, ribonuclease (2,3), has provided ample evidence that proteins should be thought of as chemical reagents susceptible to the same types of manipulation commonly applied to smaller organic molecules. The new insight allowed by this important concept provides a basis for more fundamental approaches to protein purification and characterization and lets the experimentalist take liberties with protein molecules that were unthinkable less than 30 years ago. For example, denaturation, once a process to be religiously avoided with proteins, now is a useful technique through which unusual separations can be achieved and through which the basic chemistry of proteins can be probed (4).

Like other chemical reagents, proteins respond to environmental changes in ways that minimize the free energy of the system in which they exist. In this light, a search for methods by which a protein may be isolated is, in essence, a search for energy minima in which that protein is more stable than other proteins. It is appealing to think that separation research may ultimately be shortened by computer-aided prognosis; perhaps the locations of energy minima may be predicted from limited knowledge of composition or source and from contour maps that describe the energy content of similar proteins. At present, calculation of protein conformation and conformational changes in response to environmental factors from thermodynamic considerations (5-8) depends heavily upon knowledge of primary structure and other information most easily obtained after the protein is purified. Nevertheless, abstract thought in terms of fundamental physical and chemical concepts can lead to new understanding of seed proteins and innovations in present day techniques.

Traditional protein separation methods, which have developed largely from research on mammalian proteins, commonly achieve resolution either by gradual changes in pH, ionic strength, dielectric constant, ion composition and temperature or by changes in such forces as fluid velocity, electric potential and gravity. The isolation of plant seed proteins may require amendments to this list. From the standpoint that separations are energy-controlled processes, pressure is an obvious and potentially useful variable. Josephs and Harrington  $(9)$  have already shown that hydrostatic pressure favors the dissociation of elastin aggregates into monomers with smaller partial specific volumes. Similar methods might prove useful in dissociating some of the highly aggregated protein mixtures often encountered in plant seeds.

Detailed reviews of the extensive literature dealing with the fractionation (10,11) and characterization of proteins (12-14) have been published elsewhere. This paper briefly considers some further developments in these areas of protein chemistry.

# **Disruption of Cellular Structure**

Figure 1 shows the subcellular structure present in soybean cotyledon cells (15). Protein bodies or aleurone grains  $(11)$ , which account for 70% of the protein in soybeans (16), are prominent features. These bodies, which are more than  $90\%$  protein  $(16)$ , represent the richest source of protein available from soybean. Lipids in this oil seed are mainly confined to smaller globular bodies called spherosomes. Starch granules, when present, are intermediate in size between protein bodies and spherosomes.

Since proteins are separated from lipids and carbohydrates in the native state, many problems in protein purification apparently arise through mixing that occurs as the proteins are extracted from the seed. Where purity is critical, a significant purification of protein can be achieved by isolating intact protein bodies directly. In this way, undesirable interactions of proteins with other cellular components can be avoided. From the standpoint of nutrition, the isolation of intact protein bodies from soybeans would be especially beneficial since they apparently are deficient in trypsin inhibitors (16).

Table I lists methods applicable in the selective disruption of cellular and subcellular structures in plant seeds. At present, some shearing or crushing in the form of homogenization with sand is often used to break cell walls and liberate protein bodies. Although protein bodies have been prepared from cottonseed (17), soybeans (18) and peanuts (19) without previous lipid extraction, the dispersion of spherosomes, which do not withstand homogenization, can lead to bothersome emulsions (16). With soybeans, use of defatted meal simplifies the isolation of protein bodies (16).

Grinding, crushing and other indiscriminant pro-

TABLE I **Methods That Disrupt Cellular Organization** 

Shearing	Sonic oscillation
Crushing	Freeze thaw
Heating	Microwave treatment
Hydration	Enzymic modification
Chemical treatment	Ion removal

<sup>&</sup>lt;sup>1</sup> Presented at the AOCS-AACC Joint Protein Short Course, French Lick, Ind., July 13-16, 1969.



FIG. 1. Electron micrograph of a cross section from a soybean cotyledon cell showing protein bodies (PB), spherosomes  $(S)$ , and cell walls  $(\overline{CW})$ . Source: K. Saio and T. Watanabe (15).

cesses for breaking seeds are crude by comparison to the fragile structures apparent in Figure 1. Perhaps specific chemical treatments, enzymic modifications or freeze-thaw techniques might perform the same task. Once the cell wall is broken, subcellular structure should be amenable to disruption by methods such as sonic oscillation, osmotic shock, treatment with detergents, ion removal or other techniques that have been useful in dissociating structures from mammalian cells (20-25). Protein bodies liberated from homogenized soybean cotyledon cells are stable in a 20% sucrose solution, but will dissociate further in water, and produce particles about 0.5  $\mu$  in diameter  $(16)$ .

## **Density Gradient Centrifugation**

An outgrowth of virus research (26,27), density gradient centrifugation has been especially useful in the preparation of intact protein bodies from soybeans (16) and corn (28). In equilibrium density gradient eentrifugation (29), the gradient is selfgenerating, and cellular particles and gradient components that are mixed initially are separated as equilibrium is established. Thereafter, particle migration is independent of the length of the centrifuge run. In isokinetic density gradient centrifugation, a mixture of particles with different densities is layered on a preformed gradient solution of increasing salt or sucrose concentration (30,31) and then centrifuged for  $15-30$  min at speeds that vary depending on the particles and the gradient. The particles migrate outward under the influence of centrifugal force until their movement is counterbalanced by the viscosity and buoyancy of the increasingly dense gradient solution. The distribution of an homogenate of immature corn endosperm in a sucrose gradient, as described by Christianson et al. (28), is illustrated in Figure 2. Free protein bodies are confined to a band near the bottom of the centrifuge tube. At the top are layers of lipid and soluble protein or nucleic acid. As expected of corn, a substantial starch pellet is at the bottom of the tube. Similar experiments with soybean homogenates show that the pellet can also contain clumped protein bodies, amorphous cell debris and unruptured cells.

To recover material after it has been centrifuged, either commercially available devices can be used to siphon off layers of the gradient from individual tubes or, where the pellet will permit, the tube can



# 15 to 40% Sucrose Gradient

FIG. 2. Density gradient centrifugation of an immature corn endosperm dispersion. Gradient: 15-40% sucrose. Source: D. D. Christianson, Northern Regional Laboratory.

be punctured and the contents allowed to drip directly into a fraction collector (30). For large preparations or semieontinuous operation, centrifuge heads are specially constructed. The design and operation of these rotors have been discussed in detail by Anderson (32).

Density gradient centrifugation is an invaluable tool at the laboratory level, but prospects for preparing commercial quantities of highly purified seed proteins by this method are dimmed by costs for purchase and operation of the sophisticated equipment needed. There is hope, however, that other methods for large scale separation can be developed as additional knowledge of factors which control the size, shape, uniformity and durability of protein bodies accumulates. The liquid cyclone (33) or other forms of elutriation, centrifugation (34), as well as airclassification in the dry state, deserve exploration.

## **Scanning Electron Microscopy**

The ability to visualize subcellular structure provides unique insight into the properties of cellular components and is extremely valuable in judging the efficiency of separation procedures at the particulate level. In this regard, scanning electron microscopy (35,36) is especially useful since detailed threedimensional images can be obtained (37) with comparatively little effort in sample preparation.

In transmission electron microscopy, electrons that



l~m. 3. Scanning electron micrograph of clumped **protein** bodies from soybean. Source: W. J. Wolf (39).



FIG. 4. Schematic diagram of gel filtration chromatography. Broken circles represent porous particles of column packing. Filled circles are different-sized molecules separated by selective diffusion into column packing.

pass through a carefully prepared sample form a two-dimensional image on a phosphorescent screen. In scanning electron microscopy, transmitted electrons, back-scattered electrons and secondary electrons from the sample surface, which are collected through a detector above the sample, modulate the electron beam in a cathode ray tube display, which is driven in synchrony with the beam of electrons that scan the sample. Thus surfaces nearer the detector appear brighter than surfaces farther away, and the image generated on the display screen acquires a three-dimensional character. Magnification is achieved by increasing the size of the display scan with respect to the sample scan.

Although transmission electron microscopes are capable of better resolution than scanning electron microscopes (ca. 5 A vs. 100 A) (38), recent improvements in methodology (37,38) offer prospects of detecting greater structural detail and, perhaps, of obtaining three-dimensional color micrographs (38). Scanning electron micrographs have already revealed interesting clumping properties and a patterned surface texture in protein bodies from soybeans (Fig. 3) (39) and have illustrated how protein adheres to starch grains in wheat flour  $(40)$ .

## **Filtration and Ultrafiltration**

Protein bodies from soybean vary in size from 2 to 20  $\mu$  (16), and soybean proteins have molecular weights from about 15,000 to more than 600,000. Separating protein bodies from other cellular components or fractionating the proteins themselves by filtration or ultrafiltration is especially appealing be-



FIG. 5. Molecular weight distribution of soybean proteins. Source: W. J. Wolf, Northern Regional Laboratory.

cause inexpensive durable membranes are available in a variety of molecular weight exclusion limits. Commercial membranes range downward from approximately 15  $\mu$  in pore size and upward from a 500 molecular weight (MW) exclusion limit. Although filtration and ultrafiltration methods have not been proved in use with proteins or subcellular particles from plant seeds, several examples of protein separation or concentration by these methods have appeared in the literature (41-45).

Aside from the membrane, the most important feature of commercial ultrafiltration equipment is some means by which material is prevented from accumulating on the membrane, such as a rotating stirrer just above the surface (46) or a spiral chamber through which the solution being filtered is pumped to produce circular flow parallel to the membrane surface (47).

#### **Gel Filtration Chromatography**

Gel filtration is a form of liquid chromatography that has proved invaluable in fractionating a variety of substances. Numerous applications of gel filtration in the fractionation of seed proteins are exemplified by a few studies reported since 1967 (48-52).

In gel filtration, column packings are porous materials and, as illustrated schematically in Figure 4, mixtures of particles that are percolated through the columns are separated on the basis of radius of gyration (53) or hydrodynamic size (54), provided there is no chemical or associative interaction with the column packing. Small molecules, or those whose shapes allow them to penetrate the gel freely, are retarded while larger molecules, or those with unusual shapes that are excluded from the pores of the gel, flow through unhindered.

Commercially available dextran, acrylamide, and agarose gels fractionate materials with molecular weights that range from less than 500 to more than 150 million. Unfortunately, a number of disadvantages, including the fragile nature of low density gels, susceptibility to microbial growth, poor stability in dissociating solvents, and a tendency to bleed, make these gels less than suitable for prolonged use in separating highly aggregated seed proteins under optimum conditions; namely, with small gel particles in long narrow columns under high pressure drops or high flow rates (55). The introduction of stable



FIG. 6. Dissociation and association of soybean llS proteins. Source: W. J. Wolf, Northern Regional Laboratory.

gel-like media promises, however, to make rapid gel filtration of larger molecules and particles a reality. These new materials are, commonly, porous glass beads (56,57) or pellicular materials of controlled surface porosity (58,59).

Beau et aI. (57) separated horse serum globulins from albumin in less than 1 hr at a flow rate of  $240 \text{ ml/hr/cm}^2$  using porous glass beads. By comparison, the optimum flow rate of small pore agarose, which is effective in the same size range, is about 40 ml/hr/cm% Rigid media may, thus, increase the speed of gel filtration separations of larger proteins at least sixfold.

So far porous glass beads have been prepared with average pore diameters ranging from 0.005 to 0.5  $\mu$  (57). Although pores in this size range would exclude intact protein bodies, these glass beads might be useful in purifying smaller particles released when protein bodies are ruptured in water (16).

Pellieular packings consist of uniformly porous sperical shells of various materials including alumina, cellulose, silicie acid, or crosslinked polystyrene on solid microscopic glass beads or polystyrene microspheres. The surfaces of such beads give column packings a texture like other gel filtration media, and yet their solid cores maintain stability under high pressure. The construction of these beads provides an additional advantage in that surface properties can be changed in ways not possible with glass alone.

### **Dissociation of Proteins**

In the ultracentrifuge, water-extractable soybean proteins distribute according to the pattern in Figure 5 (60,61). Their characteristic sedimentation rates are the basis of several proposed nomenclature systems for soybean proteins (62). As in extracts from defatted soybean meal, 7S and 11S globulins also predominate in protein bodies (39). Although these globulins have molecular weights from 100,000 to 400,000, they consist of smaller (63,64), perhaps dissimilar (65), subunits and are capble of undergoing association and dissociation reactions (66). Figure 6 shows conditions under which 11S globulins associate into higher molecular weight aggregates and summarizes the treatments that dissociate the 11S globulins into 7S half-molecules and, ultimately, into 2-3S subunits. In addition to these conditions, the tendency of low temperatures and specific ions (67), particularly phytie acid (68,69), to promote aggregation of the soybean globulins should also be noted.

Figure 7 illustrates how a change in aggregation

that is characteristic of the 7S globulins near neutral pH is observed in an analytical ultracentrifuge (64,70). Similar information on aggregation state can be obtained by applying frontal analysis techniques (71,72) in gel filtration chromatographic experiments.

In terms of their ability to associate or dissociate, the soybean globulins provide excellent examples of the multisubunit structure of large proteins, a more general phenomenon which is described in greater detail in extensive discussions of protein dissociation and association found elsewhere (73-75). The point is illustrated here by data in Table II that summarizes the molecular weight distribution of subunits in 53 different proteins in which polypeptide chains are<br>held together by noncovalent bonds (76). The held together by noncovalent bonds  $(76)$ . majority of these proteins, represented at one extreme by insulin (MW 11,500, two subunits) and at the other extreme by tobacco mosaic virus (MW 40 million, 2130 subunits), have subunits of less than 50,000 MW. Consequently, most large proteins probably have simple structures. Even those proteins that resist dissociation by dispersing agents or changes in pH and ionic strength may only be crosslinked by a few disulfide bonds or other structures like those shown in Figure 8.

As mentioned earlier, soybean globulins form higher order polymers by disulfide interchange reactions (66). Fortunately, from the standpoint of protein fractionation and characterization, disulfide bonds are readily cleaved by reducing agents, and the ambiguities that interchange reactions introduce can be



FIG. 7. Effect of ionic strength on the sedimentation pattern of soybean globulins at pH 7.6. Source: W. J. Wolf (62).



FIG. 8. Examples of intermolecular bonds in natural products. A, disulfide; B, peptide; C, desmosine; D, lysinoalanine; E, lanthionine. Wavy lines represent polypeptide; circles, polysaccharide.

eliminated by forming stable sulfhydryl derivatives with either acrylonitrile (77) or ethylenimine (78).

Other structures, like the peptide erosslinks (Fig. 8B) that occur in bacterial cell walls (79) or desmosine (Fig. 8C), which is present in elastin (80), are not of common concern to plant protein chemists. This situation does not, however, exclude the possibility that analogous structures might exist in seed proteins or that unusual bonds might be formed under extreme processing conditions. Fortuitous synthesis of desmosine, which normally arises from the enzymic condensation of four lysine residues, is unlikely, but the formation of lysinoalanine (Fig. 8D) or lanthioninelike structures (Fig. 8E) is distinctly possible (81,82). It remains to be seen to what extent the properties and nutritional value of seed proteins are affected by such structures.

Conjugated proteins, which contain lipid or carbohydrate, are also of interest because of unusual properties imparted by the nonprotein moieties; increased solubility in organic solvents or water, modified biological activities, altered hydrodynamic properties and varied antigenic character.

Lipoproteins apparently owe their stability to associative forces  $(\overline{83}-85)$  because the lipids can usually be extracted by organic solvents. In glycoproteins, however, carbohydrate and protein are more often joined firmly through covalent bonds (86,87)--either N-glycosides to the side chain amide of asparagine residues (Fig. 9A) or O-glycosidic bonds to hydroxyamino acids (Fig. 9B). The sugar involved in the



FIG. 9. Protein-polysaccharide bonds in glycoproteins. N-glycoside; B, O-glycoside. Wavy lines represent polypeptide,  $R = H$  or polysaccharide.

N-glycoside type is usually N-acetylglucosamine. Common examples of this kind of glycoprotein are secreted proteins from mammalian tissue, soybean heamagglutinin (87) and 7S globulin from soybean (88). The absence of specific chemical methods by which this type of bond can be selectively cleaved has prompted searches for suitable enzymes (89). After further developmental work on these enzymes, the complete removal of carbohydrate from protein preparations may be a simple task.

O-Glycoside-type glycoproteins usually contain a pentose immediately adjacent to the protein. This pentose is often xylose in human mucins and connective tissue glycoproteins (90), but in tomato plant cell walls, the sugar is arabinose (91). Similar structures are apparently present in corn pericarp (92). Where the  $0$ -glycoside-type bond involves serine or threonine, the bond can be cleaved with dilute alkali through  $\beta$ -elimination. Preferably, this cleavage is done under reducing conditions to convert the resulting dehydroamino acids into stable derivatives and to eliminate subsequent condensations with other functional groups or peptide bond cleavage through decomposition of the dehydroamino acids.

## **Immunochemical Techniques**

Descriptions of mammalian protein purifications often emphasize the importance of biological activity





a Based **on data tabulated for 53 different proteins and** 1206 **subunits** (76).





FIG. 10. Schematic diagram of purification of a protein antigen (An) by immunoadsorption onto an immobilized antibody (Ab).

as a property that can be used to judge purity and the efficacy of purification techniques. At the same *time,* in connection with plant seed protein purifications, the absence of natural biological activity is often lamented. One obvious solution to this dilemma would be to identify systems that require specific seed proteins; not an impossible task, but one that could be more formidable than purifying the proteins without knowledge of biological activity. For milk proteins, an obligatory requirement for  $\alpha$ -lactalbumin in a lactose synthetase system (93) was not established until nearly 14 years after the protein was first crystallized (94). An easier way to obtain biological activity is to create it by producing antibodies to the seed proteins. Where suitable antisera can be prepared, immunochemical reactions exhibit all the specificity of natural biologically active systems.

From a practical standpoint, the prospects for rapid protein purifications by highly specific reagents, which function as outlined in Figure 10, are especially intriguing in view of the number of enzymes, antibodies and toxins that have already been immobilized in their active forms (95). When a mixture of protein antigens is passed through a column of an immobilized antibody, proteins for which the antibody is specific are bound; the others are washed on through. The pure antigen can then be eluted from the resin by some dissociating treatment--perhaps a change in pH, salt concentration or temperature. Obviously there are many ways in which the specific properties of immobilized biopolymers can be utilized. Some of the variety is illustrated by the following studies: Catsimpoolas and Meyer (96) used insolubilized 11S soybean protein to purify specific antibodies. Werle and coworkers (97,98) isolated soybean trypsin inhibitors by adsorbing them on trypsin attached to a polymer prepared from maleic anhydride



FIG. 11. Schematic diagram of protein identification by immunodiffusion.

and ethylene. And Agrawal and Goldstein (99) demonstrated that the affinity of concanavalin A for ~-glueans facilitates its purification by adsorption onto crosslinked dextrans.

Antibodies already find general use in immunochemical techniques. The utility of immunochemical methods in studies on proteins has been described in detail by Grabar (100). Applications in soybean chemistry are illustrated by the work of Catsimpoolas et al. (96,101-103). Figure 11 shows how purity or identity can be established by immunodiffusion in agar gel. If a solution of antibodies to a whole protein extract is placed in a well cut into the center of a slab of gel in a petri dish, the antibodies will diffuse into the gel at a rate controlled by the density of the gel, the sizes and shapes of the proteins, and their concentrations. Similarly, protein antigens from the extract, placed in wells located equidistant from the center well, will also diffuse into the gel. An opaque preeipitin line forms in the gel where a protein antigen and its corresponding antibody meet at a concentration ratio of about 2:1 (antigenantibody). If two different fractions contain the same antigen, the precipitin line is continuous but where there are two different protein antigens, crosses or spurs are obtained. This method has been especially useful with sample mixtures, but it is often difficult to differentiate between several proteins that have similar gel diffusion characteristics. Resolution can be improved by combining immunodiffusion with other separation techniques; immunoelectrophoresis is the classic example.

In immunoeleetrophoresis, a mixture of proteins is first allowed to migrate in one direction in a gel under the influence of an electric potential. Then, immunodiffusion is carried out at right angles to the direction of electrophoresis by placing antiserum in a trough cut along the center or sides of the gel. Figure 12 illustrates the use of immunoelectrophoresis in the purification of soybean lipoxygenase (E.C. 1.13.1.13) (104). The pattern obtained with the purified enzyme is devoid of the extra precipitin lines present in the pattern obtained with the starting mixture of whey proteins.

An even more interesting aspect of this particular work is the isoelectric focusing method by which the enzyme was separated.



FIG. 12. Immunoelectrophoresis of soybean whey proteins A) and isolated soybean lipoxygenase (B). Source: N. Catsimpoolas (104). Copyright, Academic Press, Inc.

# Isoelectric Focusing

Isoelectric focusing, or electrofocusing, is a comparatively new and novel technique for resolving proteins on the basis of differences between their isoelectric points  $(105,106)$ . It depends upon the ampholytie properties of molecules that contain both basic and acidic groups. It became a practical reality with the development of synthetic low-molecularweight ampholytes that have isoelectric points throughout the pH scale from 3 to 10 (107).

Application of an electric potential across an ampholyte solution placed between, and in contact with, an electrode compartment containing alkali and one containing dilute mineral acid (Fig. 13) causes each ampholyte molecule to migrate toward one of the electrodes until it comes into contact with other molecules that are stronger acids or bases; thereupon, it either donates or acquires a proton and loses its charge. Conveniently, when all the ampholytes lose their net charge, no current flows and the operator merely needs to monitor current flow to determine when an electrofocusing run is complete.

Even though isoelectric-focused ampholytes carry



FIG. 13. Schematic illustration of isoelectric focusing technique **for protein** separation according to differences in **isoe]cctrlc** point.



FIG. 14. Isoelectrie focusing of a commercial lipoxygenase  $(A)$ , soybean whey proteins  $(\tilde{B})$ , and isolated soybean lipoxygenase (C) in polyaerylamide gel. Source: N. Catsimpoolas (104). Copyright, Academic Press, Inc.

no net charge, they remain ionized, and a pH gradient is thus established throughout the ampholyte solution. The slope of this gradient with respect to distance between the electrode chambers can be varied by altering the composition of the ampholyte mixture. Mixtures that represent narrower ranges of isoelectric points produce more shallow pH gradients and greater physical separation of the ampholytes. Thus two proteins that have isoelectric points at pH 4.4 and 4.5 might be confined to a single band if electrofocused in a solution of ampholytes with isoeleetric points from pH 3 to 10, but they should be separated nicely if electrofocused in the same apparatus with ampholytes that have isoelectric points from pH 4.0 to 5.0. In actual practice, electrofocusing can be used to separate proteins whose isoelectric points differ by as little as  $0.02$  pH unit (108). Simplicity of technique and apparatus makes eleetrofocusing the easiest way to estimate isoelectric points of proteins.

The design of electrofocusing apparatus can be varied widely as long as convection is controlled. In commercial equipment (108), which employs a vertical tubular separation chamber, a sucrose gradient keeps the electrode solutions separated, stabilizes the medium against density changes during electrofocusing, and provides stability while the chamber is drained. Convection due to heating is controlled by water cooling at the inner and outer surfaces of the separation chamber. Simpler designs also utilize a density gradient but do not include elaborate arrangements for cooling (105,109).

Isoelectric focusing in other media promises to be even more useful than electrofoeusing in solutions, especially for microanalytieal work (110,111). Gel electrofocusing allows rapid analysis of numerous samples and, in addition, rivals the resolving power and flexibility of similar methods--zone electrophoresis (112) and disc electrophoresis (113,114). Figure 14 shows examples of the results that may be expected of electrofoeusing in polyaerylamide gel (104).

# **Characterization of Proteins**

Improvements in protein processing or departures from traditional protein uses require knowledge of how proteins interact with other materials and how they respond to physical and chemical forces and structural modifications. Useful information on these subjects can be obtained by random testing, but more fundamental knowledge is needed to ensure that predictions of protein behavior will be accurate and that opportunities for novel applications will not be overlooked. To gain insight not provided by empirical knowledge, the plant protein chemist can turn to a substantial technology and elaborate instrumentation that has been developed to study the fundamental behavior of mammalian biopolymers. Some of these methods have already been discussed; others are listed in Table III along with references that either describe the techniques or give examples of applications. A comprehensive discussion of each method is beyond the scope of this article.

With respect to oil seed proteins, the studies on soybean 7S and IIS globulins reported by Fukushima (146) illustrate how a variety of chemical and physical methods can be used to characterize the major structural features of proteins and the forces that stabilize these structures. Studies like these are particularly useful if they help disclose the details of chemical mechanisms by which changes occur in proteins. Unfortunately, methods like sedimentation, even though improved by advances in theory and instrumentation (147-149), only provide estimates of average molecular behavior over long periods. Techniques that utilize light energy and electromagnetic energy are better suited to instantaneous measurements which yield the desired detailed data.

Protein molecules or amino acids within proteins absorb energy in ways that are influenced by neighboring molecules or atoms. Changes in the size, shape and character of protein molecules or specific sites on proteins can be followed rapidly by monitoring changes in the ways that they absorb or release

energy. Unusual interactions of tyrosine or tryptophan residues with surrounding groups or solvent molecules are readily detected through spectral shifts or changes in the intensity of UV light absorption. Similarly, correlations of the amide-II band in  $IR$ spectra of proteins are useful in estimating the amount and type of beta-structure in proteins. This same absorption band, as well as others in the IR, also allows rapid measurement of deuterium exchange in studies that yield information on protein conformations which expose peptide bonds to solvent.

Optical rotatory dispersion (ORD) studies give values that relate optical activity at absorption bands to known protein conformations, particularly helical structures. The companion technique, circular dichroism, is analogous to ORD but more sensitive because it measures absorption differences between right and left polarized light rather than merely the rotation of polarized light. Application of these methods has fostered the concept that proteins can have characteristic melting points, i.e., temperatures at which they undergo extensive and rapid conformational changes.

Transition temperatures can also be located by differential thermal analysis. This method provides the means to directly measure energy involved in conformational changes or decomposition, information that should be especially valuable in predicting the behavior of proteins during baking, cooking or other forms of processing which involve direct heat energy transfer. Descriptions of proteins in terms of funda: mental thermodynamic properties should also help in predicting their solution chemistries and the properties of chemical derivatives.

In fluorescence studies on proteins, light energy absorbed at one wavelength is emitted at another wavelength. Because the character and intensity of emitted light depend upon intervening events within the molecules, differences between incident and emitted light reflect a variety of molecular interactions. These are characterized in terms of excitation spectra, emission spectra, quantum yield, fluorescence quenching, relaxa-

Methods for Characterization of Proteins			
Method	Condition оf sample	Property measured	Reference
Titration	Solution	Hydrogen bonding	115
Hydrogen exchange	Solution	Group exposure	116
Enzymic digestion	Solution	Character and geometry	117
		of specific regions	
Reaction with bifunc-	Solution	Character and geometry	118
tional reagents		of specific regions	
Dve binding	Solution	Character and geometry	$119 - 121$
		of specific regions	
Sedimentation	Solution	Size and shape	12
	Solution	Size and shape	12
Light scattering	Solution	Size and shape	12
Viscosity measurements	Solution	Size and shape	122
Osmometry	Solution	Size and shape	123
Dialysis and diffusion	Solution	Group interaction	$124 - 126$
UV spectroscopy		Pelypeptide conformation	
IR spectroscopy	Solution-solid	Polypeptide conformation (helix-beta-structure)	125
		Hydrogen bonding	
Optical rotatory dispersion	Solution	Polypeptide conformation	14,125,127
		(helix-coil)	
Circular dichroism	Solution	Polypeptide conformation	128
		(helix-coil)	
Fluorescence measurements	Solution	Size and shape	125,129-132
		Group interaction	
		Protein-protein interaction	
Electron paramagnetic	Solution-solid	Character and geometry	133-136
resonance spectrometry		of specific regions	
(spin labeling)		Size and shape	
Nuclear magnetic resonance	Solution-solid	Group interaction	$137 - 139$
spectrometry			
(pulse labeling)			
Differential thermal	Solution-solid	Energy of conformational	$140 - 141$
analysis		changes	
X-ray crystallography	Solid	Complete structure and	$142 - 144$
		polypeptide conformation	
Sequence analysis	Solution-solid	Complete structure	145

TABLE III

tion time, polarization of fluorescence or rotational relaxation time depending upon the type of energy transition involved. Natural fluorescence of proteins is primarily due to tryptophan and, to a lesser extent, to tyrosine (131). Where natural fluorescence is unsuitable, proteins can be modified with fluorescent dyes (150) to obtain better information on protein-protein or protein-solvent interactions and on conformational changes within specific regions in the proteins.

Information like that furnished by fluorescence measurements can also be obtained by analysis of the ways in which proteins respond to electromagnetic energy. Electron paramagnetic resonance spectroscopy (EPR), which utilizes microwave energy, probes both the interior and exterior of proteins through small paramagnetic substituents on the proteins. Nuclear magnetic resonance spectroscopy (NMR), which depends on radiowave energy, may give even more valid data on protein conformation changes because it can be used in pulse-labeling experiments without modifying the proteins. Both EPR and NMR are advantageous in that they can be applied to solid materials as well as to solutions. NMR studies on fragments from red blood cells, for example, have produced direct experimental evidence of the types of chemical bonds involved in the interactions of lipids with proteins (151).

Methods of amino acid sequence analysis are uniquely important in the characterization of proteins because covalent polypeptide structure is the principal constraint on protein reactivity. With knowledge of primary amino acid sequence, it is conceivable that the chemistry of proteins and their derivatives might be predicted instantaneously from the behavior of computer-constructed models (5-8). Already computer technology allows the chemist to perturb electronic models and visually observe their responses (152) with convenience that far exceeds present laboratory experimentation. Increased computer capabilities and refinements in knowledge of secondary forces that constrain proteins may make it possible to discard many physical methods now used in the characterization of proteins.

Perhaps the ultimate means by which protein structure may be determined at present is X-ray crystallography. It provides information on peptide chain conformation as well as amino acid sequence. Much of the drudgery originally connected with recording and measuring reflections and constructing electron density maps has been eliminated by automation. A requirement for crystalline samples is the greatest problem remaining. Many times the lack of suitable forms has completely prevented X-ray analysis of protein structure. Sequence analysis by chemical means does not share this shortcoming, but it still requires a reasonable degree of purity. With the continuing introduction of labor-saving devices (153, 154), chemical sequence analysis promises to become an extremely valuable analytical tool in the protein laboratory.

#### ACKNOWLEDGMENTS

Tokuji Watanabe and Kyoko Saio, Ministry of Agriculture and<br>Forestry, Tokyo, Japan, Nicholas Catsimpoolas, Central Soya, Chicago,<br>Ill., and W. J. Wolf and Donald Christianson, Northern Regional<br>Laboratory, provided use of

#### REFERENCES

- 
- 
- 1. Sanger, F., and H. Tuppy, Biochem. J. 49, 481 (1951).<br>2. Gutte, B., and R. B. Merrifield, J. Amer. Chem. Soc. 91,<br>501 (1969).<br>3. Hirschmann, R., R. F. Nutt, D. F. Veber, R. A. Vitali, S. L.<br>Varga, T. A. Jacob, F. W. Hol
- 
- 5. Stanachard, O. N. and V. Saistehara, Disk 2, 283 (1968).<br>1832, 283 (1968).<br>1845 And V. And Stan (1971).<br>185 And V. And V. A
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 230 (1969).<br>
12. Expany, Expand B. K. Seavey, Endrocrinology<br>
12. H. Evis, U. J., E. V. Cheever and B. K. Seavey, Endrocrinology<br>
12. Hand, Ni (1967).<br>
13. Hand, Ni Robins As Al (1967).<br>
13. Hand, Ni Robins As Al (1967).<br>
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 70. Koshiyama, I., Did. 32, 879 (1968).<br>T. Stone, M. J., and H. Metzger, J. Biol. Chem. 243, 5049 (1968).<br>T. Ronalds, J. A., and D. J. Winzor, Arch. Biochem. Biophys.<br>T. Regiss 94, 456 (1969).<br>T. Reithel, F. J., Advan. Pro
- 
- 
- 
- 
- 
- 
- 
- 82. Horn, M. J., D. B. Jones and S. J. Ringel, J. Biol. Chem.<br>188, 141 (1941).<br>
83. Vandenheuvel, F. A., JAOCS 43, 258 (1966).<br>
84. Camejo, G. Biochemistry 6, 3228 (1967).<br>
85. Burton, R. M., Chairman, Symposium: "Lipid Mo
- 
- 
- 89. Tarentino, A. L., and F. Maley, Ibid. *18o,* 295 (1969). 90. Bella, A., Jr., and I, Danishefsky, J. Biol. Chem. 248, 2660
- (1968).
- 
- 91. Lamport, D. T. A., Biochemistry 3, 1155 (1969).<br>
92. Boundy, J. A., J. S. Wall, J. E. Turner, J. H. Woychik and<br>
R. J. Dimler, J. Biol. Chem. 242, 2410 (1967).<br>
93. Brodbeck, U., W. L. Denton, N. Tanahashi and K. E. E
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 97. Fritz, H., M. Gebhardt, E. Fink, W. Schramm and E. Werle,<br>
98. Hochstrasser, K., and E. Werle, 1969).<br>
98. Hochstrasser, K., and E. Werle, 1969).<br>
99. Agraval, B. B. L., and E. Werle, 1969).<br>
99. Agraval, B. B. L., and
- 
- 
- 
- 
- 
- 119. Wold, F., Bitunctional Reagents, Ibid.<br>
119. Steinhardt, J., and S. Beychok, Interaction of Proteins with<br>
Hydrogen Ions and Other Small Ions and Molecules, in "The<br>
Proteins, Yen Molecules and Other Small Ions and M
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 134. Hsia, J. C., and L. H. Piette, Arch. Biochem. Biophys. 129,<br>296 (1969).<br>135. Landgraf, W. C., and G. Inesi, Ibid. 130, 111 (1969).<br>136. Landgraf, W. C., and W. H. Orme-Johnson, Ann. N.Y. Acad. Sci.<br>155, 336 (1969).<br>13
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 140. Green, D. B., F. Happey and B. M. Watson, Differential Thermal<br>
Analysis Applied to Polypeptides, in "Conformation of Bio-<br>
polymers," Edited by G. N. Ramachandran, Vol. 2, Academic<br>
Theres, New York, 1967.<br>
141. Cri
- 
- 
- - [Received December 5, 1969]