Crystallization of Biological Macromolecules

CRYSTALLIZATION OF BIOLOGICAL MACROMOLECULES

A. McPherson
University of California, Irvine

COLD SPRING HARBOR LABORATORY PRESS
Cold Spring Harbor, New York

Printed in the United States of America

Project Coordinator: Joan Ebert
Production Editor: Christopher Bianco
Cover Design: Tony Urgo

Front cover: An orthorhombic crystal of porcine pancreatic α-amylase emerges from the surface of a kinetically favored polymorph. The growth frames on its face provide a record of thermal variation during development.

Library of Congress Cataloging-in-Publication Data

McPherson, Alexander, 1944-
Crystallization of biological macromolecules/Alexander McPherson.
p. cm.
Includes bibliographical references and index.
ISBN 0-87969-527-7 (hardcover)
1. Biopolymers. 2. Macromolecules. 3. Crystallization.
I. Title.
QD381.M69 1998
547'.7—dc21

98-26897
CIP

CRYSTALLIZATION OF BIOLOGICAL MACROMOLECULES

Preface

Rarely has a technique so thoroughly transformed a major field of research as X-ray crystallography has altered the course of molecular biology. From sequences of nucleotides on strands of DNA, we are, because of this method, now poised to visualize their ultimate products in atomic detail. A veritable revolution in biochemical research has emerged from three principal advances in the technology. The first was the advent and development of recombinant DNA methods. These permitted the expression and purification of vast amounts of otherwise unobtainable proteins and nucleic acids. The second was the unprecedented progress in X-ray diffraction instruments and procedures that radically reduced the time and effort required for structural analysis. The third, and also the subject of this book, was the development of systematic approaches for the crystallization of biological macromolecules. This technology of macromolecular crystal growth, still in its formative stages, supplies an essential bridge between recombinant DNA and X-ray diffraction analysis. Crystals of macromolecules have become keystones of modern biological science.

Crystallization, as applied to proteins, was long held in high esteem, but hardly considered science. Described in terms of "black art," "alchemy," and worse, it was relegated to the farthest back benches of the biochemistry laboratory. Because the molecules involved exhibited such perverse behavior and were so poorly characterized, conducting serious research into their crystallization was conceded as hopeless. Only when the demand for crystals by the diffractionists, and later the molecular biologists, could no longer be ignored did researchers pursue in earnest the study of macromolecular crystal growth.

The past quarter century of effort, however, has produced an appreciation that protein, nucleic acid, and virus crystal growth are indeed approachable with the traditional intellectual instruments of science. An equally important outcome, and one of more practical consequence, has

been the development of tactics, strategies, reagents, and devices that have enormously raised our expectations of crystallizing virtually all macromolecules.

From our store of shared experiences, accumulated wisdom, and tabulated data, have come rapid and efficient screening procedures and detailed recipes. Commercial kits for macromolecular crystallization are readily available, as is an expanded set of useful reagents. These have significantly broadened the community of active investigators and practitioners, extended the range of macromolecules under study, and dramatically accelerated the determination of new biological structures.

Further enhancing the impact of crystallography has been the advent of synchrotron X-ray sources of ultra-high intensity coupled with sensitive and fast X-ray detectors. These make data collection assured and rapid. In addition, these new instruments have reduced the size of crystals required for analyses to less than 50 μm on an edge, and even this measure is declining. Cryocrystallography has also had profound consequences. No longer do we need dozens of crystals to obtain a set of diffraction data, one crystal will suffice. With new phasing methods, only one set of data, hence one crystal, may be sufficient for an entire structure determination. The ultimate objective may become obtaining that single, small, but perfect crystal.

It is tempting to focus exclusively on X-ray crystallography and its successes, both realized and projected, but it is important to appreciate that macromolecular crystallization is also of value for entirely separate reasons. Protein crystals are utilized for downstream processing and purification in the industrial production of enzymes, they are involved in pharmaceutical formulation, and they are often useful in various other areas of biophysical experimentation. Of particular note is their significance in the broader field of crystal growth research. Modern techniques, such as atomic force microscopy, have shown that macromolecular crystals employ the same mechanisms for growth as do conventional crystals. Because of the much larger size of the molecules, however, the kinetics of growth are substantially attenuated. As a consequence, direct observation of nucleation and growth processes are possible that would otherwise elude detection with conventional crystals. Thus, macromolecular crystallization now serves as the best model system for general crystallization from solution.

The principal objectives of this book are to provide a comprehensive biochemical context for pursuing the crystallization of macromolecules, to instruct the reader in the practical aspects of the technology, and to lay out effective strategies for success. For consistent achievement, however, one must be more thoroughly informed. To this end, some underlying

physical and chemical principles are presented in simplified form. Illustrations are presented, both for edification and inspiration. Finally, an attempt has been made to describe the most recent advances in macromolecular crystal growth, those that have facilitated the transformation of the enterprise from "black art" to science.

This book would have been impossible without the contributions of a multitude of colleagues too vast to identify individually here. In a sense, though, the material included in this volume has thousands of authors, those who have, over the past 200 years, crystallized macromolecules and troubled themselves to record or discuss their experiences. I have only gathered and condensed their thoughts, observations, and ideas.

Several individuals, however, deserve particular recognition and thanks. They have made unusual contributions. Chief among them are Drs. Alexander Malkin and Yurii G. Kuznetsov, who taught me most of what I know regarding the physics of crystal growth, and who recorded with their own hands the atomic force micrographs found throughout this book. Thanks also goes to my able technician and friend, John Day, who carried out many of the crystallization experiments in my laboratory. Many of the illustrations were produced by Aaron Greenwood who worked tirelessly and with admirable patience. Next to myself, Debora Felix was most responsible for the organization and composition of this volume, some might argue more responsible. Without her ceaseless attention and hard work, there certainly would not have been a book.

Finally, I wish to thank the individuals of the Cold Spring Harbor Laboratory Press, especially Patricia Barker, Chris Bianco, and Joan Ebert, for their unyielding insistence on accuracy, precision, and quality in all that we attempted.

A. McPherson

Contents

1

The History and Character of Macromolecular Crystals

The history of macromolecular crystal growth extends more than 150 years (McPherson 1991). In addition to many of the principles that form the foundation of our current methods and techniques, it is replete with lessons that are still of value to us today. It is rich in methods forgotten or now ignored, and at the same time it shows that many of our modern approaches are only reinventions. The literature is filled with tricks and quirks, myths and facts that illustrate the subtleties, frustrations, and vagaries of the crystal growth process as it pertains to macromolecules.

Pioneers in the field generally were compelled, by the primitive techniques available to them, to work exclusively with proteins that could be obtained in very large amounts. Thus, we find descriptions of proteins and attendant procedures that can be useful now for the large-scale investigation of crystallization phenomena. Proteins crystallized near the turn of the century may, in fact, provide some of the best systems available for light-scattering experiments, growth-rate measurements, calorimetry, and the evaluation of new crystallization techniques and reagents.

This concise but limited review of the history of protein crystal growth, summarized in Table 1.1, may serve as an introduction to the literature and may provide starting points for those who intend a deeper exploration. Perhaps it will also suggest some useful protein systems for those who wish to put the instruction, theory, and practice of protein crystallization, found elsewhere in this book, to practical use.

THE FIRST PROTEIN CRYSTAL: HEMOGLOBIN

The first published observation of the crystallization of a protein was authored by Hünefeld in 1840. The protein was hemoglobin from the earthworm. This iron-containing protein was obtained as flat, platelike crystals when the blood of an earthworm was pressed between two slides

Table 1.1 A chronology of protein crystal growth

Protein	Investigator	Date
Hemoglobin	Hünefeld	1840
	Reichert	1849
	Leydig	1849
	Kölliker	1849
	Budge	1850
	Fünke	1851
	Lehmann	1853
	Pasteur	1863
Excelsin (reserve protein) from Brazil nut	Hartig	1855
	Maschke	1858
Globulin (reserve protein) from		
cocoa nut (cocosin)	Ritthausen	1880
castor bean	Ritthausen	1881
hemp seed (edestin)	Ritthausen	1881
sesame seed	Ritthausen	1881
squash seed (curcurbitin)	Grübler	1881
oat kernel (avenalin)	Osborne	1892
flax seed	Osborne	1894
kidney bean (phaseolin)	Osborne	1894
Hen egg albumin	Grübler	1881
	Hofmeister	1890
	Hopkins and Pincus	1898
	Wichmann	1899
	Sørensen and Hoyrup	1917
Concanavalin A (jack bean)	Sumner	1919
Concanavalin B (jack bean)	Sumner	1919
Urease (jack bean)	Sumner	1926
Insulin (pig)	Abel et al.	1927
Trypsin (beef)	Northrop and Kunitz	1931
Chymotrypsin (beef)	Kunitz and Northrop	1933
β-Lactalbumin (cow milk)	Palmer	1934
Pepsin (beef)	Northrop	1934
Old yellow enzyme (yeast)	Theorell	1934
Tobacco mosaic virus	Stanley	1935
Carboxypeptidase (beef pancreas)	Anson	1935
Chymotrypsinogen (beef)	Kunitz and Northrop	1935
Trypsin inhibitor (beef pancreas)	Kunitz and Northrop	1936
Alcohol dehydrogenase (yeast)	Negelein and Wulff	1936
Pepsinogen (beef)	Herriott and Northrop	1936
Trypsinogen (beef)	Kunitz and Northrop	1936
Catalase (beef liver)	Sumner and Dounce	1937
Papain (papaya latex)	Balls et al.	1937
Lysozyme (hen egg)	Abraham and Robinson	1937
Acetaldehyde reductase (yeast)	Negelein and Wulff	1937
Ficin (fig latex)	Walti	1937

(continued on next page)

Table 1.1 (Continued)

Protein	Investigator	Date
Tomato bushy stunt virus	Bawden and Pirie	1937
Ferritin (horse spleen)	Laufberger	1937
Tyrosinase	Dalton and Nelson	1938
Lecitinase (snake venom)	Slotta and Fraenkel-Conrat	1938
Hemocuprein and hepatocuprein (beef)	Mann and Keilin	1938
Pepsinogen (beef)	Herriott	1938
Glyceraldehyde 3 phosphate dehydrogenase (rabbit muscle)	Baranowski	1939
Serum albumin (horse)	McMeekin	1939
RNase (beef pancreas)	Kunitz	1939
Lactate dehydrogenase (beef heart)	Straub	1940
Enolase (yeast)	Warburg and Christian	1941
Fumarase (yeast)	Laki and Laki	1941
Diptheria antitoxin (human)	Northrop	1941
Pepsin inhibitor (beef pancreas)	Herriott	1941
Chymopapain (papaya latex)	Jansen and Balls	1941
Bence Jones protein (human urine)	Mall and Bersin	1941
Phosphorylase (rabbit muscle)	Green et al.	1942
Phosphate transmitting enzyme (yeast)	Bücher	1942
Catalase (sheep and beef)	Dounce	1942
Carbonic anhydrase (human)	Scott	1942
Asclepain (milkweed)	Carpenter and Lovelace	1943
Aldolase (rat)	Warburg and Christian	1943
Rennin (beef)	Berridge	1943
Rennin (beef)	Hankinson	1943
Lactoperoxidase (cow milk)	Theorell and Paul	1944
Serum mucoprotein	Bader et al.	1944
Alkaline phosphatase	Thoai et al.	1944
Southern bean mosaic virus	Price	1945
Lysozyme (hen egg)	Alderton and Fevold	1946
β-Amylase (sweet potatoes)	Balls et al.	1946
Trypsin inhibitor (soybean)	Kunitz	1946
Hexokinase (yeast)	Kunitz and McDonald	1946
Turnip yellow mosaic virus	Markham and Smith	1946
Trypsin-soybean trypsin inhibitor complex	Kunitz	1947
α-Amylase (pig pancreas)	Meyer et al.	1947

of glass and allowed to dry very slowly. In the same paper, Hünefeld refers to, but does not further describe, hemoglobin crystals (or blood crystals as they were called) from man and pig.

The observation of Hünefeld, simple though it was, demonstrated several important points. It suggested for the first time that protein crystals could be obtained by the controlled evaporation of a concentrated protein

solution; that is, protein crystals could be produced by slow dehydration. This idea, generally applied in more subtle forms, is the basis for most of the techniques we use today. The observation of hemoglobin crystals, the original similar to those seen in Figure 1.1A, suggested that proteins form crystals in much the same way as do conventional small molecules. It further demonstrated that protein crystals could, in at least some cases, be grown from relatively crude physiological preparations and that a high degree of purification was not always essential.

The first serious study of the appearance of hemoglobin crystals, however, was published by K.E. Reichert in 1849. He investigated crystals from the blood of fetal guinea pigs. The scientific community was clearly interested in the phenomenon of hemoglobin crystallization, and other reports appeared in rapid succession by Leydig (1849) and Kölliker (1849). They focused their attention on blood crystals from dog, river perch, and python. In 1850, Budge observed crystals of human hemoglobin in the stomachs of leeches (Budge et al. 1850). Other names cited in the early literature but whose papers have not yet been located are Teichman and Milne-Edwards.

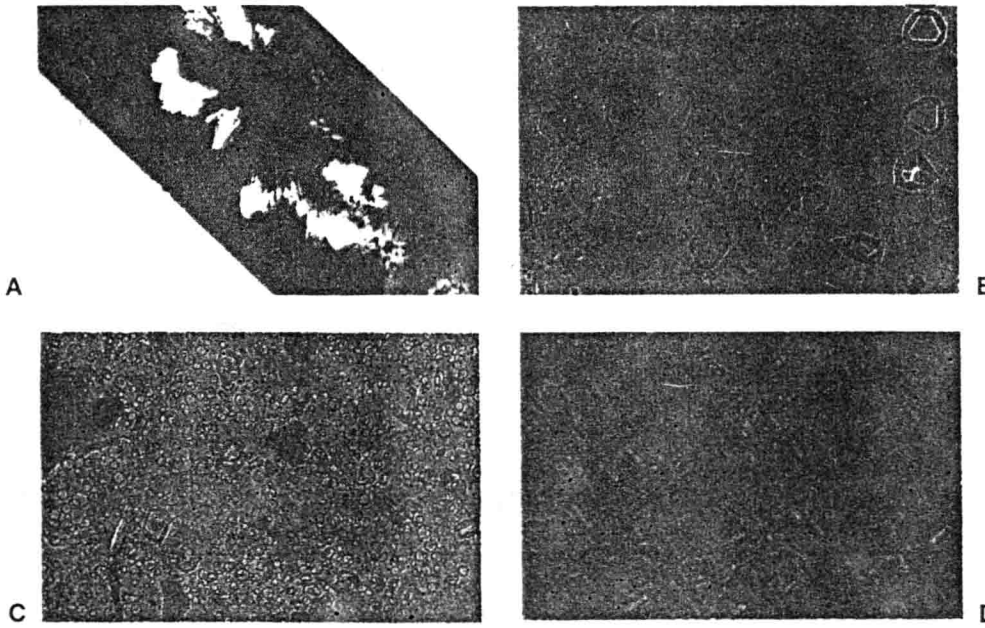

Figure 1.1. (*A*) A cluster of fine needle crystals of hemoglobin grown in a capillary. They are much like those first seen by Hünefeld in 1840. (*B*) Crystals of the storage protein excelsin from the Brazil nut, the first plant protein to be crystallized. (*C, D*) Crystals of globulin from squash and serum albumin from horse, respectively, both among the first proteins crystallized.

Through 1850, all of the blood crystals reported were observed to have grown more or less fortuitously, and no investigator had suggested any general procedure for their directed growth. The first person to actually devise successful and reproducible methods for the growth of hemoglobin crystals was O. Fünke, who began publishing a series of articles on the purposeful growth of hemoglobin crystals in 1851–1852. He described the in vitro crystallization of hemoglobin from the blood of man, horse, bullock, dog, fish, cat, pig, and pigeon. Fünke may be considered the first real student of protein crystal growth and the scientific ancestor of those of us who currently pursue this interest.

One of Fünke's methods was to successively dilute blood corpuscles with pure water, then slowly evaporate the liberated protein solution. His alternate method was based on mixing blood corpuscles with a variety of alcohol or ether solutions. This was the first example of the crystallization of proteins using organic solvents, a method still in use today. In 1853, a textbook was published by C.C. Lehmann, who extensively discussed the crystallization of blood components.

Between 1900 and 1909, E.T. Reichert, a physiologist, with the assistance of A.P. Brown, a mineralogist, made an extensive investigation of hemoglobin crystal growth and thoroughly explored methods for obtaining the crystals from the blood of several hundred different animals. Their thesis was that the outward form of a crystal reflected the composition and structure of the molecules that composed it, and that hemoglobin from different species of animals would contain differences of composition and molecular structure commensurate with their phylogenetic differences. Thus, it would follow that the kingdom of animals might be evolutionarily arranged according to the similarities and differences of the crystals their hemoglobins produced.

Although this theory may appear naive and even amusing to us today, it did, in fact, contain some important elements that are included in current genetic analyses of evolutionary relationships. Perhaps, however, the most lasting value of their extensive research and almost heroic efforts is the extraordinary volume that they produced in 1909 entitled *The Differentiation and Specificity of Corresponding Proteins and Other Vital Substances in Relation to Biological Classification and Organic Evolution: The Crystallography of Hemoglobins.* In this book of nearly 500 pages are found a detailed history of the growth of hemoglobin crystals, innumerable formulas and procedures for their preparation, line drawings and diagrams of crystals from representatives of every genus, and, most impressive of all, more than 600 light microscope photographs of hemoglobin crystals that Reichert grew in his laboratory. This book still

represents the most comprehensive and thorough account of an investigator's efforts to grow crystals of a protein.

THE RESERVE PROTEINS OF PLANT SEEDS

Following hemoglobin, the next proteins to be seriously investigated in terms of their crystallization behavior were the plant seed reserve proteins, principally the globulins. In 1855 and 1856, Hartig published a series of papers detailing his extensive investigation of plant seeds, and he reported the observation of crystals of reserve protein in the cells of the seeds. It was Maschke in 1858 and 1859, however, who succeeded in extracting the reserve protein of the Brazil nut, excelsin, and crystallizing it in his laboratory. Crystals of excelsin, and the same protein from other plant seeds, like those described by Maschke are seen in Figure 1.1B and C.

This work was extended and refined by Grübler (1881), Ritthausen (1880, 1881), and Osborne (1891, 1892, 1894, 1899, 1924) at the turn of the century. Of particular interest to us today are the proteins that they crystallized and the procedures that they employed, for indeed, the author's laboratory has reproduced most of the experiments of the pioneers and found that they yielded, in almost every case, results comparable to those claimed by the original researchers.

The methods that were developed to crystallize the plant seed proteins included: (1) extraction of proteins into salt solutions, generally about 1 M NaCl at 60°C followed by slow cooling to 20°C; (2) exhaustive dialysis of salt solution extracts of the seeds against distilled water; and (3) treatment of protein solutions with alcohol, acetone, or ether.

In these procedures, we find for the first time the exploitation of several approaches now in common use: temperature variation under otherwise constant conditions; dialysis against low ionic strength, thereby taking advantage of the salting in property of many proteins; and a further use of organic solvents as precipitating or crystallizing agents. In Osborne's 1924 monograph on the plant seed proteins, he describes one of the most fundamental procedures utilized by protein chemists during the early part of this century: "The globulins of plant seeds can be obtained as crystals by diluting their sodium chloride solutions with water heated to 50°C to 60°C until a slight turbidity forms. After warming the diluted solution until this turbidity disappears, and allowing it to cool slowly, the protein separates in well developed crystals."

The plant kingdom provided not only protein crystals, but the first crystals of intact viruses as well. Tobacco mosaic virus (TMV), a rod-shaped virus, was obtained in paracrystalline form by Stanley in 1935 and stands

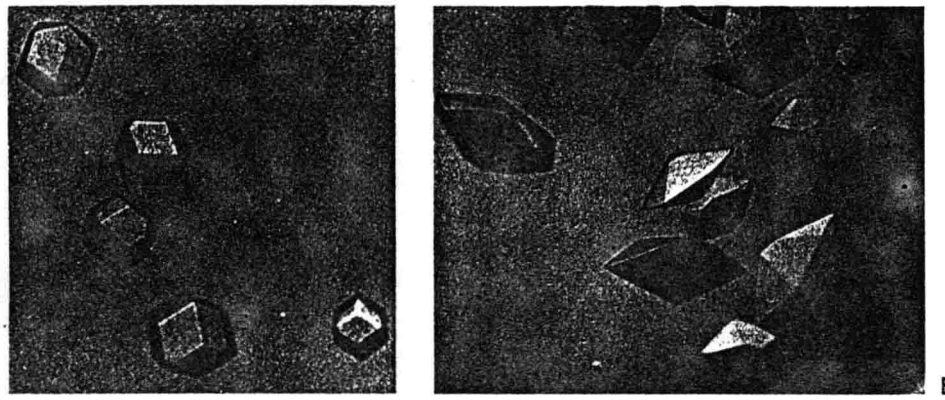

Figure 1.2. (*A*) Crystals of tomato bushy stunt virus. (*B*) Crystals of turnip yellow mosaic virus. (Courtesy of M. Canady.)

even now as a milestone in virology. The first crystals of a spherical virus were grown by Bawden and Pirie and reported in 1937. The beautiful cuboidal crystals illustrated in their paper, and shown as well in Figure 1.2A, were of tomato bushey stunt virus (TBSB), a small icosahedral virus of T = 3 symmetry that later became one of the first virus crystals studied by the fledgling science of X-ray diffraction analysis (Bernal et al. 1938). Another early example of plant virus crystals, those of turnip yellow mosaic virus (TYMV), first crystallized in 1946 (Markham and Smith 1946) are shown in Figure 1.2B.

THE ALBUMINS

At almost the same time that the work with plant seed proteins was carried out, similar efforts were under way to crystallize two animal proteins, hen egg albumin and horse serum albumin. Although a definitive description of procedures for the crystallization of ovalbumin was not published until 1898 by Hopkins and Pinkus, the growth of egg albumin crystals in the laboratory, like those seen in Figure 1.3, was well studied and had been reported by Hofmeister in 1890. In their 1898 paper, Hopkins and Pinkus refer obliquely and imprecisely to the earlier crystallization of horse serum albumin by Grübler, but no details were given. A second, useful discussion of methods for the crystallization of hen egg albumin was provided by Wichmann in 1899.

The first detailed procedures for the crystallization of horse serum albumin crystals, some of which are seen in Figure 1D (grown by the procedures of McMeekin 1939), were not published until 1915–1917 by Sørensen and Hoyrup at the Carlsberg Laboratories. The crystallization

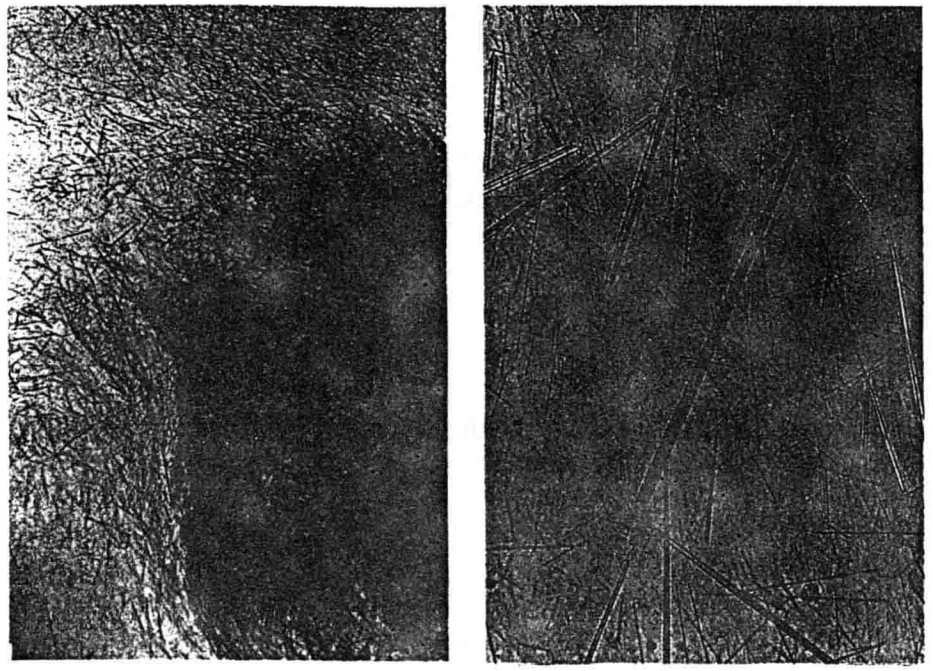

Figure 1.3. Masses of fine hen egg albumin (ovalbumin) crystals seen in *A* are magnified in *B* showing their needlelike character. This was, after hemoglobin, the second protein of animal origin to be grown in the laboratory. It was, in addition, one of the first glycoproteins crystallized.

procedures described for both egg and horse serum albumin were the same, in general terms, although the specifics such as salt concentration varied. The proteins were crystallized by the addition of ammonium sulfate or magnesium sulfate until incipient turbidity, followed by adjustment of the pH to between 4.5 and 6.0 with acetic acid. In both cases, masses of needle crystals were obtained. By using similar techniques, lactalbumin from skim milk was later crystallized by Palmer in 1934.

Procedures for crystallizing albumins took advantage of proposals by Hofmeister regarding the "salting out" of proteins by high concentrations of salt ions and their precipitation by careful adjustment of pH (Hofmeister 1888; Cohn 1925). The ideas behind this salting out procedure, responsible even today for the growth of more protein crystals than any other (see Chapter 5), were embodied in analytical form by Cohn and Ferry in 1943. Their equation (see Chapter 4) provides one of the most fundamental conceptual tools for protein crystals growers.

According to Reichert and Brown (1909), by 1910, only a few proteins were known to have been crystallized. They state, "thus far, only a very

limited number of the proteins have been obtained in crystalline form. A number of hemoglobins and hemoglobin compounds and derivatives, albumin, lactalbumin, casein, vitellin, a number of globulins from seeds and nuts, the albumin and globulin of egg white, hyalin, two proteins from abnormal urines, ichthulin from the eggs of fish, glutokyrin, hemocyanin, and phycoerythrin and phycocyanin of algea, include all, as far as we have been able to find, that have been obtained in crystal."

THE WORK OF J.B. SUMNER

Ignoring the conventional wisdom of his day, J.B. Sumner initiated efforts early in the 20th century to crystallize an enzyme. This would, he believed, demonstrate unequivocally that enzymes were proteins having unique and defined molecular structures. The principal focus of his attention was the enzyme urease, known to be found in significant quantities in a leguminous seed, the jack bean (*Canavalis ensiformis*).

Although he did not realize success in his quest until the mid 1920s, in 1919 he reported the crystallization of two more proteins, concanavalin A and concanavalin B. The former of these was the first plant lectin ever crystallized. Interestingly, we now know that concanavalin B (Morrison et al. 1984) is an enzyme, chitinase, but in an inactive form. Although unrecognized at the time, because the function was unknown, concanavalin B was, in fact, the first enzyme ever crystallized. Sumner had already achieved his objective in 1919 but was unaware of it. Some examples of crystals from these early investigations are seen in Figure 1.4.

In the course of his crystal studies, prior to 1920, Sumner also reported the isolation of a third globulin from the jack bean, but only in amorphous form. This protein, which was the major reserve protein of the seed, he named canavalin. Although Sumner was unable to crystallize canavalin, he did succeed in crystallizing urease in 1925 and demonstrated his point (Sumner 1926). For this research, Sumner shared the Nobel Prize in 1946 with W.M. Stanley, who crystallized TMV, and J.H. Northrop, who crystallized the enzyme pepsin. This remained the only Nobel Prize awarded to cultivators of protein crystals until H. Michel received the prize in 1988 for the first crystalline membrane protein.

A fine but somewhat less honored scientist, Stacey Howell, a student of Sumner, was in the 1930s given the task of further investigating an old nemesis of his mentor, canavalin, the 1919 amorphous precipitate. During the period of a week or so Howell left a solution of his jack bean canavalin in a beaker on his laboratory bench while he pursued other interests. Upon