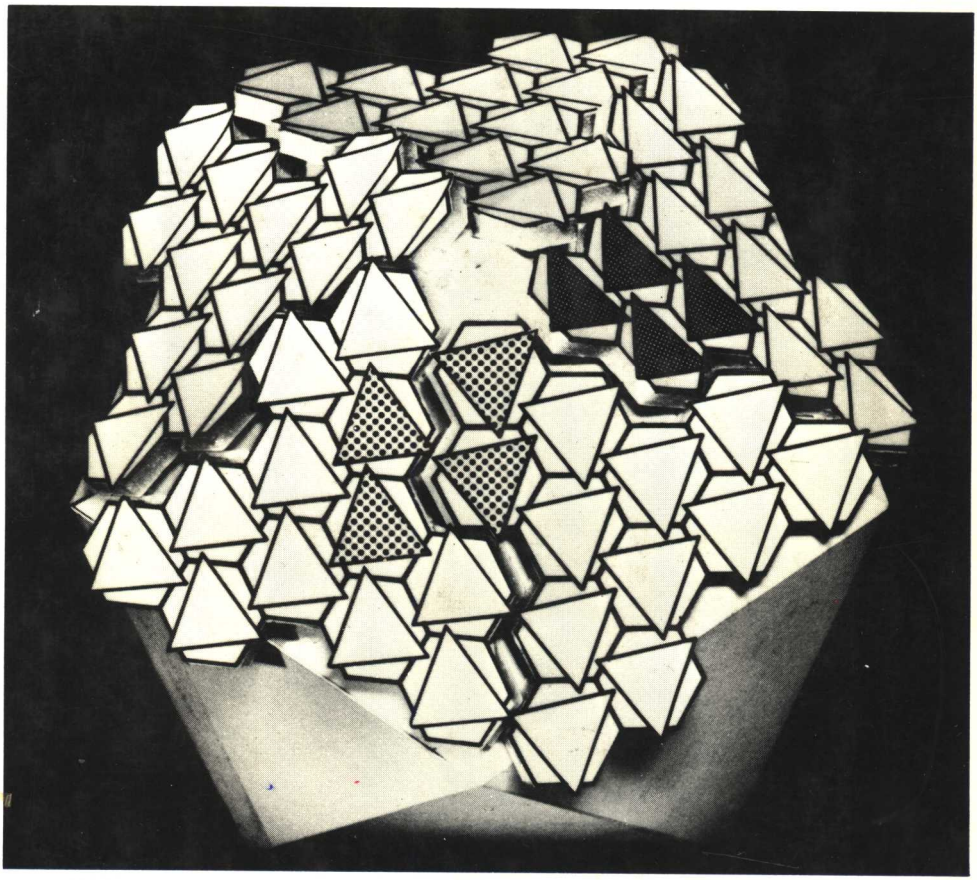
BIOLOGICAL MACROMOLECULES & ASSEMBLIES



VOLUME 1: VIRUS STRUCTURES

Edited by
Frances A. Jurnak & Alexander McPherson

BIOLOGICAL MACROMOLECULES AND ASSEMBLIES

Volume 1: Virus Structures

Editors

Frances A. Jurnak and Alexander McPherson

University of California
Riverside

A Wiley-Interscience Publication

JOHN WILEY & SONS

New York Chichester Brisbane Toronto Singapore

Library of Congress Cataloging in Publication Data:

Main entry under title:

Biological macromolecules and assemblies.

"A Wiley-Interscience publication."
Contents: v. 1. Virus structures.
Includes indexes.
1. Macromolecules—Collected works. I. Jurnak, Frances A. II. McPherson, Alexander, 1944- . [DNLM: 1. Macromolecular systems. QD 381 B615]

QP801.P64B554 1984 574.19'24 83-21732
ISBN 0-471-87077-3 (v. 1)

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Preface

This volume presents what may well be the most impressive achievements of macromolecular analysis by X-ray diffraction methods, the structures of the viruses. Represented here are examples of icosahedral or spherical plant viruses, helical plant viruses, filamentous phages, and initial investigations of the more complex icosahedral animal viruses. In addition to studies of the intact viruses, two examples are included of high resolution analyses of the structures of proteins found on the surfaces of an adenovirus and the influenza virus. These lend significant new insights into the architectural principles involved in the design and assembly of viral particles and interactions with their hosts. The investigations described here establish a sound structural basis for interpretation of the vast accumulation of data involving viruses from other sources of biochemical, physical, and physiological research. They provide the most convincing confirmation of now widely accepted virus structure theory but at the same time make evident the likelihood that these principles will see substantial alteration and refinement as larger and more complex structures become available. A not insignificant feature of these achievements is that they serve as inspiring testament to the perseverance and ingenuity of diffractionists working at the leading edge of their science with success never assured, methods marginally adequate to the task, and specimens possessing the most complex and elusive properties.

Chapter 1 is an overview of the chemical and physical properties of the spherical plant viruses and their implications for general virus structure. As such, it establishes conditions and standards by which the structure analyses must be measured. Chapters 2 and 3 focus on two small icosahedral plant viruses whose structures have now been determined in near atomic detail, southern bean mosaic virus and satellite tobacco ne-

crosis virus. These two structures, along with that of tomato bushy stunt virus, elegantly illustrate the principles of virus construction postulated by Caspar and Klug but at the same time suggest alternative bonding mechanisms by which the postulated quasi-equivalence might be achieved. Further, they beautifully illustrate the integration of protein structural principles into larger macromolecular assemblies, the economy of structure and function, and the consistency of design.

Tobacco mosaic virus, described in Chapters 4 and 5, represents a different class entirely, the viruses constructed according to precise helical symmetry. The structure of tobacco mosaic virus, the most exhaustively studied of all the viruses, illustrates the means by which small, identical protein subunits of relatively uncomplicated structure have encoded within their three-dimensional array of amino acids, the ability to recognize and bind their genomic nucleic acid, to order and protect it, and to organize themselves into a complete viral assembly with new and sophisticated properties.

In Chapter 6, the traditional crystallographic technique of isomorphous replacement is by necessity abandoned in the study of the filamentous bacteriophages. Using primarily physical constraints in conjunction with fiber diffraction data, visual images have been achieved that reveal the interaction between small α-helical subunits and the DNA genome at the core of the virus. Unlike the icosahedral and helical viruses, these bacteriophages are designed according to subtly different principles of symmetry and nonequivalence of interaction. They suggest that there may exist many variations or entirely new patterns of self-assembly in viruses.

Chapter 7 describes the low resolution analysis of an animal virus, the polyoma virus. Unexpectedly—perhaps because it so faithfully assumed the principles of Casper and Klug regarding icosahedral symmetry—the results of this study have suggested the most radical departure from those principles. Once again, the methodology was somewhat unorthodox in that it did not employ isomorphous replacement; nevertheless, it appears valid by current criteria. If the results are confirmed by other investigations or by higher resolution analyses of the polyoma virus itself, they will certainly require a reconsideration of our expectations for capsid structure and assembly in the $T = 7$ and larger virus particles.

Chapters 8 and 9 are elegant and impressive examples of traditional isomorphous replacement methods coupled with the clever application of molecular replacement and symmetry constraints. The structures emerg-

ing from these investigations, the capsid protein hexon from adenovirus and the influenza virus hemagglutinin, are striking in the manner by which recognized protein secondary structural patterns are extended and reorganized to produce unusual and unique protein molecules. These are molecules that not only contribute to our understanding of general protein structure but bring to light new principles and mechanisms for macromolecular interaction, assembly, and physiological activity.

The nine chapters presented here illustrate and describe in a reasonably comprehensive fashion what is currently known of virus structure at the atomic and near atomic level. The structures were derived principally by X-ray diffraction and the perspectives primarily are those of scientists versed in that technique. This should in no way be interpreted as ignoring the extensive achievements resulting from the many other methods that have been applied to virus structure analysis. Indeed, it should be noted that immediately following or closely entwined with the description of any structure is the correlation or rationalization of each feature in terms of benchmarks derived from biochemical and physical analyses. Perhaps the crystallographers may here repay their debt to the scientific community at large by suggesting new approaches and new experiments and at the same time provide a structural basis for their design and interpretation.

FRANCES A. JURNAK
ALEXANDER MCPHERSON

Riverside, California
March 1984

Contents

1

Chemical Stability in Simple Spherical Plant Viruses

PATRICK ARGOS
JOHN E. JOHNSON
Department of Biological Sciences
Purdue University
West Lafayette, Indiana

CONTENTS

1. OVERVIEW

1.1. Introduction

The relationship between molecular structure and function is a primary concern in molecular biology. The simple spherical plant viruses have been the subject of such study since the early 1950s. It was only in the 1970s, however, that an investigative level was reached that permits analysis of this relationship at atomic resolution. The tertiary structures of three spherical plant viruses and a helical plant virus have yielded substantial knowledge of subunit association and suggested the nature of protein–RNA interactions. The functional aspects of chemical stabilization and its relation to assembly processes have been exhaustively reviewed by Kaper (1975) through 1974. Since then, studies similar to those reported by Kaper have been extended and new methods of analysis have been brought to bear on the virus problem. These include neutron diffraction, nuclear magnetic resonance spectroscopy, circular dichroism, and primary structure analysis of viral proteins and nucleic acids.

The goal of this chapter is to describe plant virus structure and assembly processes established through biochemical and biophysical methods other than single crystal X-ray diffraction techniques. The viral systems chosen for discussion are taken from the simple spherical plant viruses, which have undergone extensive chemical study. The rod-shaped tobacco mosaic virus has been reviewed recently and exhaustively (Hirth and Richards, 1981). The first part of this chapter provides a comparative analysis of the virus systems; the second presents the detailed experimental evidence and related references.

1.2. Capsid Protein

Table 1 lists the gross structural parameters for the type members of various virus groups. In the native state all except alfalfa mosaic virus (AMV) are icosahedral with hexamer-pentamer clustering dominating the capsomeric distribution. With the exception of satellite tobacco necrosis virus (STNV), cowpea mosaic virus (CpMV), and AMV, all display a triangulation number of 3 and contain 180 identical subunits in the protein shell. The $T = 3$ viruses are approximately the same size ranging from 260 to 310 Å in diameter. CpMV particles fall in this range despite their

Table 1. Physical and Chemical Properties of Selected Plant Viruses

	Particle					
	Diameter (Å)	Triangulation Number	Sedimentation Coefficient (S)	State of Structure Determination	Multipartite	*In Vivo* Capsids
TYMV	280	3	54 (T) 117	Crystallized[a] 40 Å	No	Yes
CpMV	280	1	58 (T) 95 (M) 115 (B)	Crystallized	Yes	Yes
STNV	176	1	50	3.0 Å	No	No
SBMV	284	3	115	2.8 Å	No	No
TBSV	308	3	131	2.9 Å	No	No
BMV	260	3	87	None[b]	Yes	No
CMV	290	3	99	Crystallized	Yes	No
AMV	190[c]	1	94[d] 82 73 66	4.5 Å[c]	Yes	No

	Protein					
	Molecular Weight	Protein Sequence	Basic Amino Terminus	Tertiary Structure	State of Soluble Disassembled Protein	Polymorphism
TYMV	20,133	Known	No	?	?	No
CpMV	22,000	Known	?	?	?	No
	42,000	Known				
STNV	21,600	Known	Yes	β barrel	?	No
SBMV	28,200	Known	Yes	β barrel	dimer	Yes
						$T = 1$, $T = 3$
TBSV	40,000	Unknown	Yes	β barrel	dimer	Yes[e]
						$T = 1$, $T = 3$
BMV	20,300	Known	Yes	?	dimer	Yes
						$T = 1$, $T = 3$, tubes, shells
CMV	26,200	Known	Yes	?	?	No
AMV	24,240	Known	Yes	β barrel	dimer	Yes
	(21,100)[e]					$T = 1$

Table 1. (*Continued*)

	RNA					
	Subunit Molecular Weight ($\times 10^6$)	3′ Terminus	5′ Terminus	Sequence	Secondary Structure	Subgenomic Messenger
TYMV	2.0	tRNA-like valine chargeable	7-Methyl-guanosine cap	Partial	~60%	Yes
CpMV	1.39	poly (A)	Protein bound	Partial	>60%	No
	2.39	poly (A)	Protein bound	Partial	>60%	
STNV	0.4	No special structure	No special structure	Known	?	Yes
SBMV	1.2	No special structure	Protein bound	Partial	~60%	Yes
TBSV	1.5	?	?	No	?	Yes
BMV	1.1	t-RNA like tyrosine chargeable	7-Methyl-guanosine cap	Partial	~60%	Yes
	1.0			Partial	~60%	
	0.7			Known	~60%	
	0.3			Known	~60%	
CMV	1.27	t-RNA like tyrosine chargeable	7-Methyl-guanosine cap	Partial	?	Yes
	1.13			Partial	?	
	0.82			Known	?	
	0.35			Known	?	
AMV	1.04	All 3′ are similar but not chargeable and no poly (A)	7-Methyl-guanosine cap	Partial	?	Yes
	0.73			Partial	?	
	0.62			Partial	?	
	0.28			Partial	?	

Stability and Assembly

	Swelling	Primary Stabilizing Forces				Sensitivity to RNase	Reversible Disassociation
		Hydrophobic	pH	Metal	Protein–RNA		
TYMV	?	Yes	No	No	No	No	No
CpMV	?	Yes	No	No	Yes	No	?
STNV	Yes	No	Yes	Yes	Yes	No	?
SBMV	Yes	No	Yes	Yes	Yes	No	Yes
						(swollen yes)	
TBSV	Yes	No	Yes	Yes	Yes	No	?
						(swollen yes)	
BMV	Yes	No	Yes	Yes	Yes	No	Yes
						(swollen yes)	
CMV	No	No	No	No	Yes	Yes	Yes
AMV	No	No	No	No	Yes	Yes	Yes

[a] Belladonna mottle virus has been crystallized.
[b] Cowpea chlorotic mottle virus has been crystallized.
[c] A reassembled protein capsid.
[d] AMV contains four cylindrical components of different size and RNA content.
[e] TCV displays polymorphism.

$T = 1$ icosahedral symmetry. The thickness of the protein shell, approximately 40 Å, is also similar for the viruses. The structure, sequence data, and proteolytic cleavage patterns indicate that the N-terminal region of southern bean mosaic virus (SBMV), tomato bushy stunt virus (TBSV), STNV, AMV, cowpea chlorotic mottle virus (CCMV), and cucumber mosaic virus (CMV) is composed predominantly of basic residues associated with the nucleic acid. When the unassembled capsid protein is in solution, it generally exists as a dimer for all the viruses studied.

The capsid protein serves primarily a protective role. Table 1 lists the molecular weight and certain properties of the coat protein sequences. The molecular weights range from 19,600 for the bromoviruses to 44,000 for the large protein of CpMV. TBSV protein is the largest of the $T = 3$ viruses. The three-dimensional structures determined for TBSV, SBMV, and STNV reveal that a β-barrel formed by antiparallel strands is the predominant feature of their tertiary architectures. The molecular weight of the barrel is about 19,000 for all three viruses with any large variations accounted for by the length of the amino terminal arms (roughly 8000 for TBSV and SBMV and 3000 for STNV) and an additional barrel domain in TBSV (12,000). Circular dichroic studies of many of the other viral proteins suggest a large quantity of β-sheet; predictions of TYMV secondary structure indicate the β-fold. Many of the subunit molecular weights shown in Table 1 are near 20,000 daltons, suggesting a similar β-topology in their backbones. In the case of CpMV, which contains two polypeptide chains in the capsid, the large protein could consist of two domains similar to the basic β-fold and the small protein one such domain, resulting in a near $T = 3$ lattice of surface β-barrels.

1.3. Nucleic Acid

The RNA molecules for the spherical viruses fall into three size classes: 2.0×10^6 daltons or larger, $0.6–1.4 \times 10^6$ daltons, and less than 0.5×10^6 daltons. The smallest RNA molecules are likely monocistronic subgenomic species that are the functional messenger for the capsid protein. In AMV, CCMV, CMV, SBMV, and TYMV, the sequences for the coat protein are also present in a larger RNA; nonetheless, encapsidation of the subgenomic RNA occurs and yet is not essential for infection. The exception is CpMV, which does not generate a subgenomic messenger.

Many of the RNA molecules occupy only 20–30% of the volume available within the capsid, assuming a partial specific volume for RNA of

0.55 $Å^3$/dalton. The viruses containing a high percentage of RNA (CpMV bottom component and the TYMV group) have approximately half the capsid interior occupied.

The termini of the viral RNAs display a variety of chemical structures. The bromovirus and CMV molecules are capped with 7-methylguanosine at their 5′ termini, and their 3′ termini display substantial sequence homology and probably similar secondary structure. The 3′ termini can be charged with tyrosine in the presence of the appropriate tRNA synthetase. Tymovirus RNA can be charged with valine. SBMV and CpMV RNAs have proteins covalently linked to their 5′ termini. CpMV contains a sequence of polyadenylic acid at the 3′ terminus. SBMV contains 130 noncoding bases that do not fold into a particularly stable secondary structure. In fact, the bean and cowpea strains of SBMV show no sequence homology in this region. In contrast, STNV RNA does not display any of these various terminal structures. It seems probable that the diverse features of the RNA are closely related to the replication and translation strategy employed in their hosts.

The RNA molecules for most of the viruses have been the subject of *in vitro* translation studies. In general more than one protein is synthesized and the coat protein for all but the comoviruses is produced from the subgenomic messenger RNA. In all but STNV the RNA probably codes for a subunit of the RNA replicase molecule as well as the coat protein. CpMV RNA codes for a protease that cleaves a viral polyprotein into viable molecules.

The RNA is highly organized within the particle. Circular dichroism indicates at least 60% of the polynucleotide is ordered in secondary structures. Isolated SBMV RNA when stretched into a fiber gives an X-ray diffraction pattern consistent with A DNA, as expected for double helical RNA. While the RNA is highly organized, it is not necessarily in its most compact form. SBMV and TYMV RNA can undergo *in situ* rearrangements that make the nucleic acid more compact but less infectious. If the RNA is denatured, it regains full infectivity. Apparently the structure of encapsidated RNA provides a balance between minimum packing volume and rapid unfolding for processing.

A problem of RNA packaging at physiological pH is the cancellation of negative charge associated with phosphate groups. There are three sources of positive charge for neutralization: cations such as potassium or calcium, which are present in the cytoplasm, polyamines such as

spermine and spermidine, and basic residues of the N-terminal arm of the coat protein.

1.4. Capsid Stabilization

Strong protein–protein interactions fall into three categories: metal ion mediated, pH dependent, and hydrophobic interactions. Those involving metal ions are also pH dependent. SBMV, STNV, TBSV, and the bromoviruses show stability from pH 3.0 to 10.0 in the native state; however, swelling of the particles generally occurs above neutrality with metal ion removal by EDTA. The tertiary structures of SBMV, STNV, and TBSV show the metal ions are associated with acidic residues at the subunit interfaces. Deprotonation of these residues in the absence of metal ions results in a local buildup of charge causing the subunits to separate with protein–RNA interactions primarily maintaining the structure. All viruses that contain removable metal ions display this structural transition.

The CpMV group shows pH dependence in capsid stability, but metal ions do not play a significant role. Above pH 8.0 an alteration takes place in the RNA-free particles, making them susceptible to disassembly by low concentrations of denaturants while the nucleoprotein particles remain stable. TYMV particles, also stabilized by very strong protein–protein interactions, will, under conditions of high salt and pH, release RNA from a momentarily expanded structure. A stable empty capsid nonetheless remains after RNA expulsion. However, another tymovirus, belladonna mottle virus (BDMV), will release its RNA at pH 7.0 under only mild buffering conditions. All evidence points to extended regions of hydrophobic interactions between the subunits as the stabilizing force in CPMV and TYMV.

An N-terminal region not involved in RNA interactions is important in establishing the architecture of some virions. The SBMV and TBSV X-ray structures show the formation of a β annulus structure about the icosahedral threefold axes which are quasi-sixfold axes. Using the nomenclature of Harrison, three of the C subunits contribute to the hydrogen bonded structure whereas the quasi-equivalent A and B subunits do not as they display a disordered structure in this region. Thus, considering only the C subunits and the β annuli, a $T = 1$ structure is formed. The functional role of this region is clearly demonstrated in SBMV and TBSV in which the protein will not assemble into a $T = 3$ structure if the arm